Storage and Retrieval Techniques for Multimedia Data

(annotated bibliography)

by
P. Zezula and C. Galindo-Legaria
IEI-CNR, Via S. Maria 46
56126 Pisa, Italy

Abstract

The success of multimedia information systems to adequately meet the needs of accessing and presenting essential information from a large multimedia information space, depends heavily on the existence and proper application of storage techniques suitable for this task. This bibliography contains surveys of some papers which have recently appeared on this subject. The main emphasis is given to the modeling, content retrieval, and management of continuous multimedia objects. However, distributed algorithms and storage techniques for database mining and mobile wireless computing, i.e. the emerging applications which are closely related to the multimedia domain, are also briefly referenced.

1 Introduction

The existence of massive volumes of multimedia data makes the performance aspect of multimedia information management systems the critical implementation issue. Several very important application environments, such as digital video libraries or medical information systems, have claimed requirements for storage of many trillions of bytes of data. Because of the huge amount of data available, such environments naturally demand the use of a storage hierarchy, and a carefully optimized environment for data placement and retrieval of real-time, delay-sensitive, and synchronized multimedia data is necessary.

But, it is not only the large volume and the real-time characteristics of the multimedia data which make the performance problem complex. Another serious performance related problem concerns efficiency in which the information, a user is looking for, is provided to him. This demands not only the use of powerful, uniform, and extensible models for describing the concepts stored in the information base, but also index structure extensions for effective and efficient content-based retrieval of information.

Even though the area of storage techniques for multimedia data is quite new, the available literature is already extensive. What makes the situation even more complex is that the boundaries of this topic are not very well defined, because the current research in such areas as distributed algorithms, parallel data processing, mobile wireless computing, database mining, and computer networks, does consider multimedia data as well.

The following annotated bibliography should be considered as a summary of literature which the authors of this paper have studied in order to enter this new, interesting, and promising area of research. For this reasons, there are no claims for completeness or optimal structuring of the bibliography. But, since we felt that such text might be useful also for other people interested in this subject, we have decided to share the knowledge we have learned.

2 Multimedia data modeling

Data modeling is essential to the design and maintenance of information systems, as it provides the conceptual structure for the interpretation and use of stored data. Recent work has started to address the issue of modeling of multimedia data, whose *continuous* nature cannot be handled appropriately by traditional database modeling tools. This media continuity is made manifest in two senses. First, time flow is central to the meaning of items, which have a specific duration and whose relationships are heavily time-dependent —e. g. sound is recorded on a CD at 44100 samples per second, and its reproduction must be coordinated in time with a video signal in case the sound consists of an explanation of the video images. Second, items are not always decomposable into discrete, atomic units, from a conceptual point of view —e. g. a digital video signal may be composed of a sequence of bits, but only subsequences, images, or image fragments may make sense at the conceptual level; there are no natural "primitive components."

A collection of high-level modeling primitives applicable, in principle, to different time-based data is presented in [22]. The theme of the paper is that databases should handle multimedia data not as BLOBs (i. e. large sequences of uninterpreted bits), but as *streams*, with explicit modeling of time, derivation, and composition dependencies between items.

An modeling technique for video/audio data based on film theory is presented in [26]. The paper adapts known concepts from film analysis, emphasizing the meaningful composition of elements in terms of a narrative. It also shows the difficulty of building a unique conceptual structure for a multimedia item, whose human interpretation is open-ended in comparison with the highly schematic tuple of string/integer values traditionally used in databases.

A set of n-ary relation primitives to describe time-dependencies between multimedia items is presented in [32]. Dependencies modeled include simultaneity of sequentiality, for example. A logical interpretation of the primitives is described, and algorithms to create and traverse this representation are given.

There is relatively little theoretical work on multimedia databases. Some initial work on databases for geometrical information is presented in [36]. The paper tries to adapt the basic theoretical concept of "genericity" of relational databases to the case where data does have some intrinsic semantics attached to it (geometrical, in this case).

summary:

- modeling of time-based media [22]
- video information contents and film theory [26]
- interval-based models for time-dependent multimedia data [32]
- theory of spatial database queries [36]

3 Content-based retrieval

Access structures which support efficient content-based retrieval of multimedia objects, form one of the most important parts of multimedia information management systems. One important stream of research effort in developing indexing structures for spatial and temporal data is reviewed in [41]. In the following, we briefly survey some recent techniques suitable for image and text data indexing.

3.1 Image and space retrieval

The problem of multi-dimensional range searching in both main memory and secondary memory has been the subject of much research. Many elegant data structures (such as the priority tree, segment tree, and interval tree) have been proposed for use in main memory for special cases of 2-dimensional range searching. The path caching technique, proposed in [37], is able to convert many of these in-core data structures into secondary storage structures. These new data structures have optimal query response time at the expense of small storage overhead.

When points are stored in a spatial access method, such as an R-tree, an important problem for a query optimizer is the estimation of the search effort for executing range queries. Traditional assumptions for such methods are the uniformity and independence, which make the analysis tractable, but which do not reflect the reality where the data are indeed skewed. A formula that estimates the number of disk accesses for range queries for R-trees, given only the fractal dimension of the point set and its count, is developed in [17]. Experiment on real data sets show that the formula is very accurate: the relative error is usually below 5%, and it rarely exceeds 10%.

The content-based image indexing problem is formulated in [15] as a multi-dimensional nearest-neighbor search problem. An optimistic vantage-point tree algorithm is developed that can dynamically adapt the indexed search process to the characteristics of given queries. The performance study indicates that the system typically needs to touch less than 20% of the index entries for a query.

summary:

- skewed space data distribution [17]

- path caching (external 2-dimensional searching) [37]
- content-based image indexing [15]
- indexing spatial and temporal data [41]

3.2 Text retrieval

The application of database technology is seen essential to the operation of a conventional business enterprise. However, there is a universe of business information, namely text, which is currently stored, accessed, and manipulated in an add hoc fashion with none of the consistency and discipline of the database approach. A tutorial on theory, practice and experience in text dominated databases can be found in [23].

Free text retrieval through signature files is an important indexing technique which can significantly benefit from a parallel architecture. A parallel bit-sliced signature file method is proposed in [35], which can index very large files where the size of the signature file exceeds the available (parallel) main memory. The method uses a partial fetch slice swapping algorithm, thus accesses only the necessary number of bit slices. Arithmetic examples show that the method is able to handle a 128GB database with a 2sec response time on a machine with the characteristics of the Connection Machine.

The fundamental problem in the signature construction is to find the optimum signature weight to minimize the false drop probability. Even though the problem was extensively studied in the past, a new formula for computing the false drop probability is presented in [44]. The formula is easier to analyze so that optimal solutions can be more adequately derived. Performance results show that these solutions are better than the solutions obtained before.

summary:

- text dominated databases [23]
- parallel bit-sliced signature files [35]
- performance analysis of superimposed signature files [44]

4 Management of multimedia data

Recent years have witnessed a significant increase in electronic management of information. In particular, many types of information that traditionally were considered to be analog, such as image and sound, are now processed and utilized in digital form. According to [9], this advancement comes with tremendous opportunities and challenges for information systems to better meet information users' needs to manipulate multimedia information in a natural and effective manner.

To embed various media into divers applications, multimedia information systems (MMIS), which allow for creation, processing, storage, management, retrieval, transfer, and presentation of multimedia information, have been proposed. These systems resulted from the synergism of many different fields, including databases, digital signal processing, image processing, optical communication, mass storage, artificial intelligence, new paradigms for programming languages, and text processing.

SuiteSound, [40], is a tool for manipulating multimedia flows, such as digital audio. It supports development of collaborative applications using multimedia by integrating flexible, easy-to-use digital audio with a flexible, easy-to-use object-based system. SuiteSound applications readily combine live or recorded audio with editable representations of user objects.

A low-cost storage architecture for a movie on demand (MOD) server that relies principally on disks is presented in [34]. The high bandwidths of disks in conjunction with a clever strategy for striping movies on them is utilized in order to enable simultaneous access and transmission of different portions of a movie.
summary:

- multimedia information systems [9]
- a system for distributed collaborative multimedia [40]
- storage server for a movie on demand [34]

4.1 Continuous media management

Though the basic storage support for multimedia data types may indeed be the same as for conventional database types, such as records, the kinds of data manipulations and application interfaces needed, especially for continuous media types, are vastly different from those found in conventional data processing. According to [33], addressing these requires significant changes in storage system architectures, both for database management systems and for the underlying operating system and network facilities.

For continuous media, the mode of delivery of data is all important. The human eye, and more so the human ear, are sensitive to shifts and skips in the time domain, so rates of presentation are significant. There are also synchronization demands, because most continuous media presentations combine more than one data stream. The data storage volumes demand tertiary storage, which introduces seconds of latency on some accesses. Data rates are high enough, relative to hardware capabilities, that buffering of data in the storage system or application memory is only suitable for overcoming very small latencies or time shifts. In fact, getting data too early becomes as much a problem as getting data too late, so the schedule with which data is delivered to an application becomes critical.

A study of the ways, in which storage system architectures must change in order to provide the constrained-latency storage access on continuous media (taking into account operating system and network support as well as database management) is presented in [33]. The problems of temporal synchronization of various data streams in multimedia information, exchanged between users over a high speed network, are addressed in [39].

Acme, [6], is a network server for digital audio and video I/O. It lets users specify their synchronization requirements through an abstraction called a logical time system.

The Continuous Media File System, CMFS [7], supports real-time storage and retrieval of continuous media data (digital audio and video) on a disk. CMFS clients read or write files in "sessions," each with a guarantied minimum data rate. Multiple sessions, perhaps with different rates, and non-real-time access can proceed concurrently. Issues, such as real-time semantics of sessions, disk layout, an acceptance test for new sessions, and disk scheduling policy are also addressed in [7].

The problems which appear in integrating storage and transmission of multimedia data with computing are discussed in [38]. The design of a high-performance multimedia storage server that addresses the above complex design issues, is the subject of this paper.

Physical storage organizations for time-dependent multimedia data are also studied in [11]. A model is derived to relate disk characteristics to different media recording and playback rates in order to derive their storage patterns. These storage organizations guarantee that as long as a multimedia delivery process is running, "starvation" never appears.

Optical disks are among the most promising secondary storage devices for multimedia data. The placement of multimedia data on optical disk is of primary importance because this data is being extracted from the optical disk in real-time (e.g. real-time for audio data means that system hiccups greater than 30 msec cannot be tolerated). Methods for an efficient placement of audio data on optical disks for real-time applications are developed in [48].
summary

- audio data placement on optical disks for real-time applications [48]
- I/O server and its synchronization mechanism [6]
- synchronization of multimedia data streams [39]
- a file system for continuous media [7]
- storage techniques for continuous multimedia [38]
- organizations for time-dependent multimedia data [11]
- storage system architectures for continuous media data [33]

4.2 Distribution, replication, and storage hierarchies

The special requirements of multimedia data require novel hardware configurations, and algorithms tuned for such configurations. First, the large storage requirements introduce the need of tertiary storage not as an infrequently used archive, but as an essential component of the normal operation of a system. Second, the strict rate-of-delivery requirements introduce, on one hand, the need of parallel disk access to increase bandwidth; and, on the other hand, the need of scheduling algorithms sensitive enough to deliver multiple streams of data with little variance on a per-stream specified rate-of-delivery, in order not to overflow the buffers of the reproduction device.

Various research results derived from a terrain-visualization project developed at the Lawrence Berkely Laboratory are reported in [45, 12, 13]. An overview of the system architecture, which declusters data over a high-speed, geographically distributed network is given in [45]. The architecture relies on ATM technology to merge multiple slow streams arriving from disks into a single combined stream of the appropriate bandwidth. Only off-the-shelf, relatively inexpensive hardware is required. Motivated by this project, [12] deals with the problem of how to optimize the time of retrieval of a set of partially replicated items —i. e. which replica to read of each item. Also, [13] deals with the optimization of allocation of object streams on sequential tertiary storage. Optimal allocation must balance the time required to mount, say a tape, with the time required to seek the required data within such tape.

A series of papers by Ghandeharizadeh and his colleagues examines in depth the alternative architectures and policies to deliver multiple streams of data retrieved from tertiary storage. [19] opens this series of articles by presenting the problem, and proposing that objects should be downloaded and replicated from tertiary storage into multiple disks dynamically, depending on the order of requests, free resources, and required bandwidth of each stream. Then, [20, 21] examine, using simulation, several policies for the replacement of data in the disk buffers, which are different from traditional buffer replacement policies due to timing constraints. Finally, [10] refines and improves the earlier work using striping.

A related paper, [14], describes the problems and some solutions introduced by the conceptually simple fast-forward replay, for browsing purposes. Modifying the rate-of-delivery affects architectures and policies designed under different assumptions. In addition, an appropriate solution has to be media-dependent: It has to take into account encoding standard (e. g. differential coding of frames in MPEG), and also a "reasonable" appearance to users.

summary:

- distributed parallel data storage systems [45]
- optimal response time for retrieval of replicated data [12]
- optimizing allocation of objects in tertiary storage [13]
- continuous retrieval of streams using parallelism [19]
- replicas in parallel multimedia information systems [20]

- hierarchical storage structures [21]
- staggered striping [10]
- video server fast-forward functionality [14]

5 New perspectives

New applications, such as the multimedia data processing, and new hardware/software technology are major factors to drive database research to new directions. A survey of the effects of up-to-date storage hardware technology on data structure and algorithms is presented in [30]. Specifically, the following memory technologies are discussed in this paper: high-speed DRAM, flash memory, dielectrical thin film memory, dual-port RAM, and content addressable memory.

5.1 Distributed algorithms

Data-intensive computer applications are posing ever-increasing demands in terms of storage and performance capacity. One cost-effective approach to meeting such requirements is to exploit distributed storage and computing resources in a client-server architecture. The most salient feature of this approach is that its communication overhead is largely independent of the number of servers and clients in the system; it is considered a scalable approach.

More precisely, the objective of scalability means the following. Starting with a system where one server manages a file of a specific size that is accessed by a specific number of clients at a specific rate, a scalable approach can efficiently manage a file that is n times bigger and accessed by n times more clients at the same per-client rate, by adding servers and distributing the file across these servers. Furthermore, the response time of the clients' requests should be as good as in the one-server case.

Distributed dynamic hashing, [16], extends the idea of dynamic hashing by using a novel autonomous location discovery algorithm that learns the bucket locations instead of using a centralized directory. Unlike many other works in this field, [46] considers explicitly the cost/performance ratio of the system by aiming to minimize the number of servers that are acquired to

provide the required performance. The advantage of the distributed search tree method, [31], is that it can perform well also for non-uniform data request distributions, such as the linear order of keys, nearest neighbour, or range queries.
summary:

- distributed dynamic hashing [16]
- scalable distributed file organization [46]
- distributed search tree [31]

5.2 Mobile wireless computing

In the mobile wireless computing environment of the future, a large number of users, equipped with low powered palmtop machines, will query databases over the wireless communication channels. However, the physical requirements of such environment make the problem of organizing wireless broadcast data different from data organization on the disk.

One specific problem is how to improve battery utilization because small palmtop units typically run on small AA batteries. Several indexing schemes are designed in [28, 29]. These papers also demonstrate that the proposed algorithms lead indeed to significant improvement of battery life while still retaining a low access time.

Palmtop based units will often be disconnected for prolonged periods of time due to the battery power saving measures, but palmtops will also frequently relocate between different cells and connect to different data servers at different times. Caching of frequently accessed data items will be an important technique that will reduce connection on the narrow bandwidth wireless channel. However, cache invalidation strategies will be severely affected by the disconnection and mobility of the clients. In fact, the server may no longer know which clients are currently residing under its cell and which of them are currently on. A taxonomy of different cache invalidation strategies is proposed in [8] and the impact of client's disconnection times on their performance analyzed.

Other way to reduce the expensive wireless communication is to apply data replication, which has been proposed and analyzed in [27].
summary:

- caching strategies [8]
- data replication [27]
- indexing on air [28, 29]

5.3 Database mining

Database mining refers to the efficient construction and verification of models of patterns embedded in large, typically multimedia, databases, and is emerging as a major application area for database technology. The problem is that current database systems do not provide the necessary functionality for all kind of data which have already been collected. There are many cases reported in which massive amount of data have been stored on tertiary storage, but are very slowly migrating to database systems. Quest, [3, 1], is the name of a project aiming at enhancing database technology to address this problem. Specifically, Quest focuses on three classes of database mining problems involving classification, association, and sequences.

The importance of sequence query processing is motivated in [42] where also a framework for the optimization of sequence queries based on several novel techniques is presented. An indexing method for processing similarity queries on time sequences is proposed in [2]. Time sequences are mapped to a frequency domain by considering the first few coefficients of the Discrete Fourier Transformation. The transformed sequences are organized as R-tree, and in this way, efficient answer to similarity queries is obtained. The method is generalized in [18] to answer approximate-match queries for subsequences of arbitrary length. Combinatorial pattern discovery or combinatorial data mining method of [47] is able to find structural or topological patterns in data that can lead to important conclusions or prediction of new phenomena.
summary:

- tutorial [1]
- text database discovery [24]
- combinatorial pattern discovery [47]
- subsequence matching [18, 42]

- similarity search in sequence databases [2]
- Quest project [3]

5.4 Databases in networks

Various types of large, complex, and expensive real-time computer systems contain a database engine as a critical component. These systems share some of common database issues with conventional applications, but they also exhibit rather unique characteristics that present challenging database issues. According to [4], major database issues for network management include choosing the right data model, handling two different kinds of data terms of integrity and recovery constraints, supporting temporal queries, satisfying real-time performance and high availability requirements, and several miscellaneous issues. Some of these issues have been investigated in various areas of database research stage. Advances in these areas that result in actual integrated implementations for data-intensive, real-time and temporal applications are eagerly awaited.

Research on distributed databases, multimedia databases, federated databases, will enable integration of databases to a greater or lesser extent, but does not contribute towards problems such as application interoperability and distributed system management [25]. For these functions, the research on distributed objects shows greatest promise.

summary:

- telecommunication networks [4]
- distributed objects [25]

References

[1] Agrawal, R., Tutorial on Database Mining. *Proceedings of ACM PODS'94*, ACM Symposium on Principles of Database Systems, Minneapolis, Minnesota, USA, May 1994, pp. 75–76.

[2] Agrawal, R., Faloutsos, C., and Swami, A., Efficient Similarity Search in Sequence Databases. *Lecture Notes in Computer Science*, No. 730, David B. Lomet (Ed.). Proceedings of FODO'93, Chicago, Illinois, USA, October 1993, Spriger-Verlag, pp. 69–84.

[3] Agrawal, R., Carey, M., Faloutsos, C., Ghosh, S., Houtsma, M., Imielinski, T., Iyer, B., Mahboob, A., Miranda, H., Srikant, R., and Swami, A., Quest: A Project on Database Mining. *Proceedings of the ACM SIGMOD'94*, Minneapolis, Minnesota, USA, May 1994, pp. 514-514.

[4] Ahn, I., Database Issues in Telecommunications Network Management. *Proceedings of the ACM SIGMOD'94*, Minneapolis, Minnesota, USA, May 1994, pp. 37–43.

[5] Ahn, I., Filtered Hashing. *Lecture Notes in Computer Science*, No. 730, David B. Lomet (Ed.). Proceedings of FODO'93, Chicago, Illinois, USA, October 1993, Spriger-Verlag, pp. 85–100.

[6] Anderson, D.P. and Homsy, G., A Continuous Media I/O Server and Its Synchronization Mechanism. *IEEE Computer*, October 1991, 24(10):51–57.

[7] Anderson, D.P., Osawa, Y., and Govindan, R., A File System for Continuous Media. *ACM Transactions on Computer Systems*, November 1992, 10(4):311–337.

[8] Barbara, D. and Imielinski, T., Sleepers and Workaholics: Caching Strategies in Mobile Environments. *Proceedings of the ACM SIGMOD'94*, Minneapolis, Minnesota, USA, May 1994, pp. 1–12.

[9] Berra, P.,B., Golshani, F., Mehrotra, R., and Liu-Sheng, O.R., Guest Editors' Introduction: Multimedia Information Systems. *IEEE Transactions on Knowledge and Data Engineering*, August 1993, 5(4):545–550.

[10] Berson, S. and Ghandeharizadeh, S., Staggered Striping in Multimedia Information Systems. *Proceedings of the ACM SIGMOD'94*, Minneapolis, Minnesota, USA, May 1994, pp. 79–90.

[11] Chen, H. and Little, T.D.C., Physical Storage Organizations for Time-Dependent Multimedia Data. *Lecture Notes in Computer Science*, No. 730, David B. Lomet (Ed.). Proceedings of FODO'93, Chicago, Illinois, USA, October 1993, Spriger-Verlag, pp. 19–34.

[12] Chen, L.T. and Rotem, D., Optimal Response Time Retrieval of Replicated Data. *Proceedings of ACM PODS'94*, ACM Symposium on Principles of Database Systems, Minneapolis, Minnesota, USA, May 1994, pp. 36–44.

[13] Chen, L.T. and Rotem, D., Optimizing Storage of objects on Mass Storage Systems with Robotic Devices. *Lecture Notes in Computer Science*, No. 779, M. Jarke, J. Bubenko, and K. Jeffery (Eds.). Proceedings of EDBT'94, Cambridge, United Kingdom, March 1994, Springer-Verlag, pp. 273–286.

[14] Chen, M.S., Kandlur, D.D., and Yu, P.S. Support to Fully Interactive Playout in a Disk-Array-Based Video Server. *Proceedings of the ACM Multimedia 94*, Second ACM International Conference on Multimedia, October 1994, San Francisco, California, pp. 391-398.

[15] Chiueh, T., Content-Based Image Indexing. *Proceedings of the 20th VLDB Conference*, Santiago, Chile, 1994, pp. 582–593.

[16] Devine, R., Design and Implementation of DDH: A Distributed Dynamic Hashing Algorithm. *Lecture Notes in Computer Science*, No. 730, David B. Lomet (Ed.). Proceedings of FODO'93, Chicago, Illinois, USA, October 1993, Spriger-Verlag, pp. 101–114.

[17] Faloutsos, C. and Kamel, I., Beyond Uniformity and Independence: Analysis of R- tree Using the Concept of Fractal Dimension. *Proceedings of ACM PODS'94*, ACM Symposium on Principles of Database Systems, Minneapolis, Minnesota, USA, May 1994, pp. 4–13.

[18] Faloutsos, C., Ranganathan, M., and Manolopoulos, Y., Fast Subsequence Matching in Time-Series Databases. *Proceedings of the ACM SIGMOD'94*, Minneapolis, Minnesota, USA, May 1994, pp. 419-429.

[19] Ghandeharizadeh, S. and Ramos, L., Continuous Retrieval of Multimedia Data Using Parallelism. *IEEE Transactions on Knowledge and Data Engineering*, August 1993, 5(4):658–669.

[20] Ghandeharizadeh, S. and Shahabi, C., Management of Physical Replicas in Parallel Multimedia Information Systems. *Lecture Notes in Computer Science*, No. 730, David B. Lomet (Ed.). Proceedings of FODO'93, Chicago, Illinois, USA, October 1993, Spriger-Verlag, pp. 51–68.

[21] Ghandeharizadeh, S. and Shahabi, C., On Multimedia Repositories, Personal Computers, and Hierarchical Storage Systems. *Proceedings of the ACM Multimedia 94*, Second ACM International Conference on Multimedia, October 1994, San Francisco, California, pp. 407-416.

[22] Gibbs, S., Breiteneder, C., and Tsichritzis, D., Data Modeling of Time-Based Media. *Proceedings of the ACM SIGMOD'94*, Minneapolis, Minnesota, USA, May 1994, pp. 91–101.

[23] Gonnet, G.H., Tutorial on Text Dominated Databases, Theory Practice and Experience. *Proceedings of ACM PODS'94*, ACM Symposium on Principles of Database Systems, Minneapolis, Minnesota, USA, May 1994, pp. 301–302.

[24] Gravano, L., Garcia-Molina, H., and Tomasic, A. The Effectiveness of GlOSS for Text Database Discovery Problem. *Proceedings of the ACM SIGMOD'94*, Minneapolis, Minnesota, USA, May 1994, pp. 126–137.

[25] Herbert, A., Databases in Distributed Systems: The New Frontier. *Lecture Notes in Computer Science*, No. 779, M. Jarke, J. Bubenko, and K. Jeffery (Eds.). Proceedings of EDBT'94, Cambridge, United Kingdom, March 1994, Springer-Verlag, pp. 2–6.

[26] Hjelsvold, R., Video Information Contents and Architecture. *Lecture Notes in Computer Science*, No. 779, M. Jarke, J. Bubenko, and K. Jeffery (Eds.). Proceedings of EDBT'94, Cambridge, United Kingdom, March 1994, Springer-Verlag, pp. 259–271.

[27] Huang, Y., Sistla, P., and Wolfson, O. Data Replication for Mobile Computers. *Proceedings of the ACM SIGMOD'94*, Minneapolis, Minnesota, USA, May 1994, pp. 13–24.

[28] Imielinski, T., Viswanathan, S., and Badrinath, B.R., Energy Efficient Indexing on Air. *Proceedings of the ACM SIGMOD'94*, Minneapolis, Minnesota, USA, May 1994, pp. 25–36.

[29] Imielinski, T., Viswanathan, S., and Badrinath, B.R., Power Efficient Filtering of Data on Air. *Lecture Notes in Computer Science*, No. 779, M. Jarke, J. Bubenko, and K. Jeffery (Eds.). Proceedings of EDBT'94, Cambridge, United Kingdom, March 1994, Springer-Verlag, pp. 245–258.

[30] Kambayashi, Y., Takakura, H., and Meki, S., Data Structure and Algorithms for New Hardware Technology. *Lecture Notes in Computer Science*, No. 730, David B. Lomet (Ed.). Proceedings of FODO'93, Chicago, Illinois, USA, October 1993, Spriger-Verlag, pp. 164–196.

[31] Kroll, B. and Widmayer, P. Distributing a Search Tree Among a Growing Number of Processors. *Proceedings of the ACM SIGMOD'94*, Minneapolis, Minnesota, USA, May 1994, pp. 265–276.

[32] Little, T.D.C. and Ghafoor, A., Interval-Based Conceptual Models for Time-Dependent Multimedia Data. *IEEE Transactions on Knowledge and Data Engineering*, August 1993, 5(4):551– 563.

[33] Maier, D., Walpole, J., and Staehli, R., Storage System Architectures for Continuous Media Data. *Lecture Notes in Computer Science*, No. 730, David B. Lomet (Ed.). Proceedings of FODO'93, Chicago, Illinois, USA, October 1993, Spriger-Verlag, pp. 1–18.

[34] Ozden, B., Biliris, A., Rastogi, R., and Silberschatz, A., A Low-Cost Storage Server for Movie on Demand Databases. *Proceedings of the 20th VLDB Conference*, Santiago, Chile, 1994, pp. 594–605.

[35] Panagopoulos, G. and Faloutsos, C., Bit-Sliced Signature Files for Very Large Text Databases on a Parallel Machine Architecture. *Lecture Notes in Computer Science*, No. 779, M. Jarke, J. Bubenko, and K. Jeffery

(Eds.). Proceedings of EDBT'94, Cambridge, United Kingdom, March 1994, Springer-Verlag, pp. 379–392.

[36] Paredaens, J., Van den Bussche, J., and Van Gucht, D., Towards a Theory of Spatial Database Queries. *Proceedings of ACM PODS'94*, ACM Symposium on Principles of Database Systems, Minneapolis, Minnesota, USA, May 1994, pp. 279–288.

[37] Ramaswamy, S. and Subramanian, S., Path Caching: A Technique for Optimal External Searching. *Proceedings of ACM PODS'94*, ACM Symposium on Principles of Database Systems, Minneapolis, Minnesota, USA, May 1994, pp. 25–35.

[38] Rangan, P.V. and Vin H.M., Efficient Storage Techniques for Digital Continuous Multimedia. *IEEE Transactions on Knowledge and Data Engineering*, August 1993, 5(4):564–573.

[39] Ravindran, K. and Bansal, V., Delay Compensation Protocols for Synchronization of Multimedia Data Streams. *IEEE Transactions on Knowledge and Data Engineering*, August 1993, 5(4):574–589.

[40] Riedl, J., Mashayekhi, V., Schnepf, J., Claypool, M., and Frankowski, D., SuiteSound: A System for Distributed Collaborative Multimedia. *IEEE Transactions on Knowledge and Data Engineering*, August 1993, 5(4):600–610.

[41] Salzberg, B., On Indexing Spatial and Temporal Data. *Information Systems*, 1994, 19(6):447–465.

[42] Seshadri, P., and Livny, M., and Ramakrishnan, R., Sequence Query Processing. *Proceedings of the ACM SIGMOD'94*, Minneapolis, Minnesota, USA, May 1994, pp. 430–441.

[43] Stone, H.S., Parallel Querying of Large Databases: A Case Study. *IEEE Computer*, October 1987, 20(10):11–21.

[44] Sung, S.Y., Performance Analysis of Superimposing-Coded Signature Files. *Lecture Notes in Computer Science*, No. 730, David B. Lomet (Ed.). Proceedings of FODO'93, Chicago, Illinois, USA, October 1993, Spriger-Verlag, pp. 115–129.

[45] Tierney, B., Johnston, W.E., Herzog, H., Hoo, G., Jin, G., Lee, J, Chen, L.T., and Rotem, D., Distributed Parallel Data Storage Systems: A Scalable Approach to High Speed Image Servers. *Proceedings of the ACM Multimedia 94*, Second ACM International Conference on Multimedia, October 1994, San Francisco, California, pp. 399-405.

[46] Vingralek, R., Breitbart, Y., and Weikum, G., Distributed File Organization with Scalable Cost/Performance. *Proceedings of the ACM SIGMOD'94*, Minneapolis, Minnesota, USA, May 1994, pp. 253–264.

[47] Wang, J.T., Chirn, G.W., Marr, T.G., Shapiro, B., Shasha, D., and Zhang, K., Combinatorial Pattern Discovery for Scientific Data: Some Preliminary Results. *Proceedings of the ACM SIGMOD'94*, Minneapolis, Minnesota, USA, May 1994, pp. 1–12.

[48] Yu, C., Sun, W., Bitton, D., Yang, Q., Bruno, R., and Tullis, J., Efficient Placement of Audio Data on Optical Disks for Real-Time Applications. *Communications of the ACM*, July 1989, 32(7):862–871.