MULTIMEDIA INFORMATION SYSTEMS

Maria Luisa Sapino
University of Torino

K. Selcuk Candan
Arizona State University

Synonym: *Multimedia systems, multimedia data management*
Definition: *Multimedia Information Systems (MIS)couple data management capabilities (effective and efficient methods for storage, indexing, querying, and retrieval) with media management (media representation, data compression, standardization, and transmission)*

Multimedia applications require representation and management of non traditional data, such as text documents, images, audio and video data, possibly together with traditional (e.g. relational or object oriented) data. In particular, in multimedia applications, different data types are often coexisting within the same application domain. Hence data management systems which are able to treat, in an integrated and uniform way, different types of data are needed.

Historically, multimedia systems and databases emerged as separate disciplines addressing the needs of different application domains. However, as these application domains matured over the years, the scientific disciplines also moved closer. For instance, on the media management side, although MPEG1, MPEG2, and MPEG4 standards were primarily focused on video compression and decompression tasks, researchers concentrated on the video-content description and indexing issues while developing the MPEG7 standard (which is currently in the works)[1]. On the data management side, commercial database management systems that were once primarily targeting traditional business applications, are today including media and Web engines within their packages. Applications such as digital libraries and dynamic Web content, necessitate a common understanding of both of these application domains. As a result, since the mid nineties, there has been a flurry of activities in the so called area of *multimedia information management systems*. These activities resulted in many techniques useful in extracting various features of media [56, 48] and providing means for indexing/querying [57, 80, 81, 82, 47] based on these features.

Multimedia information management systems, by definition, are not *dealing* with only a single media data type, but in fact need to be designed with the semantics and requirements of multiple media and modalities in mind. The heterogeneity of integrated-media based systems, such as multimedia document systems, have two main

---

[1] See *"Multimedia content description and standards"* by Alberto Messina and Maurizio Montagnuolo

reasons. First, the semantics of the information captured in different media types may be different. Consequently, constructing or interpreting a document that uses multiple media types is not straightforward. The lack of a predetermined and user-independent structure and semantics in many of the richer data types compounds this problem. A multimedia system requires intelligent information storage and delivery mechanisms to deal with such differences. Since the resource requirements of new media tend to be multiple orders larger than text, development of optimal data placement, storage, and retrieval techniques for newer media types is crucial for effective utilization of these resources.

Rather than being independent, the challenges associated with the semantic and resource heterogeneities of media are very much related and dependent on each other. For instance, once they are known, the semantics and structures of media objects can be used to develop optimal placement techniques for media components that will minimize the resource requirements. On the other hand, optimal retrieval techniques for multimedia information can only be developed keeping semantics, structures, and resource requirements of media types, media objects, and multi-media documents in mind. Since knowledge about users' needs, preferences, and profiles can help solving the semantic heterogeneity by putting the objects and documents within a context and since many of the resource-based challenges, such as data placement or buffer allocation, are most of the time user-oriented tasks, knowledge about users preferences and profiles can help the system tackle both semantic and resource-related challenges.

Multimedia objects are composed of media-objects of different data types, such as video, audio, text, and images. Unlike traditional documents, multimedia objects have temporal, spatial, and interactive properties. Thus a multimedia information system commonly requires two complementary components: (1) a multimedia specification/presentation sub-system that enables users to create and view multimedia objects, and (2) a storage, retrieval, communication component to maintain multimedia objects (as well as the individual media objects included within them).

Figure 1 depicts the outline of a multimedia information system (MIS) architecture, consisting of four sub-systems: A multimedia authoring system (MAS), a media sensing system (MSS), media processing system (MPS), multimedia communication system (MCS), a multimedia visualization and interaction system (MVIS), a multimedia object database (MODB), a profile and context manager (PCM), and a digital rights management system (DRMS).

- o MAS is responsible for authoring of multimedia objects. This involves describing visual, temporal, spatial, hierarchical, and interactive properties of a complex multimedia object.
- o MSS uses environmentally distributed sensing devices to collect multimedia data relevant for a given application.
- o MCS communicates media and multimedia objects between the various components of a multimedia information system, providing appropriate quality of service (QoS) guarantees.
- o Since row, sensed data is often not directly usable, MPS processes media objects for the benefit of the other components. For example, multiple media objects

may be fused to obtain a composite object, a media object may be processed to adjust its quality, or certain features may be extracted from a multimedia object for indexing purposes.

o MVIS benefits from the user specifications to create presentation schedules which maximize the utilization of available resources. It also enables the user to interact with the media objects and explore the multimedia information space. It performs document and media scaling to match the resource requirements to the resource availabilities. It also benefits from document structure, priorities, user preferences, and quality/cost trade-offs to develop object prefetching and caching strategies for document presentation.

o MODB is responsible for storage and retrieval of media objects and multimedia documents. MODB enables multimedia information to be queried and retrieved, efficienty and effectively. For this purpose it maintains appropriate index structures to support queries. Since multimedia retrieval is subjective, unlike traditional databases, MODB employs fuzzy or probabilistic query processing techniques and ranking algorithms to present the query results according to their relevance.

o PCM is responsible for keeping the user, context, and task profiles to improve the processing and presentation of multimedia information.

o DRSM ensures the intellectual property rights of the users who contribute multimedia objects into the multimedia information system by providing digital signatures and copy detection mechanisms.

In the rest, we describe the above components of a multimedia information system.

1. Multimedia Authoring System (MAS)
A multimedia information system needs to enable users to create multimedia objects by combining individual media objects as well as other multimedia objects. This process can be explicit (i.e., due the user provides the individual objects to be used) or declarative (the user provides specifications of the objects and the system instantiates the appropriate objects in run time)[2].

There are various multimedia authoring standards as well as commercial tools available for users. Standard Generalized Markup Language (SGML) is an international standard (ISO 8879) for describing the structure of documents have been accepted in 1986 [2]. The most essential feature of this standard is the separation of document content and structure from the presentation of the document. One of the most notable applications of the SGML standard is the HyperText Markup Language (HTML), the current standard for publishing on the Internet, which dates back 1992. Recently, Extensible Markup Language (XML) is developed by W3C Generic SGML Editorial ReviewBoard [3]. XML is a subset of SGML, especially suitable for creating interchangeable, structured Web documents.

---

[2] See *"Automatic generation of video documentaries"* by Stefano Bocconi

The Hypermedia/Time-based Structuring Language (HyTime) is another international multimedia standard (ISO 10744) [4], based on SGML. Unlike HTML or its derivatives, however, HyTime not only aims to describe the hierarchical and link structures of multimedia documents, but it also aims to describe temporal synchronization requirements between objects within a document. Synchronized Multimedia Integration Language (SMIL), also developed by the W3C [5, 6], provides a language for interactive multimedia presentations: authors can describe spatio/temporal properties of objects within a multimedia document and associate hyperlinks with them[3]. MHEG, a standard developed by the Multimedia and Hypermedia Experts Group [7], focuses on platform-independent interchange and presentation of multimedia presentations. The spatio-temporal relationships between objects and the interaction specifications form the structure of a multimedia presentation. Virtual Reality Modeling Language, VRML and its successor, X3D, are designed to provide standardized ways to describe interactive three dimensional (3D) information for web based applications and has soon evolved into the international standard for describing 3D content [8].

Unlike the above mentioned standards, which aim to describe the content of authored documents, MPEG7 (Multimedia Content Description Interface) aims to describe the content of captured media objects, such as video, along with authored content. It is a follow up to the previous MPEG standards, MPEG1, MPEG2, and MPEG4, which were mainly concerned on video compression.

2. Media Sensing System (MSS)
Media sensing systems range from high-end audio-visual equipment (such as microphone arrays, which can localize sources of audio signals in space, motion tracking devices, which can track objects in 3D space, and 3D-scanners which can create a precise 3D representation of an object) to low-end sensors (such as web-cams). The quality of the captured media may also depend on the environmental factors, such as the air quality, environmental noise, number of objects in the environment, or the mobility of the objects being recorded. Sensing devices can also be categorized as continuous, live devices which generate a stream of multimedia objects or as still devices, which takes a single on-demand recording of the environment.

A particular challenge with environmentally embedded, wireless live media sensing devices (such as motes) is that continuously capturing and communicating media to processing devices may require significant amount of resources taxing the limited battery power of many such devices. This requires media sensing systems to be sufficiently intelligent to be able to adapt their capture and communicate rates to their resources and criticality of the tasks.

3. Media Processing System (MPS)
Media processing system operates on the captured media to make it useable within the context of an application.

---

[3] See"*Interactivity in multimedia documents and systems*" by Ombretta Gaggi

Media processing includes media transformations, such as sharpening, blurring, rotating cutting, resizing, and warping, which transforms a media object without changing its type. Other media transformations may take one media as input and return and other media as its output. This includes, analysis operations (such as video shot detection and key-frame extraction) and feature extraction operations (including color histogram extraction, texture extraction, and object detection and tracking). Media fusion operators takes multiple media as input to create one single output. These include stereoscopic depth detection, panaromic image and splice generation through multiple images, and media enhancement and super-resolution.Other media processing tasks include lossy or lossless compression for efficient meia communication and watermaking for digital rights management.

Media processing workflows describe how a media processing task can be performed by combining various available operators, such as sensors, filters, fusion operators, communication operators, and environmental actuators. A distributed middleware, such as ARIA, support deployment of media processing workflows to capture, stream, and process various audio, video, and motion data in such sensory/reactive environments[4].

4. Multimedia Communication System (MCS)

Multimedia objects, by their nature, are large, heterogeneous, and in many cases, streaming. Consequently, multimedia communication systems must provide a large spectrum of services to support diverse applications, from telephony applications that require 10-20Kbits per second to gigabit applications, such as high definition television. The heterogeneity of the multimedia data streams necessitate further protection against bursts in the data volume. Thus the components and networking protocols must be able to scale to accommodate the heterogeneous volume of communication between the various components of multimedia information systems.

Since many multimedia applications require synchronization of streaming data at the receiving end, multimedia networks need to provide certain quality of service (QoS) guarantees, including low end-to-end latency limits in the data loss due to network bottlenecks or data compression, and protections against large jitter (i.e., variance of media transmission delays which may be perceived as the user as quality degradation). Thus, multimedia communication protocols need to be more informed about the characteristics and requirements of multimedia applications (including how reductions in network QoS reflect onto the users' perception of application quality), as well as the current state of the communication network. For example DiffServ (Differentiated Services) relies on cooperation of network routers to provide different QoS to different applications. The real-time transport protocol (RTP), which provides efficient (low network overhead) transport mechanism for streaming media data, such as audio and video, captures the payload type and its encoding scheme to inform the network components regarding the media data type being transmitted. The real time control protocol (RTCP) monitors QoS of RTP streams to inform the receiver and the sender regarding the quality of the media communication between each other.

---

[4] See "*Streaming multimedia information processing*" by K. Selçuk Candan

Since streaming multimedia communications require long term commitments from the underlying network, resource reservation protocols (such as RSVP) let multimedia applications to identify and reserve network resources. Other protocols rely on virtual private network (VPN) technologies to ensure protection of the QoS service of ongoing multimedia communications.

In many multimedia applications, the same multimedia object is consumed by multiple clients simultaneously. This gives rise to multicasting and broadcasting protocols which aim to preserve network resources by reducing redundant communications. While some of these protocols, such as IP-Multicast, operates at the network level others are implemented at the application level using overlay networks.

Some other multimedia applications (such as sensor network) require collection and integration of media streams from multiple sources and their delivery to a single receiver (such as a control center or a database). These applications require converge-cast protocols to reduce redundant communications from the sources to the destination. In many sensing applications, network nodes may filter and fuse incoming data before passing it along to the next node to reduce communication requirements..

5. Multimedia Visualization and Interaction System (MVIS)
The multimedia visualization system uses the presentation and media content specifications to visualize complex multimedia objects. It locates the media objects to be included within the presentation, creates a presentation schedule which does not conflict with the available resources, and handles user interaction.

If the specifications include explicit UID of a media object, then the system can make a directory lookup to locate it. In case the object is replicated, however, the most suitable, in terms of resource requirements and quality, object has to be located with the minimum communication between clients and servers. If the user provides a description of the media object, on the other hand, the system must identify a set of matching objects and choose one or more among them. To locate the matching media objects, the presentation system commonly uses the query processing and query refinement services provided by the underlying MODB. The challenge is to identify the media and multimedia objects, among the set of candidates, that fit into the presentation specifications the best.

Presentation of a temporal multimedia objects requires creation of presentation and object retrieval schedules that satisfy the resource limitations. It is the responsibility of the scheduler to identify the best way to fetch the media objects to be included in the presentation. The scheduler uses flexibilities in the specification, the context, and the user preferences to identify retrieval schedules that fit into the available resources. If the presentation specifications are not flexible, then using lesser matching media objects, also with less resource requirements, or using lower quality (e.g. lossily compressed) versions of the objects in the presentation may be acceptable if these replacements do not cause unacceptable visual and semantic degradations.

Many of the resources may need to be pre-negotiated between the presentation system and the resource brokers to guarantee certain quality of service. Resource negotiator performs one-to-many (one client, many resources) and, if there are multiple collaborators using the same resources, many-to-many (many clients, many resources) negotiations to optimize the collaborative resource utilization.

Users may interact with the multimedia information system in one of three ways: (1) change of the presentation parameters (e.g., speed, direction), (2) change of the presentation path (e.g., use of hyperlinks), and (3) interaction with the media objects within the multimedia presentation (e.g., opening a door in a 3D virtual world). Such user interactions may require the update of the presentation specifications. Furthermore, the scheduler may need to change the retrieval schedules. On the other hand, many multimedia applications require that the presentation is updated in less than 100 milliseconds to achieve true interactivity. To achieve this goal, the interaction manager collaborates with the scheduler, which creates new schedules incrementally, and the prefetch/cache manager.

The prefetch/cache manager retrieves multimedia objects that are likely to be viewed by the users into the unused resources or prefetch/cache buffers. To identify these objects, it uses the multimedia object structure, object priorities, statistics, and the user profiles. Since every missed object requires the scheduler to reshape the document or fill the gap with a replacement, each object is assigned a miss-penalty depending on how important it is for the structure of the document or for the viewer. Also, since each prefetched but unused object is a waste of valuable resources, each object is assigned a waste-penalty depending on the amount of resources it occupies. Using these penalties, the document structure, and the user interaction profiles, the prefetch/cache manager identifies the most desirable objects and retrieves them. Also, if an object is likely to be re-referred, the prefetch/cache manager maintains the object in the cache buffers for future use.

6. Multimedia Object Database (MODB)

MODB allows users to query for multimedia objects, multimedia documents, and their contents[5]. Due to the possibly redundant ways to sense the environment, the alternative ways to process, filter, and fuse multimedia data, and the subjectivity involved in the interpretation of data and query results, multimedia data quality is inherently imprecise. Feature extraction algorithms that form the basis of content-based multimedia data querying are generally imprecise. For example, high error rate is encountered in motion capture data due to the multitude of environmental factors involved, including camera and object speed. It is also rare that a multimedia querying system relies on exact object matching. Instead, in many cases, multimedia databases leverage similarities between features to identify data objects that are similar to the query. Due to the sheer size of the data, many systems rely on imperfect clustering and classification algorithms for pruning during query processing. Furthermore, in many cases, query formulation methods are not able to capture user's intention perfectly. For example, in Query by Example (QBE), which features, feature value ranges, feature combinations, or the similarity notion are to be used for processing is left to the system to figure out through feature significance

---

[5] See *"Image and Video Information Retrieval"* by Antonio Picariello

analysis, user preferences, relevance feedback **Error! Reference source not found.Error! Reference source not found.**, and/or collaborative filtering **Error! Reference source not found.** techniques, which are largely statistical and probabilistic in nature.

Processing multimedia queries requires assessment of different sources of imprecision and merging them into a single value. Traditional databases are not able to deal with imprecision since they are based on Boolean logic: predicates treated as propositional functions, which return either *true* or *false*. Multimedia information systems, on the other hand, require data models and query evaluation mechanisms which can take into account imprecision in the evaluation of the query criteria.

6.1 Data Representation

Captured media data is typically non structured or its structure is not known in advance, but structures (such as object segments or video sequences) are commonly pre-extracted using media processing algorithms. Authored or synthesized multimedia objects on the other hand have explicit authoring structures. Multimedia data models thus need to represent the visual features of the data as well as the underlying structures.

In order to associate multimedia data with information about their content, a conceptual representation needs to be given for every object. Such conceptual representation consists of a set of structured data describing the multimedia data, both in terms of their (hierarchical) structure, and of their semantic content. The structured description represents a surrogate of the corresponding object, and it is chosen as a representative of it at query/retrieval time. One way to define the surrogate of a multimedia object simply consists in providing a set of attributes characterizing the object itself. Content-based attributes (features) are usually automatically extracted by the system, and provide information about the content of the object they refer to. Examples of such attributes include color histograms, shape/texture descriptors for visual objects, mel-frequency cepstral coefficients for audio, term frequencies for text documents. The set of features used for describing the media objects in the database is naturally a function of the media type. Colors, textures, and shapes are commonly used to describe images. Time and motion are used in video databases. Terms (also referred to as keywords) are often used in text retrieval. The features that are selected to represent the objects in a database are selected based on three criteria

- o *application requirements:* some image database applications rely on color matching whereas in some other applications, texture is a better feature to represent the image content,
- o *power of discrimination:* since the features will be used during query processing to distinguish those objects that are similar to the user's query from those that are different from it, the features that are selected must be able to discriminate the objects in the database,
- o *human perception:* not all features are perceived equivalently by the user. For example, some colors are perceived more strongly than the others by the human eye [?]. Human eye is also more sensitive to contrast then colors in the image [?].

In addition, the *query workload* (i.e., which features seem to be dominant in user queries) and *relevance feedback* (i.e., which features seem to be relevant to a particular user in a particular query) need also be considered. Many multimedia applications also rely on

descriptive attributes (tags) manually associated to the objects to provide descriptions, mostly capturing their users' interpretations of them. These are commonly complemented with automated media-annotation and collaborative-filtering schemes which propagate available tags onto other non-tagged and non-annotated multimedia objects in the system. Ontologies, such as WordNet, also support retrieval of multimedia objects based on semantics similarities of the annotations and tags.

6.2 Components of an MODB

Feature Extractor. Given a media object, the feature extractor will identify which features are most significant and will extract these features. The relative importance of those features will be used during query processing.

Semantics Extractor. The semantics extractor identifies the real world meanings of media objects for semantic indexing For an image, its semantics may corresponds to the meanings and spatial relationships of the entities within the image.

Structure Manager. Temporal, spatial, and interaction structures of complex multimedia objects need to be indexed for querying purposes.

Index/Cluster Manager. The extracted features, semantics, and the structures have to be indexed for quick reference. An index/cluster manager chooses the best indexing mechanism for different features. Quad trees, k-d trees, R-trees are typical index structures to index multimedia data that can be modeled as vectors in a space. Edit distance algorithms are commonly used to measure differences between media objects that can be represented as sequences, graphs, or trees. Suffix-trees are commonly used for indexing strings. Hash-based signature files are commonly used for quick and dirty lookup. Indexes and clusters may over-prune the space resulting in misses or may return false positives. While in general, misses are less acceptable than false-positives, which need to be eliminated during post-processing

Query Processor. Multimedia queries have different characteristics from queries in traditional databases. One major difference is the similarity- (or quality-) based query processing requirement, as finding exact matches is either undesirable or impossible due to imperfections in the media processing functions. Another difference is that some of the user defined predicates, such as the media processing functions, may be very costly.

One common characteristic of multimedia information systems is the *uncertainty* or *imprecision* of the data. The models that can capture the imprecise and statistical nature of multimedia data and query processing are fuzzy and probabilistic in nature. Therefore multimedia data query evaluation requires fuzzy and probabilistic data and query models as well as appropriate query processing mechanisms. Probabilistic models rely on the premise that the sources of imprecision in data and query processing are inherently statistical and thus they commit onto probabilistic evaluation. Fuzzy models are more flexible and allow various different semantics, each applicable under different system requirements, to be selected for query evaluation.

The traditional approach to query optimization is to use database statistics to estimate the query execution cost for different execution plans and to choose the cheapest one. In

the case of a database for media objects and documents, the expected quality of the results is also important. Since different query execution plans may cause results with different qualities, the quality statistics must also be taken into consideration. In addition, query optimizers must take into account expensive user-defined predicates. In particular, multimedia predicates may have different alternative implementations, each with drastically different execution costs.

Since media-based queries are mostly similarity-based and user dependent, the results have to be ordered accordingly. Furthermore, the visualization and resource requirements of the objects may also need to be used for ranking. Since a combination of search criteria can be specified simultaneously, the quality results from each criterion must be merged to create the final ranking.

Visualization Specification. MODB queries may be accompanied by result visualization specifications. Providing specifications for result visualization in multimedia databases can be thought as authoring (i.e., complex media object composition). Thus, for visualization specification, MODB uses the MVIS services.

Feedback/Query Refinement. To guide the user in the search for a particular piece of information, MODB need to provide query verification, feedback, and refinement services. Multimedia information retrieval systems face an objective-subjective gaps: given a user query which features of the data are relevant (and how much so) for the user's query may not be known in advance, and the number of candidate matches in the database can be large. These challenges are dealt through relevance feedback. A relevance feedback cycle enables the multimedia information retrieval system to learn the user's interests and focus onto a suitable feature set through a query-driven, transparent, and iterative process (Figure 2): (1) Given a query, using the available index structures, the IR system (2) identifies an initial set of candidate results. Since the number of candidates can be large, the system presents a small number of samples to the user. (3) This initial sample and (4) user's relevance/irrelevance inputs are used for (5) learning user's interests (in terms of relevant features) and this information is used for (6) updating the query or the retrieval/ranking scheme.

Based on selectivity and the quality statistics, MODB may also provide users with system-feedback, including an estimated number of matching images, strictness of each query condition, and alternative query conditions. Users can relax or reformulate queries by selecting alternatives.

7.Profile and Context Manager (PCM)
Since many parameters, such as the semantics of the media objects, significance of various features, and tolerance to query execution time, are all user-dependent, PCM must maintain profiles. These user profiles include (1) object access profiles for the users (to optimize object caching and interaction management tasks), (2) user feature profiles (which describe those features that are more significant for a given user), and (3) semantics profiles (which describe the context in which a given user is working). In addition to the user profiles, PCM must maintain profiles about (4) resources available for users, (5) query execution statistics, and (6) content of the database. These profiles,

along with the user profiles, media type descriptions, and document and query specifications, are used in similarity- and cost- based optimization and feedback creation. The profile manager must use a different profiling technique for each type of profile. Note however that although they look different, many of these profiles are related. For instance, object access profiles can be estimated using the database content, user/feature, and user/semantic profiles. Similarly, resource distribution profile can be estimated using the object access profiles and the query execution statistics. To reduce the profiling cost, to deal with missing profiles in an effective manner, and to check the consistency of the existing profiles, PCM needs to allow the reconstruction of the unknown profiles from the known ones.

8. Digital Rights Management System (DRMS)

A major challenge faced in multimedia information systems is digital rights management. The fact that a multimedia object can be modified through transformations (such as scaling, cutting, or resizing) to obtain new objects makes it hard for the creator of the original object to assert her intellectual property rights and copyrights.

Digital signatures, such as watermarks, are imperceptible modifications made on multimedia objects that are hard to remove through common transformations without significantly affecting media object quality. Thus, the creator of the media object can assert her ownership by establishing that the media object under dispute contains her watermark.

A second approach to digital rights management is duplicate detection. In this approach, the original copy of the media object is registered by its owner. Then, a duplicate detection algorithm is run (for example over the web) to check whether there is any unauthorized copy of this object in a publicly available venue.

See:
 " Interactivity in multimedia documents and systems" by Ombretta Gaggi
 "Image and Video Information Retrieval" by Antonio Picariello
 "Automatic generation of video documentaries" by Stefano Bocconi
"Multimedia content description and standards" by Alberto Messina and Maurizio Montagnuolo
 "Streaming multimedia information processing" by K. Selçuk Candan

References

[1] S. Adalı, M.L. Sapino, and V.S. Subrahmanian. *Plan Databases: Model and Algebra*, private communication.

[2] ISO 8879:1986.

[3] World Wide Web Consortium (W3C). *Extensible Markup Language 1.0 Recommendation*, 10 Feb, 1998.

[4] ISO/IEC 10744.

[5] World Wide Web Consortium (W3C). (1998) *Synchronized Multimedia Integration Language Recommendation*, 15 June, 1998.

[6] http://www.w3.org/TR/REC-smil/

[7] ISO/IEC 13522-5.

[8] VRML (ISO/IEC 14772-1)

[9] VRML 200x or ISO/IEC 14772:200x.

[10] R. Hamakawa and J. Reikomoto, *Object Composition and PlaybackModels for HandlingMultimedia Data* ACM Multimedia Conference '93, Pages 273-281, 1993.

[11] J.F. Allen. *Maintaining Knowledge about Temporal Intervals*, Communications of the ACM, vol. 26, no. 11, pp. 832-843, November 1983.

[12] J.F. Allen. *Towards a General Theory of Time and Action*, Artificial Intelligence, 23, pp. 123-154, 1984.

[13] T.D.C. Little and A. Ghafoor. *Synchronization and Storage Models for Multimedia Objects*, IEEE J. on Selected Areas of Communications, vol. 8, no. 3, April 1990, pp. 413-427. April 1990.

[14] T.D.C. Little and A. Ghafoor. *Interval-Based Conceptual Models for Time-Dependent Multimedia Data* IEEE Transactions on Knowledge and Data Engineering, vol. 5, no. 4, pp. 551-563, Aug 1993.

[15] L. Li, A. Karmouch, and N.D. Georganas. *Multimedia Teleorchestra With Independent Sources: Part 1 and Part 2*, ACM/Springer-Verlag Journal of Multimedia Systems, vol. 1, no. 4, February 1994, pp.143-165. February 1994.

[16] M. Vilain and H. Kautz (1986) *Constraint Propagation Algorithms for Temporal Reasoning*, Proceedings of AAAI-86, Artificial Intelligence, pp. 377-382, August 1986.

[17] Claudio S. Pinhanez, Kenji Mase, and Aaron F. Bobick. *Interval Scripts: a Design Paradigm for Story-Based Interactive Systems*. CHI 1997, pp. 287-294, 1997.

[18] Claudio S. Pinhanez and Aaron F. Bobick. *Fast Constraint Propagation on Specialized Allen Networks and its Application to Action Recognition and Control*. Submitted to AAAI'98. Perceptual Computing Tech. Rep. No. 456. January 1998.

[19] K.S. Candan, B. Prabhakaran, and V.S. Subrahmanian. *A Framework for Supporting Multimedia Document Authoring*

and Presentation, The Fourth ACM International Multimedia Conference '96, pp. 329-339, Boston,
November 1996.

[20] K.S. Candan, B. Prabhakaran, and V.S. Subrahmanian. *Retrieval Schedules Based on Resource Availability and*
Flexible Presentation Specifications, ACM-Springer Multimedia Systems Journal, Vol. 6, No. 4, pp. 232-250,
1998.
[21] M.C. Buchanan and P.T. Zellweger (1993) *Automatic Temporal Layout Mechanisms* ACM Multimedia 93, pp.
341-350, 1993.
[22] M.C. Buchanan and P.T. Zellweger. *Scheduling Multimedia Documents Using Temporal Constraints*, Network
and Operating System Support for Digital Audio and Video, P. Venkat Rangan (ed.), Springer-Verlag, Berlin
Heidelberg New York, pp. 237-249, 1993.
[23] M.C. Buchanan and P.T. Zellweger. *Automatically Generating Consistent Schedules for Multimedia Documents*,
ACM/Springer-Verlag Journal of Multimedia Systems, vol. 1, no. 2, 1993.
[24] M.Y. Kim and J. Song. *Hyperstories: Combining Time, Space and Asynchrony in Multimedia Documents*,
RC19277(83726) (revised 1995) IBM Computer Science/Mathematics Research Report, 1993.
[25] M.Y. Kim and J. Song. *Multimedia Documents with Elastic Time*, ACM Multimedia Conference '95, 1995.
[26] P. van Beek. *Approximation Algorithms for Temporal Reasoning*, Proceedings of the 11th International Joint
Conference on Artificial Intelligence. Detroit, MI, USA, Morgan Kaufmann, pp. 1291-1296, August 1989.
[27] B. Aspavall and Y. Shiloach. *A Polynomial Time Algorithm for Solving Systems of Linear Inequalities with Two*
Variables per Inequality, SIAM Journal of Computing, vol. 9, no. 4, pp. 827-845, November 1980.
[28] R. Dechter, I. Meiri, and J. Pearl. *Temporal Constraint Networks*, Artificial Intelligence, vol. 49, pp. 61-95, 1991.
[29] G. Ö¨zsoyog˜lu, V. Hakkoymaz, and J.D. Kraft. *Automating the Assembly of Presentations from Multimedia*
Databases, Twelfth International Conference on Data Engineering, pp. 593-601, New Orleans, February 1996.
[30] V. Hakkoymaz and G. Ö¨zsoyog˜lu. *A Constraint-Driven Approach to Automate the Organization and Playout of*
Presentations in Multimedia Databases, Journal of Multimedia Tools and Applications, Kluwer Pub., Vol. 4, pp.
171-197, 1997.
[31] V. Hakkoymaz, J. Kraft, and G. Ö¨zsoyog˜lu. *Constraint-Based Automation of Multimedia Presentation Assembly*,

ACM Multimedia Systems Journal, to appear, 1998.

[32] M. Vazirgiannis and S. Boll. *Events in Interactive Multimedia Applications: Modelling and Implementation Design*,
IEEE International Conference on Multimedia Computing and Systems, June 1997.

[33] S. Wirag, *Scheduling of adaptive Multimedia Documents*, ICMCS'99, pp. 307-311, June 1999.

[34] Y. Song, M. Mielke, and A. Zhang. *NetMedia: Synchronized Streaming of Multimedia Presentations in Distributed Environments,* ICMCS'99, pp. 585-590, June 1999.

[35] I. Mirbel, B. Pernici, and M. Vazirgiannis. *Temporal Integrity Constraints in Interactive Multimedia Documents,*
ICMCS'99, pp. 867-871, June 1999.

[36] R.M. de Lima, P.A. da S. Gonc̦alves, F.P. Junqueria, and O.C.M.B. Duarte. *SAMM: An Integrated Environment to Support Multimedia Synchronization of Pre-orchestrated Data*, ICMCS'99, pp. 929-933, June 1999.

[37] C. Keramane and A. Duda (1996) *Interval Expressions - a Functional Model for Interactive Dynamic Multimedia Presentations*, IEEE International Conference on Multimedia Computing and Systems '96, pp. 283-286,
Hiroshima, June 1996.

[38] J. Song, Y.N. Doganata, M.Y. Kim, and A.N. Tantawi (1995) *Modeling Timed User-Interactions in Multimedia Documents*, IBM Computer Science Research Report, 1995. Also appeared in IEEE International Conference on
Multimedia Computing and Systems '96, pp. 407-416, Hiroshima, June 1996.

[39] K. Harada, E. Tanaka, R. Ogawa, and Y. Hara (1996) *Anecdote: A Multimedia Storyboarding System with Seamless Authoring Support*, The Fourth ACM International Multimedia Conference '96, pp. 341-351, Boston,
November 1996.

[40] B. Prabhakaran and S.V. Raghavan. *Synchronization Models for Multimedia Presentation With User Participation*,
ACM/Springer-Verlag Journal of Multimedia Systems, vol.2, no. 2, August 1994, pp. 53-62. Also in the
Proceedings of the First ACM Conference on Multimedia Systems, Anaheim, California, pp.157-166, August
1993.

[41] K.S. Candan. *A Framework for Distributed Multimedia Collaborations*, Ph.D. Thesis, University of Maryland at
College Park.

[42] K.S. Candan, E. Lemar, and V.S. Subrahmanian. *Interactive Result Visualization and View Management in Multimedia Databases*, Accepted for publication in the VLDB Journal on Multimedia Systems.

[43] K.S. Candan, B. Prabhakaran, and V.S. Subrahmanian. *Collaborative Multimedia Documents: Authoring and*

Presentation, Special issue of the International Journal of Intelligent Systems on Multimedia Computing Systems,
Volume 13(12), 1998.
[44] K.S. Candan, V.S. Subrahmanian, and P. Venkat Rangan.*Collaborative Multimedia Systems: Synthesis of Media
Objects.* IEEE Trans. on Knowledge and Data Engineering, Vol. 10, Nr. 3, pp. 433-457, May-June 1998.
[45] E. Oomoto and K. Tanaka. (1993) *OVID: Design and Implementation of a Video-Object Database System*, IEEE
Trans. on Knowledge and Data Engineering, 5, 4, pp. 629-643.
[46] W. Niblack, et. al. *The QBIC Project: Querying Images by Content Using Color, Texture and Shape*, IBM Research
Report, Feb. 1993.
[47] M. Arya,W. Cody, C. Faloutsos, J. Richardson and A. Toga. *Design and Implementation of QBISM, a 3DMedical
Image Database System*, in: (V.S. Subrahmanian and S. Jajodia, eds.) "Multimedia Database Systems: Issues and
Research Directions", Springer 1995.
[48] F. Idris and S. Panchanathan. *Review of Image and Video Indexing Techniques*, Journal of Visual Communication
and Image Representation-Special Issue on Indexing, Storage and Retrieval of Images and Video - Part II, vol. 8,
no. 2, pp. 146 - 166, June 1997.
[49] F. Idris and S. Panchanathan. *Storage and Retrieval of Compressed Images*, IEEE Transactions on Consumer
Electronics , vol. 41, no. 3, pp. 937 - 941, August 1995.
[50] F. Golshani and Y.C. Park. *Content-based Image Indexing and Retrieval in ImageRoadMap*, Proc. SPIE s International
Symposium on Voice, Video and Data Communications Multimedia Storage and Archiving Systems II,
Dallas, TX, November 1997.

[51] M. Mandal, F. Idris, and S. Panchanathan. *Image and Video Indexing in Compressed Domain - A Critical Review*,
to appear in the Journal of Image and Vision Computing - special issue on Content-based Image Indexing, January
1998.
[52] N. Dimitrova and F. Golshani. *Rx for Semantic Video Retrieval* Proceedings of the ACM Multimedia Conference,
San Francisco, October 1994, ACM Press, pp. 219-226, 1994.
[53] K.L. Liu, P. Sistla, C. Yu, and N. Rishe. *Query Processing in a Video Retrieval System* in Proceedings of the 14th
International Conference on Data Engineering, pp. 276-283, Orlando, Florida, USA, February 23- 27, 1998.
[54] H. Jiang, D. Montesi, and A.K. Elmagarmid. *VideText Database System*, 4th IEEE Int. Conf. on Multimedia
Computing and Systems, pp. 344-351, 1997.

[55] H. Jiang and A.K. Elmagarmid. *A Semantic Content-based VideoText Database System on the World Wide Web*,
To Appear in IEEE Transactions on Knowledge and Data Engineering.
[56] E. Wold, T. Blum, D. Keislar, and J. Wheaton. *Content-based Classification, Search, and Retrieval of Audio*,
IEEE Multimedia Magazine 1996.
[57] C. Faloutsos. *Searching Multimedia Databases by Content*, Kluwer Academic
Publishers, 1996.
[58] V. Oria, P. J. Iglinski, and M. T. Öˇ zsu. *A Framework for Multimedia Database Systems*,
4th African Conference
on Research in Computer Science, Dakar, Senegal, October 1998.
[59] V.S. Subrahmanian. *Principles of Multimedia Database Systems*, Morgan Kaufman
Press, Dec. 1997.
[60] J.R. Bach, C. Fuller, A. Gupta, A. Hampapur, B. Horowitz, R. Jain, and C.-F. Shu. *The Virage Search Engine: An*
open Framework for Image Management, SPIE- The International Society for Optical
Engineering: Storage and
Retrieval for Still Image and Video Databases IV, San Jose, CA, February 1996.
[61] W. Niblack, et. al. *The QBIC Project: Querying Images by Content Using Color, Texture and Shape,* IBM Research
Report, Feb. 1993.
[62] M. Flickner, et. al. *Query by Image and Video Content: The QBIC System*, IEEE
Computer, 28(9), pp. 23-32,
September 1995.
[63] M.T. Roth, M. Arya, L. Haas, M. Carey, W. Cody, R. Fagin, P. Schwarz, J. Thomas, E. Wimmers.*The Garlic*
Project, Proceedings of the 1996 ACM SIGMOD Conference, May 1996.
[64] M. Carey, L. Haas, V. Maganty, and J. Williams. *PESTO: An Integrated Query/Browser for Object Databases*,
Proceedings of the 1996 VLDB Conference, Bombay, India, Sept. 1996.
[65] S. Adalı, P. Bonatti, M.L. Sapino, and V.S. Subrahmanian. *A Multi-Similarity Algebra*,
Proceedings of the 1998
ACM SIGMOD Conference, pp. 402-413, June 1998.
[66] Wen-Syan Li *et al. WebDB Hypermedia Database System*, Submitted to IEICE
Transactions on Information
Systems - Special Issue on New Generation Database Technologies, 1998.
[67] V. Oria, M.T. ¨Ozsu, L. Liu, X. Li, J.Z. Li, Y. Niu, and P. Iglinski, *Modeling Images for Content-Based Queries:*
The DISIMA Approach, Second International Conference on Visual Information Systems,
pp. 339-346, San
Diego, CA, December 1997.
[68] M. T. Öˇ zsu, P. Iglinski, D. Szafron, S. El-Medani, M. Junghanns. *An Object-Oriented SGML/HYTIME Compliant*
Multimedia Database Management System, Fifth ACM International Multimedia Conference (ACM Multimedia
'97), pp. 239-249, Seattle, WA, November 1997.

[69] W.-S. Li, K.S. Candan, K. Hirata, and Y. Hara. *Hierarchical Image Modeling for Object-based Media Retrieval*,
Data and Knowledge Engineering, Elsvier Science, Elsvier Science, Vol:27, Nr:2, pp. 139-176, July 1998.
[70] W.-S. Li, K.S. Candan, K. Hirata, and Y. Hara. *Supporting Efficient Multimedia Database Exploration,* to apper
in TKDE Journal, 1999.
[71] W.-S. Li, K.S. Candan, K. Hirata, and Y. Hara. *A Hybrid Approach to Multimedia Database Systems through Integration
of Semantics and Media-based Search*, Springer-Verlag Lecture Notes in Computer Science - Worldwide
Computing and Its Applications, T.Masuda, Y.Masunaga, M.Tsukamoto (Eds.), Vol. 1274, P. 182-197, August,
1997. Also, invited talk at the 1st International Conference on The Worldwide Computing and Its Applications
'97 (WWCA97), March 10-11, 1997, Tsukuba, Japan.
[72] W.-S. Li, K.S. Candan, K. Hirata, and Y. Hara. *Facilitating Multimedia Database Exploration through Visual
Interfaces and Perpetual Query Reformulations* Proceddings of the 23rd International Conference on VLDB, pp.
538-547, August 26-29,1997, Athens, Greece.
[73] W.-S. Li and K.S. Candan. *SEMCOG: A Hybrid Object-based Image Database System and Its Modeling, Language,
and Query Processing* in Proceedings of the 14th International Conference on Data Engineering, pp.
284-291, Orlando, Florida, USA, February 23- 27, 1998.
[74] W.-S. Li and K.S. Candan.*SEMCOG: A Hybrid Object-based Image and Video Database System and Its Modeling,
Language, and Query Processing,* Accepted for publication in the Special Issue on Object-oriented Technology
in Advanced Applications of the Theory and Practice of Object Systems (TAPOS) Journal, 1998.
[75] W.-S. Li, K.S. Candan, K. Hirata, and Y.Hara. *Facilitating Multimedia Database Query and Exploration*, to
appear at the VLDB Journal on Multimedia Systems, 1998.
[76] S. Adalı, K.S. Candan, S.-S. Chen, K. Erol, and V.S. Subrahmanian. *Advanced Video Information System: Data
Structures and Query Processing.* ACM-SpringerMultimedia Systems Journal, Vol. 4, pp 172-186, August 96.
[77] W.-S. Li and K.S. Candan. *Integrating Content Search with Structure Analysis for Hypermedia Document Retrieval
and Management*, accepted for publication at the Computing Surveys.
[78] W.-S. Li, J. Shim, K.S. Candan, and Y. Hara. (1998) *WebDB: A Web Query System and its Modeling, Language,
and Implementation.* IEEE ADL'98 Conference, 1998.
[79] W.-S. Li, K.S. Candan, and Q. Vu. *Retrieval by Information Unit on the Web*, accepted for publication at the

VLDB Journal.

[80] C.T. Yu and W. Meng. (1998) *Principles of Database Query Processing for Advanced Applications.* Morgan
Kauffman Publishers, 1998.

[81] A. Yoshitaka and T. Ichikawa, *Survey on Content-based Retrieval for Multimedia Databases*, IEEE Transactions
on Knowledge and Data Engineering, volume 11, number 1, pp. 81-93, January 1999.

[82] Y.A. Aslandogan and C.T. Yu, *Techniques and Systems for Image and Video Retrieval*, IEEE Transactions on
Knowledge and Data Engineering, volume 11, number 1, pp. 56-63, January 1999.

[83] S. Adalı, K.S. Candan, Y. Papakonstantinou, and V.S. Subrahmanian. *Query Caching and Optimization in Distributed
Mediator Systems.* Proc. of the SIGMOD Conference on Management of Data, Montreal, June 96.

[84] P. Yamuna and K.S. Candan. *Similarity-based Retrieval of Temporal Documents.* Accepted for publication in the
Multimedia Information Retrieval workshop, 2000.

[85] V. Kobla, D.S. Doermann, and C. Faloutsos, *Developing High-Level Representations of Video Clips using Video-
Trails*, Storage and Retrieval for Image and Video Databases SPIE 1998, pp.81-92, 1998.

[86] O. Etzioni, K. Golden, and D. Weld. (1997) *Sound and Efficient Closed-world Reasoning for Planning.* Artificial
Intelligence, 89(1-2), 113-148.

[87] S. Chaudhuri and L. Gravano. (1996) *Optimizing Queries over Multimedia Repositories.* SIGMOD' 96, pp. 91-
102, 1996.

[88] S. Chaudhuri and K. Shim. (1996) *Optimization of queries with User-defined Predicates.* VLDB'96, pp. 87-98,
1996.

[89] J.M. Hellerstein andM. Stonebraker. (1993) *PredicateMigration: Optimizing Queries with Expensive Predicates.*
SIGMOD'93, pp. 267-276, 1993.

[90] J.M. Hellerstein. (1994) *Practical Predicate Placement.* SIGMOD'94, pp. 325-335, 1994.

[91] R. Fagin. *Fuzzy Queries in Multimedia Database Systems.* 17th ACM Symposium on Principles of Database
Systems, pp. 1-10, June 1998.

[92] R. Fagin. *Combining Fuzzy Information from Multiple Systems.* 15th ACMSymposium on Principles of Database
Systems, pp. 216-226, June 1996.

[93] K.S. Candan and W.-S.- Li. *On Similarity-based Ranking and Query Processing in Multimedia Databases*, accepted
for publication at the DKE Journal.

[94] S. Adalı, M.-L. Sapino, V.S. Subrahmanian. *AMultimedia Presentation Algebra*, SIGMOD'99, pp. 121-132, June
1999.

[95] A. Del Bimbo, E. Vicario, and D. Zingoni, *Symbolic Description and Visual Querying of Image Sequences Using*

Spatio-Temporal Logic, IEEE Transactions on Knowledge and Data Engineering, Vol. 7, n.4, August 1995.

[96] G. Ahanger and T.D.C. Little, *Data Semantics for Improving Retrieval Performance of Digital News Video Systems*,
DS-8 1999, pp. 47-64, 1999.

[97] M. Vazirgiannis, Y. Theodoridis, and T.K. Sellis, *Spatio-Temporal Composition and Indexing for Large Multimedia*
Applications Multimedia Systems 6(4), pp. 284-298, 1998.

[98] E. Schwalb and R. Decther, *Processing Disjunctions in Temporal Constraint Networks,* Artificial Intelligence 93,
pp. 29-61, 1997.

[99] T. Milo and D. Suciu. *Index Structures for Path Expressions*, ICDT 1999, pp. 277-295, 1999.

[100] T. Milo, D. Suciu, and V. Vianu. *Typechecking for XML Transformers*, PODS 2000, pp. 11-22, 2000.

[101] D. Suciu, *Semistructured Data and XML*, Proceedings of 5th International Conference of Foundations of Data
Organization (FODO'98), Kobe, Japan, November 1998.

[102] K.-C. Tai. *The Tree-to-Tree Correction Problem*, Journal of ACM, pp. 422-433, vol 26, March 1979.

[103] K. Zhang and D. Shasha. *Simple Fast Algorithms for the Editing Distance between Trees and Related Problems*,
SIAM Journal of Computing, 18(6), pp. 1245-1262, 1989.

[104] D. Shasha and K. Zhang. *Fast Algorithms for the Unit Cost Distance between Trees*, Journal of Algorithms, 11,
pp. 581-621, 1990.

[105] K. Zhang, J.T.L.Wang, and D. Shasha. *On the EditingDistance between Undirected Acyclic Graphs*, International
Journal of Computer Science, 7(1), pp. 43-57, 1996.

[106] S. Chawathe. *Comparing Hierarchical Data in External Memory*, to appear in Proceedings of the Twenty-fifth
International Conference on Very Large Data Bases, Edinburgh, Scotland, U.K., 1999.

[107] S. Chawathe and H. Garcia-Molina. *Meaningful Change Detection in Structured Data,* Proceedings of the ACM
SIGMOD International Conference on Management of Data , pp. 26-37, Tucson, Arizona, May 1997.

[108] S. Selkow. *The Tree-to-tree Editing Problem*, Information Processing Letters, 6(6), pp. 184-186, 1977.

[109] F. Luccio and L. Pagli. *Approximate Matching for Two Families of Trees,* Information and Computation, 123(1),
pp. 111-120,1995.

[110] M. Farach and M. Thorup. *Sparse Dynamic Programming for Evolutionary-tree Comparison,* SIAM Journal on
Computing, 26(1):210-230, January 1997.

[111] E. Myers. *An* O(ND) *Difference Algorithms and its Variations.* Algorithmica, 1(2), pp. 251-266,1986.

[112] S. Nestorov, J. Ullman, J. Wiener, and S. Chawathe. *Representative Objects: Concise Representations of*
Semistructured, Hierarchical Data, in Proceedings of the International Conference on Data Engineering, pp.
79-90, U.K., 1997.
[113] D. Florescu, A. Levy, I, Manolescu, and D. Suciu. *Query Optimization in the Presence of Limited Access Patterns,*
SIGMOD'99, pp. 311-322, June 1999.
[114] S.L. Hakimi, *Steiner's Problem in Graphs and its Implications*, Networks, volume 1, pp. 113-131, 1971.
[115] , G. Reich and P. Widmayer, *Approximate Minimum Spanning Trees for Vertex Classes*, author=G. Reich and P.
Widmayer, Technical Report, Inst. fur Informatik, Freiburg Univ., 1991.
[116] G. Reich and P. Widmayer. *Beyond Steiner's Problem: A VLSI Oriented Generalization*, Proceedings of the 15th
Int. Workshop on Graph Theoretic Concepts in Computer Science WG89, Lecture Notes in Computer Science,
vol. 411, pp. 196-210, 1990.
[117] P. Slavik. *The Errand Scheduling Problem*, Technical Report 97-2, Department of Computer Science, SUNY,
Buffalo NY, March 14, 1997.
[118] C.D. Bateman, C.S. Helvig, G. Robins, and A. Zelikovsky. *Provably Good Routing Tree Construction with Multiport*
Terminals, Proceedings of the ACM/SIGDA International Symposium on Physical Design, Apr 1997.
[119] N. Garg, G. Konjevod, and R. Ravi. *A Polylogarithmic Approximation Algorithm for the Group Steiner Tree*
Problem, Proceedings of the 9th Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 253-259, 1998