10 Month Report

Semantic Multimedia Database Information Retrieval

Yi Li BSc, MSc
Email: liy@sbu.ac.uk

Supervisors: Prof. Marios C. Angelides (Director of studies)
Dr. Harry W. Agius (Second supervisor)

05 March 1999

Centre for Multimedia
School of Computing, Information Systems & Mathematics
South Bank University

Contents

1 Background Literature Review

1.1 What is Multimedia?

Many people when talking about multimedia refer to it as "the multimedia industry." It is extremely hard to find a clear division between multimedia and any of the other industries mentioned in the same breath-namely, the entertainment, computer hardware, software, publishing, communications and music industries. It looks like everything is multimedia [11]. At a simplistic level, multimedia can be defined as the combination of more than one medium-text, graphics, sound, animation, and video-commonly assumed to be in digital format.

However, we'd better give it a more exact definition to conduct the following research. The most general definition states that **Multimedia** is the seamless integration of data, text, images of all kinds and sound within a single, *digital* information environment. Further refinement of a multimedia system definition should include *interactive* character [8]. Two important features are described as follows:

Digital environment
Before computer world, the term 'multimedia' was already in use. The products certainly offered multiple media - text, images and sound - but each was delivered as an independent element in the package. *Electronic technology* (digital environment) provides a single medium with the power to integrate diverse types of information [12].

Interactively
Interactivity is one of the most obviously unique features which multimedia offers. Because this 'non-linearity' is one of its most powerful advantages over traditional, linear media such as film or video, it is hardly surprising that proponents of multimedia promote interactivity as a vital ingredient [12].

Interactively is really another word for the ways in which a user can search and browse through an electronic database, the process being more or less constrained by the control software.

The following four important factors brought out by multimedia and will affect the development of it in all multimedia systems.

1. Need very large memory stores.
2. Handle retrieval, processing and display of high volumes of information.
3. Output both sound and images to the required standards of any given application.
4. Easy navigation.

1.2 Digitising Information

Whether you like it or not, digitalisation will dominate the whole information world. It is not only because the computer only know digital data. It also gives us incredible power to control the digitised information. - It is a practical issue because by taking information out of the analogue world, the "real" world, comprehensible and palpable to human beings, and translating it into the digital world, we make it infinitely changeable [12].

Table 1 shows the conventional classification of media types. It is easy to digitising synthesised media through some coding and specifications e.g. ASCII. For media captured from real world more work needs to be done on digitising. Sounds first need sampling and then followed by quantisation to finish digitising. Images are digitised by pixels, the smallest element of resolution of the image, which have a numerical value each. The size of digitised captured media is much bigger than that of synthesised media.

	Captured (from real world)		**Synthesised** (by computers)	
Discrete		Still images	Text	Graphics
Continuous	Sound	Moving images		Animation

Table 1. Conventional classification of media types

1.3 Data Compression

In the case of video, processing uncompressed data streams in an integrated multimedia system leads to secondary storage requirements in the range of at least giga-bytes, and in the range of mega-bytes for buffer storage. The throughput in a multimedia system can be as high as 140 Mbits/second, which must be transferred between different systems. This kind of data transfer rate is not realisable with today's technology, or in the near future with reasonably priced hardware. However, the use of appropriate compression techniques considerably reduces the data transfer rates, and fortunately research, development and standardisation have rapidly progressed in this area during the last few years.

There are two modes of compression: lossless and lossy compression. Two compression methods are used broadly (see Table 2): entropy encoding and source encoding, which optimises the compression according to the semantics of the original data. [13]

Entropy coding	Repetitive sequence suppression	Zero suppression Run-length encoding
	Statistical encoding (most frequent bit/ch identified)	Pattern substitution Huffman-like encoding
Source coding (mathematical transformation)	Transform encoding	FFT DCT Others
	Differential encoding	DPCM (differential pulse code modulation) Delta modulation (1 bit to code the difference) ADPCM (adaptive differential pulse code modulation)
	Vector quantisation	

Table 2. The major data compression techniques

JPEG is probably the most popular compression standard for continuous-tone grey-scale or colour images. DCT (discrete cosine transform) is the main technique used together with other techniques e.g. Huffman encoding during JPEG process. Figure 5 shows the process steps of JPEG.

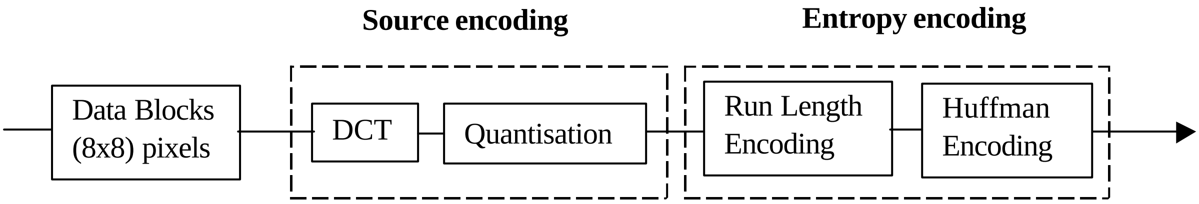

Figure 5. Processing steps of JPEG/sequential lossy mode

MPEG is compression standard for moving images based on JPEG. Considering two types of redundancies (correlations), the spatial redundancies and the temporal redundancies, the MPEG standard has three major components: MPEG-Video, MPEG-Audio, and MPEG-System which specifies how the two streams are multiplexed and synchronised.

More details on data compression can be found in Ralf Steinmetz's excellent book [14].

1.4 Networked Multimedia

Progress in networking pushes the development of multimedia, which gives it extremely power to change people's life, e.g. videophone, virtual shopping, remote teaching. The following techniques are main, not least, critical for networked multimedia.

1.4.1 Client/Server Architecture

Client/server application architecture was introduced to address issues of cost (client side- PC rather mainframe computer) and performance (C/S applications allowed for applications to run on both the user workstation and the server). This kind of architecture needs network support and has a great use in database applications. The Client sends a SQL request to the Server. After procession of an application running locally on the server machine, the Server sends the result back to the Client.

There are mainly two kinds of models of C/S architecture, two-tier and three-tier model. In the two-tier model, logic is split between these two physical locations, the client and the server. Business logic for application must physically reside either on the client or be implemented on the back-end within the DBMS in the form of triggers and stored procedures. This simple model has three critical limitations: not scalable, unmanageable, and poor performance. The three-tie model may tackle with these problems. The Service Model is logically grouped into three tiers: User services, Business services, and Data services. But three-tier development is not the answer to every situation. Good partitioning and component design take time and expertise, both items that are in short supply. Additionally, three-tier client/server development requires the support and commitment of the enterprise's powers that be. Two-tier C/S development is a much quicker way of taking advantage of SQL database engines and can fit the bill if both money and time are running out [15].

1.4.2 ATM

Networked multimedia needs more broadband of the network than other applications. ATM (asynchronous transfer mode) seems to be the most suitable technique to serve networked multimedia. It uses two different concepts. One is call which should be set up prior to the use of a communication, and the other is data packets which is not sent continuously, but chopped into a succession of fixed units – cells. ATM normally has three levels of usage. Firstly, it is used as a transport technology used by the PNOs (Public Network Operators). Secondly, it is used to be a long-distance technology brought into to the user premises. Thirdly, it is used to build or upgrade existing local-area networks. Clients can use it to get multimedia services e.g. video-on-demand.

1.4.3 Internet and WWW

The World Wide Web, or simply the Web, is a most influential Internet data transfer method that came into existence not long ago. Thanks to this recent invention, Internet users all over the world are able to access remote servers of multimedia information from any regular personal computer. As a result, Internet has been growing very rapidly. Actually this new technology is changing the whole pattern of human culture and communication. The reason the World Wide Web model is so popular and powerful is that it offers means to distribute or access any form of digital data (multimedia) easily and inexpensively. It is available to both companies and individual consumers. Obviously this technology has profound implications for every aspect of our life - business, culture and society. However, the current Internet cannot support most advanced multimedia services due to its narrow bandwidth. The information superhighway is suggested to tackle with this problem.

The promise multimedia services supported by information superhighway are as follows:

1. Videophone
2. Multimedia e-mail
3. Inter net meeting
4. Digital/High-definition TV broadcast
5. Video-on-demand
6. Movie-on-demand
7. Multi-player Internet games
8. Access to digital library/museum
9. Customised multimedia-based learning
10. Virtual shopping/banking/ticketing

1.5 Virtual Reality

Just several years ago, it seemed that only a small group of people at California indulged in it. Now it gets broadly notice for its useful and practical future. Through the use of VR interfaces, a person actually “merges” into a computer-generated scene to roam through a synthetic environment. Many books are available for it now.

1.6 Research Areas

Since Multimedia covers most computer areas and some industry areas, research in it has lots of areas. The following is one way to list most current active research areas.

1.6.1 Technical

1. <u>Compression</u> – in the multimedia world it is necessary to capture, move, and store visual information in a compressed form. Direct manipulation of compressed images and video

based on a set of block-level transforms is now possible [15]. There are several compression techniques, mainly syntactic compression. These are proposals for semantic compression.

2. Distributed multimedia systems – large volume of data put a great pressure on the communication infrastructure. ATM, megabit switching, ISDN and online compression are used to improve the operation of the network [12, 16].
3. Multimedia databases – they comprise storage, query language design, presentation, search, retrieval and multimedia database indexing.
4. Multimedia Languages – A new generation of markup and programming language is designed for creating uniform and reconfigurable multimedia web sites with minimal or no programming effort on the part of user. Other languages enable finding and extracting information for making multimedia publications usable and useful are also needed. The most popular language, the query language, requires algorithms that are provably correct in processing and whose efficiency can be appropriately evaluated. Sometimes a software framework is used for composing distributed multimedia applications supported by a variety of platforms and a WWW-based system.
5. Protocol design and implementation – it is obviously inefficient to transfer real time data between clients and multimedia servers over Internet by using HTTP/TCP. To improve the performance of TCP over the ATM may be a solution [17].
6. Virtual Reality – The impact will be very widespread e.g. in entertainment, medicine, advertising, engineering, science, training, accident simulator and etc. The growth of Web and the interface between Java and VRML give more power to virtual reality [18].

1.6.2 Socio-Technical

1. Digital Libraries – their works range from major historical archives, to audio archives, virtual museums, and virtual art galleries.
2. Multimedia Conferencing – developments on multimedia systems and networking technology show that using desktop multimedia conferencing for group decision making on wide area networks such as the Internet is now part of the corporate suite [19].
3. Video on Demand – it will be one of the most important commercial applications of distributed multimedia systems. The major limiting constraints on VoD is the ability to satisfy the huge bandwidth and capacity requirements of VoD. The current model normally uses batching and buffer sharing techniques in video servers to support a large number of VoD services [20, 21].

1.6.3 Organisational

1. Quality of Service – this is an extremely popular research area that tries to guarantee the specific requirements over the network. Researches also address the issue of an overall QoS architecture (end-to-end) for multimedia communications [22].
2. Online Marketing – business online is a key factor to boost the World Wide Web. Electronic catalogues, electronic malls are boom on the net. New models on the web based on advertising, fees, or transactions are begun to test [17].

Here is another way to classify the research areas in Multimedia for reference (see Table 3).

Multimedia Databases	Multimedia Tools	Multimedia Applications
Content-based indexing/*retrieval* Digital libraries Image, video, audio content analysis Browsing and visualisation	Special hardware devices MPEG, QuickTime, API standards Java, VRML, Multimedia languages Multimedia authoring tools Multimedia software engineering tools Animation and computer graphics Pattern recognition, image processing	Educational applications Art and Multimedia Cultural heritage and Multimedia Medical Multimedia applications Electronic commerce 3D audio and video
Distributed Multimedia	**Operating System Support**	**Human Computer Interaction**
Multimedia on the Internet Web servers and services Intelligent network architectures Mobile network architectures	Network & resource management Quality of service control/scheduling Audio and video compression Multimedia database management	Advanced man-machine interfacing Visual languages and computing Multimodal interaction Virtual and augmented reality

Table 3. Multimedia research areas

2 Research Proposal

Multimedia data in traditional databases are stored in the form of raw alphanumeric data. Quite often we are interested in only certain information inside these raw data, for example, "What is happening within the media at time *t*? ". Since they are raw data, we cannot get it directly from the databases. Therefore, a new kind of database system, which I call semantic multimedia database, is needed that stores the semantic meaning of the raw data such that we are able to query the content of the data. In a semantic database, reference does not need to be made to the entire raw data but only to the selected content.

Multimedia semantics refers to the meaning depicted within videos, audios, and etc. and Semantic Multimedia Database (SMDB) systems are thus intended to integrate semantic information of a wide variety of formats, i.e. text animation, audio and video [1]. Agius and Angelides (1997, 1999) suggest a semantic content-based model that integrates syntactic and semantic information of multimedia [2, 3]. It consists of a syntax m-frame (multimedia frames) layer and a semantic m-frames layer. A syntax m-frame of each frame's (video/audio) content is created to describe the syntactic content occurring within that frame. This is what is traditionally being stored in a database system. Semantic m-frames are generated based on syntax m-frames and a kind of object model that consists of three parts: description, events and actions, which describe the object, its activities and the events in which the object is engaged in. The proposed semantic database will be developed to accommodate the semantic m-frame.

Integrating multimedia information to a database has a great impact on its design and functions. If we only store multimedia as files, then a multimedia file server with pointers maintained by a relational database such as Oracle will be enough [4]. If more functions such as indexing, searching and querying based on semantics are required from the database, then new designs and techniques have to be developed. The design of this multimedia database requires [5]:

1. The SMDB conceptual model
2. Indexing structure and techniques
3. SMDB content-based query language
4. Visual interface for content-based information retrieval

The **aims and objectives** of this research are as follows:

- Develop an indexing structure for semantic multimedia databases
- Develop a content-based query language for semantic multimedia databases
- Develop a visual interface for posing content-based queries to the semantic multimedia database

Semantic multimedia databases have far more advantages than the traditional databases in audio and video content requirements. Applications can benefit from SMDB because retrieval does not result in whole audio/visual document retrieval and thus placing the task of meaning interpretation on the user, but the system is able to respond to specific content query [6].

A semantic multimedia database should be set up first. Then the index structure will be built based on file structure of the database. A new query language will be added to the database to ease the retrieval. Finally, a complicated sample should be used to test the retrieval system. The following is the research timetable:

Time Table			
ID	Task Name	Starting Date	Ending Date
1	**Review and Background Reading**	**Mar. 1998**	**Nov. 1998**
2	Multimedia introductory domain reading	Mar. 1998	May 1998
3	Multimedia databases reading and reviewing semantic retrieval	Jun. 1998	Nov. 1998
4	**Build Up Semantic Multimedia Database**	**Dec. 1998**	**Jun. 1999**
5	Build up the database prototype	Dec. 1998	Apr. 1999
6	Set up the database model	May 1999	Jun. 1999
7	**Set Up Retrieval Model**	**Dec. 1998**	**Feb. 2000**
8	Set up the retrieval model	Dec. 1998	Apr. 1999
9	Build up query interface	May 1999	Jun. 1999
10	Build up query language transform mechanism	Jul. 1999	Sep. 1999
11	Build up search mechanism	Oct. 1999	Dec. 1999
12	Build up the presentation prototype	Dec. 1999	Feb. 2000
13	**Test**	**Feb. 2000**	**Jun. 2000**
14	Collect test results	Feb. 2000	Mar. 2000
15	Modify the retrieval system accordingly	Apr. 2000	Jun. 2000
16	**Writing up of Thesis**	**Jul. 2000**	**Feb. 2001**

The rest of this paper includes section 2 which discusses semantic multimedia database development, section 3, which discusses semantic multimedia database information retrieval, and section 4, which presents a review of general knowledge of multimedia.

3 Semantic Multimedia Databases

3.1 Multimedia Databases

Multimedia database systems is a relatively new area of research. It tries to incorporate different kinds of media objects, such as audio and video data, into the database in addition to alphanumeric information.

There are currently four techniques that have been used in multimedia data management [7]:

1. Local Storage –multimedia data is stored in files on the local system drive. The advantage is that there is no network or system delays imposed on the delivery of the multimedia data. The disadvantage is that there is no sharing of data storage locations, therefore, moving the data difficult to update.
2. Media Server – it is a shared storage facility that is analogous to a file server with the added capability of delivering multimedia data. Its function is limited to responding the client's request by opening the multimedia data file and delivering the multimedia content in an isochronous fashion.
3. Binary Large Objects – a relational database stores multimedia data by using binary large objects (BLOBs) as an attribute of its relations. The advantage is obvious if we treat the multimedia content as a single large object. But the database system can not apply delivery optimisation techniques because the BLOB is untyped and there is no method of working with or modifying its structure.
4. Object-Oriented Methodologies – to use object-oriented methodologies is a good method to overcome the problems with BLOBs. It provides a framework for defining extensible user defined data types and the ability to support complex relationships in the object-oriented database. But the OO does not adequately solve all problems associated with multimedia management.

3.2 Multimedia Databases Applications

Applications can benefit from multimedia databases, including [6]:

Medical information systems: contain medical imaging (X-ray, CAT scan), monitoring information (EKG recordings), as well as photographs of characteristic physical symptoms.

Engineering information systems: include both manually generated and computer-generated blueprints, sketches diagrams and illustrations. Photos documenting construction stages are also useful.

Office and library information systems: information on paper can be scanned and stored in an image database. Non-paper objects can be photographed or video taped and stored in a multimedia database.

Consumer catalogues: not only contain pictures and textual descriptions, but may also contain verbal commentary and video demonstrations of goods and services.

Training and education: contain video clips demonstrating how things work, how to repair things, and how to assemble things.

Geographic databases: maps of all kinds, as well as aerial and satellite photographs, can be stored and analysed by geographic database systems.

Reference works: include encyclopaedias containing news clips, audio clips, and digitised photographs.

3.3 Semantic Multimedia Databases

The problem with storing multimedia information is the lack of providing enough information for the users because we do not know what is inside these raw data (e.g. BLOBs). One of good way to solve the problem is to store the semantic meaning of the raw data as well to let the users be able to query and interact with the content of the raw data. This new kind of database system is called a semantic multimedia database. In a semantic database, reference is not made to the entire raw data but only to the selected content.

3.3.1 Architecture of semantic multimedia database

Multimedia semantics refers to the meaning depicted within videos, audios, and etc. and Semantic Multimedia Database (SMDB) systems are thus intended to integrate semantic information of a wide variety of formats, i.e. text, animation, graphics, audio and video [8]. Figure 1 shows the architecture of semantic multimedia database management.

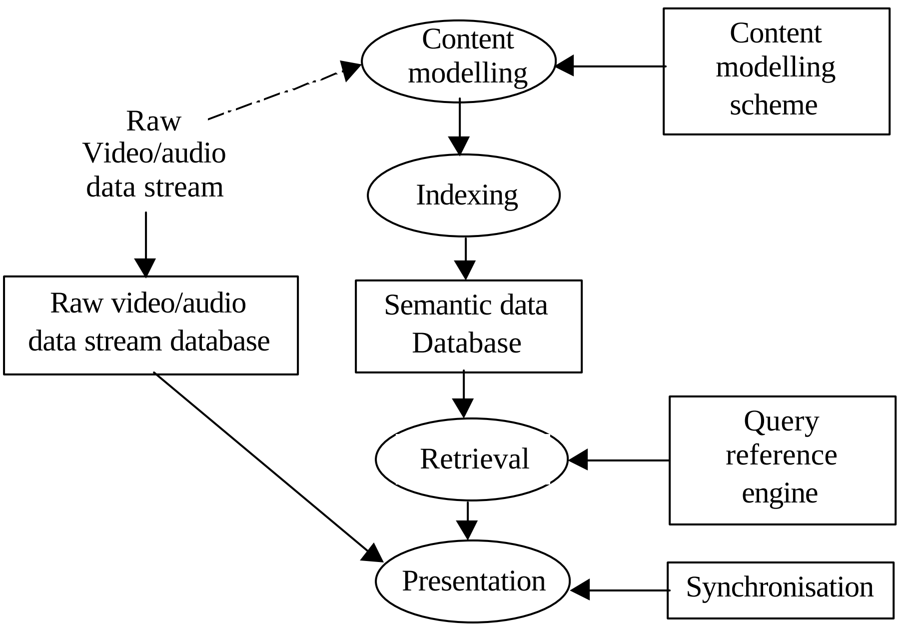

Fig. 1 Diagram of semantic multimedia database management architecture

3.3.2 Basic concepts and assumptions of semantic multimedia database

Here are some important basic concepts and assumptions pertaining to multimedia data type in semantic database.

A *frame* is a single frame of video or audio (1/25 sec).

Key frame is a frame of video in which the object picture or their spatial relationship has changed. It becomes the smallest unit for indexing.
A *shot* is defined as an arbitrary sequence of continuous frames that are related in that together they constitute some form of continuity in meaning within the sequence.
A *Slot* is a field provides descriptive information. An element distinguishes the topical characteristic for each object represented by a frame.
Multimedia *syntax* refers to the organisation and representation of multimedia information.
Multimedia *semantics* refers to the meaning depicted within videos and audios.

Entities of interest integrate semantic content-based information about raw video and audio is traditional knowledge.

Most segments of information are difficult to exact from single frames because they have meaning over time and are also often meaningless when taken out of context. Moreover, it is not always possible to attribute events or actions based on a single frame.

3.3.3 Semantic aspects

The main semantic aspects of a semantic content model which will be included in the SMDB are as follows:

Objects
There are advantages in describing multimedia information as objects. First because of its multiple nature, and also because of the frequently huge volumes of multimedia information to be handled, which leads us to breaking them down into smaller components called as "objects."

Spatial relationships between objects
The different locations of objects have much semantics. Spatial relationships within existing models can be determined from the co-ordinates of the objects. But the on-screen location for hidden objects should also be determined in some other ways.

Events and actions
Events often consist of one or many actions that make it clear that one or more objects are involved. The search sequence is events first, then actions.

Temporal aspects of objects
The sequence and timing of objects (events and actions) within the media, e.g. A happens before B. Without temporal relationships, the representation of objects becomes ambiguous and leading to no meaning.

Explicit media structure
Separating the sequences of video and audio into meaningful segments and then combining them into flat or hierarchical structures creates explicit media structure. To build up the EMS from multimedia data, especially video, it needs some knowledge outside the media to construct the structure.

Integration of syntactic and semantic information

The integration of syntactic and semantic information can express all the meaning of the video and audio. The integration mainly happens: between the events and actions, and the temporal relationships; between the objects, and the spatial relationships.

3.3.4 COSMOS - A new kind of semantic content model

Agius and Angelides (1997) suggest a semantic content-based model that integrates the above six aspects of semantic which they call COSMOS [2]. It consists of a syntax m-frame (multimedia frames) layer and a semantic m-frames layer. A syntax m-frame of each frame's (video/audio) content is created to describe the syntactic content occurring within that frame. This is what is traditionally being stored in a database system. Semantic m-frames are generated based on syntax m-frames and a kind of object model that consists of three parts: description, events and actions, which describe the object, its activities and the events in which the object is engaged in. Figure 2 shows a simple structure of the model.

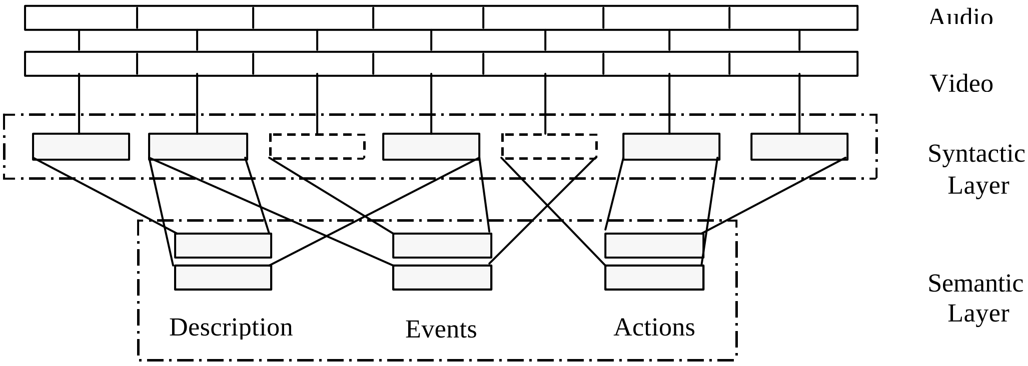

Figure 2. COSMOS

One of important goals for multimedia data management is to provide separation between the application's logical view of data organisation and the physical organisation of the stored data. This new model has very clear line to separate physical and logical view of the multimedia data.

3.3.5 A simple case study for COSMOS

The following is some frames of a football clip. Six aspects of the clip are built up by using COSMOS:

Figure 3. A football clip

1. Objects

Key Frame No.	***Object Name***	**Location**	*Meta-Data*
0-5 6-10	football player	66, 120, 110, 207 50, 150, 85, 207	sex: male, age: adult
11-21	football players	59,42,227,166	sex: male, age: child
22-27 28-35	football player	173,88,212,196 116,52,175,181	sex: female, age: child
36-47	goalkeeper	112, 48, 167, 202	sex: male, age: adult
48-58 59-66	football player	116,53,187,208 158,28,221,164	sex: male, age: child
67-86	football player	77,70,186,205	sex: male, age: adult
87-102	football player	200, 48, 238, 180	sex: male, age: adult
	goalkeeper	130, 106, 150, 155	sex: male, age: adult
103-108	football player	100,100,166,206	sex: male, age: adult
109-122	football player	126, 118, 196, 213	sex: male, age: adult
123-127 128-132	football player	205,160,232,206 163,142,199,188	sex: female, age: adult
133-138	football player	126,56,181,181	sex: female, age: child
139-153	goalkeeper	104,34,202,208	sex: male, age: adult
154-156	defender forward	164,121,197,177 139,125,166,192	sex: male, age: adult sex: male, age: adult
157-158	defender forward	169,109,196,174 141,119,167,177	sex: male, age: adult sex: male, age: adult
	goalkeeper	37,101,65,162	sex: male, age: adult
159-160	defender forward goalkeeper	166,112,195,169 142,116,162,171 92,95,108,146	sex: male, age: adult sex: male, age: adult sex: male, age: adult
161-162	defender forward goalkeeper	123,95,164,166 123,95,164,166 123,95,164,166	sex: male, age: adult sex: male, age: adult sex: male, age: adult
163-167	defender forward goalkeeper	64,97,90,154 42,103,67,159 138,103,156,158	sex: male, age: adult sex: male, age: adult sex: male, age: adult
168-181	goalkeeper	92,100,240,177	sex: male, age: adult
182-196	football player #1	93,89,134,185	sex: male, age: adult
193-196	football player #2	180,84,222,171	sex: male, age: adult
197-213	football player #1 football player #2	122,95,151,182 122,95,151,182	
214-221	football player #1 football player #2	94,90,136,172 152,90,189,178	
222-234	football player	52,26,158,177	sex: male, age: adult
235-246	football player	108,46,193,198	sex: male, age: child
247-255	football players	57,88,189,153	sex: male, age: child
256-266	football players	48,51,209,158	sex: male, age: child
267-279	football players	181,79,203,99	sex: male, age: adult
283-300	Goalkeeper	160,88,194,112	sex: male, age: adult

2. Spatial Relationship

Key Frame No.	***Object 1***	***Relationship***	***Object 2***
87-102	football players	$\Uparrow >$	goalkeeper
154-156	Defender	$\Uparrow <$	forward
157-160	Defender	$\Uparrow <$	forward
	Defender	$\Uparrow >$	goalkeeper
	Forward	$\Uparrow >$	goalkeeper
161-162	Defender	$\Downarrow =$	forward
	Defender	$\Uparrow =$	goalkeeper
	Forward	$\Uparrow =$	goalkeeper
163-167	Defender	$>=$	forward
	Defender	$<$	goalkeeper
	Forward	$<$	goalkeeper
193-196	football players #1	$\Uparrow <$	football players #2
197-213	football players #1	$\Downarrow =$	football players #2
213-221	football players #1	$\Downarrow <$	football players #2

Notes:

X touches Y: $X = Y$; Y touches X : $Y = X$;
X above Y : $X \uparrow Y$; Y beneath X : $Y \downarrow X$;
X inside Y : $X \subseteq Y$; Y encapsulates X : $Y \supseteq X$;
X left Y : $X < Y$; Y right X : $Y > X$;
X before Y : $X \Uparrow Y$; Y behind X : $Y \Downarrow X$;

3. Events and Actions

Frame No.	***Events***	***Actions***	*Frame No.*
0-10	throw-in	throw-in	0-10
11-21	tap ball	tap ball	11-21
36-47	goal kick	goal kick	36-47
67-86	overhead kick	overhead kick	67-86
87-102, 235-246	control ball by thigh	control ball by thigh #1 control ball by thigh #2	87-102 235-246
103-108	control ball by chest	control ball by chest	103-108
109-138	header	header	109-138
139-167	clean catching	clean catching	139-167
168-181	save	save	168-181
182-195,222-234	dribbling	dribbling dribbling	182-195 222-234
196-221	intercept	intercept	196-221
247-255	training	training	247-255
22-35,48-66,256-266	easy play	easy play #1 easy play #2 easy play #3	22-35 48-66 256-266
267-300	scored goal	player shoots the ball goalkeeper misses the ball the ball in net	267-285 286-289 290-300

4. Temporal Relationships (between events and actions)

scored goal:

player shoots the ball	(<, [ )	goalkeeper misses the ball
goalkeeper misses the ball	(<, [ )	ball in the net

Notes:

If there are two event/actions (start A, end A) and (start B, end B), then the temporal relationships between the two event/actions can be expressed in (TR1, TR2).

Where TR1, TR2 $\in$ { <, [, #,], >, * }, and

If start A – start B < 0 : <
If start A – start B = 0 : [
If start A – start B > 0 && start A - end B < 0 : #
If start A – end B = 0 :]
If start A – end A > 0 : >

5. Explicit media structure

```
Objects
   |
   |------football player-- |---goalkeeper
   |                        |
   |                        |---field player-|----defender
   |                                         |-----forward
   |                                         |-----player
   |
   |
   |------event ------------|--- throw-in
   |                        |--- tap ball
   |                        |--- goal kick
   |                        |--- overhead kick
   |                        |--- control ball by thigh
   |                        |--- control ball by chest
   |                        |--- header
   |                        |--- clean catching
   |                        |--- save
   |                        |--- dribbling
   |                        |--- intercept
   |                        |--- training
   |                        |--- easy play
   |                        |--- scored goal ---|---player shoots the ball
   |                                            |---goalkeeper misses the ball
   |                                            |---the ball in net
```

6. Integration of syntactic and semantic information

Frame No.	*Objects*	*Events*
1	Player	Throw-in
2	Player	Throw-in
3	Player	Throw-in
…	…	…

4 Semantic Multimedia Database Information Retrieval

4.1 Retrieval Language

Semantic multimedia databases require retrieval facilities to extract individual multimedia portions from the documents. Retrieval systems require a specification language with which the required multimedia data are described. A logical query language will be developed to express queries requiring multimedia accesses.

4.1.1 Basic requirement of the language

The language needs at least to be able to describe the following information:

1. Frame type/id: video/audio/image/…
2. Description: author/creation_date/…/entity of interest/…
3. Spatial: up/down/right/left/before/after/inside/outside/…:
4. Temporal: before/after/duration/…
5. Boolean: AND/OR/NOT
6. Area-field: Events/meeting/…/Actions/applause/…
7. Keyword: text/?/*/…

4.1.2 Query a semantic database

To know what kind of queries will be asked is the key to design the query language. The semantic multimedia database queries may be clarified into the following categories (It will be shown by a couple of examples):

1. Media Content (without media structure)

   This kind of query is like normal SQL. The following only shows the typical query that may be raised.

   Network
   What is the relationship between whale and mammals?
   Find relationship
   From video
   Where relationship (whale, mammals)

   Spacial
   Who stood left of Clinton during his inauguration from the video #108?
   Find who
   From video #108
   Where person *left* Clinton
   And Events = inauguration

   Temporal
   What happened before the chairman’s lecture?
   Find Events
   From video
   Where *before* the chairman’s lecture

Object
Find all the title of Beethoven's music.
Find title
From all audio
Where composer = "Beethoven"

2. Media content (with media structure)

This is the most important query for the multimedia databases. Because the users need to retrieval the exact media data (the media frame here) that they want. For example:

Events
Find all clips with goal scoring from World Cup 98.
Find frame No.
From the World Cup98 video
Where Events = "Goal"

Network
Show clip to prove whale is kind of mammal.
Find frame No.
From video
Where is_kind_of (whale, mammal)

Spacial
Show the person who stood left of Clinton during his inauguration from the video #108?
Find frame No.
From video #108
Where the person *left* Clinton
And Events = inauguration

Temporal
What happened before the chairman's lecture?
Find frame No.
From video
Where Events (*before* Events = "the chairman's lecture")

Object
Find all the frame No. of Beethoven's music.
Find frame No.
From all audio
Where composer = "Beethoven"

3. Advanced query

A more advanced query would be one which combine the above two kind of queries into. For example:

Find all the talks of whose declared to have improper relationships with Bill Clinton.
Find frame No.

```
From      the video
Where     Events = talk
And       person = ( Find      who
                     From      the video
                     Where     improper_relationship (Clinton)
                   )
```

4.2 Indexing

4.2.1 General

In COSMOS, the semantic data is stored in file format. An indexing structure to store multimedia content should be defined. The index is the most important map for locating semantic information in a SMDB. The multimedia raw data can be represented using m-frames [3].

With the m-frames, the information we are interested in is represented by a collection of three m-frames. They are: (1) Description m-frames which describes the entity of interest, (2) Events m-frames which model the events that are associated with the entity of interest, and (3) Action m-frames which model the constituent actions of the events modelled in the Events m-frames. Most segments of information are difficult to extract from single frames of video and audio (they have meaning over time and are also often meaningless when taken out of context) and it is not always possible to attribute events or actions based on a single frame. Therefore, each entity in m-frames will be given a set of frames (e.g. 104-151) to locate in the multimedia database. For example, if we have an ‘in the Andalusian costume’ segment at “Prince Andrew”: 104-151 and an ‘at a ball’ segment at “Prince Andrew” 33-145, Then we get “Prince Andrew”: 104-145 (=$[104, 151] \cap [33, 145]$) which has content of Prince Andrew in the Andalusian costume at a ball.

4.3 User Interfaces

Graphical user interfaces (GUI) can simplify human-machine interaction. A user visual query interface needs to be developed on top of the proposed language.

4.3.1 Steps of posing a query

To pose the content-based query by a visual interface will go through the following steps [9]:

1. Formulation: In this stage, users input where/what to search for. It is important that what the computer interprets the input should be exactly the same as what users want. When users want variants to be accepted to pose a flexible query, the user interface should make it clear how variants are handled.
2. Action: Searches may be started by a Search button to initiate the search and then wait for the results or by a method called “dynamic queries” to get the answer step by step.
3. Review of results: Search interfaces should provide helpful messages to explain search results and to support progressive refinement in addition to contents, sequencing of documents.
4. Refinement: it should support successive queries.

4.3.2 Design of the interface

The main points in the design of the user interfaces should consider the following:

1. To determine the appropriate information content to be communicated.
2. To represent the essential characteristics of the information, e.g. temporal characteristics.
3. To coordinate different media and assembling techniques within a presentation, e.g. audio/video.
4. To provide interactive exploration of the information presented, e.g. browse.

The user friendliness, the main property of the interface, will be evaluated based on the following aspects:

1. Easy to learn instructions
2. Context-sensitive helps functions
3. Easy to remember instruction rules
4. Effective Instructions and aesthetics.

Shneiderman has an excellent description about the diversity of the users when designing the interface [10]:

> "First-time users need an overview to understand the range of services … plus buttons to select actions. Intermittent users need an orderly structure, familiar landmarks, reversibility, and safety during exploration. Frequent users demand shortcuts or macros to speed repeated tasks and extensive services to satisfy their varied needs."

To design the presentation of query result, issues like content selection, media and presentation technique selection and presentation co-ordination must be considered. Individual functions must be placed together in a meaningful fashion. This occurs through alphabetic ordering or logical grouping.

5 References

1. R. Aston and J. Schwarz (1994) *Multimedia: gateway to the next millennium*, AP Professional.
2. H. W. Agius and M. C. Angelides. (1997) "Integrating logical video and audio segments with content-related information in instructional multimedia systems", *Information and Software Technology*, 39, 679-694.
3. H. W. Agius and M. C. Angelides (1999) "Developing knowledge-based intelligent multimedia tutoring systems using semantic content-based modelling", *Artificial Intelligence Review*, Vol. 13, No. 1, pp. 55-83.
4. S. T. Campbell and S. M. Chung (1996) "Database Approach for the Management of Multimedia Information", *Multimedia Database Systems Design and Implementation Strategies*, Kluwer, Boston.
5. V.S. Subrahmanian and J. Sushil (1996) *Multimedia Database Systems, Issues and Research Directions,* Springer.
6. J. A. Larson (1995) *Database Directions from Relational to Distributed, Multimedia, and Object-Oriented Database Systems*, Prentice Hall.
7. M. ADIBA (1996) "Storm: An Object-Oriented Multimedia DBMS", *Multimedia Database Systems Design and Implementation Strategies*, Kluwer, Boston.
8. M. C. Angelides and S. Dustdar. (1997) *Multimedia Information Systems*, Kluwer, Boston.
9. B. Shneiderman (1997) "A User-Interface Framework for Text Searches", *D-Lib Magazine,* January 1997.
10. B. Shneiderman (1992) *Designing the User Interface: Strategies for Effective Human-Computer Interaction: Second Edition,* Addison-Wesley, Reading, MA.
11. K. Fromm et al. (1995) *Careers in Multimedia*, Ziff-Davis, California
12. T. Feldman (1994) *Multimedia*, Blueprint, London.
13. F. Fluckiger (1995) *Understanding Networked Multimedia: Applications and Technology*, Prentice Hall, London.
14. R. Steinmetz and K. Nahrstedt (1995) *Multimedia: Computing, Communications and Applications,* Prentice Hall.
15. N. Jerke and etc., (1997) *Visual Basic 5 Client/Server How To*, Waite Group Press
16. D. D. Roure and W. Hall (1997) "Distributed Multimedia Information Systems", *IEEE MultiMedia*, Oct-Dec.
17. R. Goyal (1998) "Improving the performance of TCP over the ATM-UBR service", *Computer Communications*, 21, 898-911.
18. J. Vince (1995) *Virtual Reality Systems*, Addison-Wesley, England.
19. S. Dustdar and R. Huber (1998) "Group Decision Making on Urban Planning Using Desktop Multimedia Conferencing", *Multimedia Tools and Applications*, 6, 33-46.
20. T. Wen Jiin and L. Suh Yin (1998) "Dynamic Buffer Management for Near Video-On-Demand Systems", *Multimedia Tools and Applications*, 6, 61-83.
21. W. J. Liao and V. O.K Li. (1997) "The Split and Merge Protocol for Interactive Video-on-Demand, *IEEE MultiMedia*, Oct-Dec, 51-62.
22. A. Cristina, T. C Andrew, et al. (1998) "A survey of QoS architectures", *Multimedia Systems*, 6, 138-151.

6 Appendix A – Research notes

Jamec A. Larson (1995) *Database Directions Grom Relational to Distributed, Multimedia, and Object-Oriented Database Systems*, Prentic Hall.

Multimedia DBMS also need to manage a variety of new data types, including text, image, audio, and video data types. These new media types introduce special problems, including large data objects, continuous temporal data objects, and the problems of synchronisation of multiple streams of temporal data such as audio and video.

Applications can benefit from multimedia databases, including:
Medical information systems. Medical databases contain medical imaging (X-ray, CAT scan), monitoring information (EKG recordings), as well as photographs of characteristic physical symptoms.
Engineering information systems, including both manually generated and computer-generated blueprints, sketches diagrams and illustrations. Photos documenting construction stages are also useful.
Office and library information systems. Information on paper can be scanned and stored in an image database. Nonpaper objects can be photographed or video taped and stored in a multimedia database.
Consumer catalogues. These databases not only contain pictures and textual descriptions, but may also contain verbal commentary and video demonstrations of goods and services.
Training and education. Databases can contain video clips demonstrating how things work, how to repair things, and how to assemble things.
Geographic databases. Maps of all kinds, as well as aerial and satellite photographs, can be stored and analysed by geographic database systems.
Reference works, including encyclopaedias containing news clips, audio clips, and digitised photographs.

Ralf Steinmetz and Klara Nahrstedt (1995) *Multimedia: Computing, Communications and Applications*, Prentice Hall, Upper Saddle River.

A narrow definition, a multimedia system is any system that supports more than a single kind of media. It is nascence. It only makes sense to have a definition in computer world.

Michael Vazirgiannis (1996) "An Object-Oriented Modelling of Multimedia Database Objects and Applications", *Multimedia Database Systems Design and Implementation Strategies*, Kluwer, Boston.

[Overview] an object oriented data base model (MOAP – Multimedia Object and Application Model) that aims at representation of multimedia objects and applications is proposed. The important feature of it is the approach for integrated modelling of the multimedia objects as well as of the applications.

[Key points] the main issures which multimedia database management researchers/designers need to face include:
development of sophisticated conceptual models which are rich in their semantic capabilities to represent complex multimedia objects and express their synchronisation requirements. A transformation from models to database scheme is then needed. The object retrieval algorithms is needed to specify.

Designing multimedia query languages which are not only powerful enough to handle various manipulation functions for multimedia objects but also simple in handling user's interaction for these functions.
Designing powerful indexing and organisation techniques for multimedia data.
Multimedia data base modelling: Graphical models, Petri Net models and object oriented models.

A key issue in the representation of multimedia applications is the description of spatial and temporal composition of objects participating in the application.

Stacie Hibino, Elke A. Rundensteiner (1996) "A Visual Multimedia Query Language For Temporal Analysis of Video Data", *Multimedia Database Systems Design and Implementation Strategies*, Kluwer, Boston.

[Overview] the focus of the research is to exploit the temporal continuity and combined spatio-temporal characteristics of video data for the purpose of video analysis. The primary contributions include 1) a visual information seeking (VIS) approach to video analysis, 2) the temporal visual query language (TVQL) for specifying relative temporal queries and for facilitating temporal analysis, 3) a transformation function for deriving temporal diagrams, 4) a description of the automated maintenance of interdependencies between the temporal position query filters, and 5) a formal annotation model for abstracting temporal, spatial, and content-based characteristics from video data.

[Key points] Visual query approach will be more effective than a forms-based language for the purpose of searching for temporal trends in the video data.

Databases, which handle temporal media, tend to focus on semantic or text-based queries as well as on locating information rather than analysing it. the drawback of this approach is that it does not take advantage of the temporal and/or spatial characteristics inherent in the media.

The advantages of using annotations are that 1) they allow users to abstract both temporal and spatial information from the data, 2) they simplify analysis by reducing the amount of information to be processed, and 3) when layered on top of the original data, they allow users to preserve context without corrupting the original data.

Cyril Orji (1996) "Multimedia DBMS – Reality or Hype? ", *Multimedia Database Systems Design and Implementation Strategies*, Kluwer, Boston.

[Overview] This is an overview of the whole book and presentation of the author's general points of view. These issues include proper and accurate characterisation of multimedia data, multimedia data integration, and multimedia query language and processing. Others include multimedia data management and storage issues, and multimedia retrieval and indexing.

[Key points]
Multimedia database management – there is a strong need to manage multimedia meta-data and presentation with a DBMS. Formal development of a data model would facilitate the construction of a multimedia DBMS. One of the challenges for the evolution of future DBMS is their ability to handle multimedia data in an integrated and efficient way.

Multimedia integration – the integration makes the handling of multimedia database management system more efficient since the storage, retrieval, buffering, and playout management is performed by one consistent system.

Query language and processing – a temporal visual query language and a specification language for multimedia query are proposed for data processing.

Multimedia storage issues – different of multimedia servers are surveyed.

Multimedia retrieval and indexing – two ways on multimedia retrieval and indexing are proposed, Model for Interactive Retrieval of Videos and Still Images and MB^+-Tree.

Scott T. Campbell and Soon M. Chung (1996) "Database Approach for the Management of Multimedia Information", *Multimedia Database Systems Design and Implementation Strategies*, Kluwer, Boston.

[Overview] a multimedia database system needs to extend the traditional query response role and provide multimedia specific data modelling, delivery modelling, access modelling and storage modelling. A novel temporal query script methodology is developed to support the incorporation of the role of media servers with isochronous multimedia data delivery capabilities.

[Key points]
Underlying idea – the query script that is a tool to enable optimise the retrieval and delivery of the multimedia streams to clients creates a novel client-database interface that allows the database system to better manage system resources through multimedia data delivery scheduling. Query script's temporal modelling ability also helps database systems maintain the separation between the database system's data model and the application's data model.

Current multimedia data management – local storage, media server, binary large objects and object-oriented methodologies are used based on multimedia information temporal and synchronisation characteristics.

Multimedia databases – the client application's MHEG document defines the multimedia content's presentational and relationship information. The MMDBMS locates the content and in conjunction with the client interface, initiates and performs a stream delivery. It can also use its knowledge about the content structure for optimal delivery.

A multimedia query language requires a rich set of features to support multimedia content specification and retrieval.

Michel ADIBA (1996) "Storm: An Object-Oriented Multimedia DBMS", *Multimedia Database Systems Design and Implementation Strategies*, Kluwer, Boston.

[Overview] STORM, a multimedia DBMS, is developed based on the object-oriented approach on top of O_2. It provides facilities for describing and storing time-based objects, and for building sequential or parallel presentations of multimedia data. Two issues are addressed: (1) modelling and management of time-based data, and (2) capabilities, languages and interfaces for building, querying and updating multimedia data.

[Key points]

HyTime is included in SGML and provides a collection of abstract semantic constructs associated with syntactic conventions. OODBMS technology can bring a lot of benefit to future multimedia document management: modelling concepts, concurrency control, high-level query facilities, etc.

The extension of MDBMS should provide a way to build multimedia presentations expressing temporal and synchronisation constraints between objects. One database object can have different presentations. Presentations are themselves considered as database objects.

Each object appearing in a presentation, a pair of temporal elements, **duration** and **delay** each of them being either free or bound and constituting the **Temporal Shadow TS** of the object.

The extensions of O2SQL language concern principally: (1) query on temporal attributes, (i.e. the Temporal Shadow); (2) query on collections of time-based objects with specific synchronisation; (3) query on correlated lists for continuous time-based data.

Rune Hjelsvold, Roger Midtstraum, and Olav Sandsta (1996) “Searching and Browsing a Shared Video Database”, *Multimedia Database Systems Design and Implementation Strategies*, Kluwer, Boston.

[Overview] VideoSTAR (Video Storage And Retrieval) has been developed to show issues related to searching and browsing a shared video database. Video databases architectures, video algebra operations, video querying, and video browsing are discussed based on characteristics of video information and video database applications.

[Key points]
Audio/video data and related meta-data, in contrast to traditional data types, may have temporal relationships to each other.

A key question for context handling is how meta-data, especially structural and content data, can be shared in a consistent way when media data are shared or parts of video documents are reused in other documents.

Content-based retrieval can be done by using advanced feature extraction/matching tools or by providing tools and methods that can enhance manual indexing

The three-level VideoSTAR architecture consists of specialised repositories, generic data model and video database API. The integrated video tool environment consists of a video player, a tool manager, and tools for searching, browsing, and registration of meta-data.

Marios C. Angelides and Schahram Dustdar. (1997) *Multimedia Information Systems*, Kluwer, Boston.

[Overview] it comprehensively defines multimedia information systems and its emerging architecture. It is a essential reading for all people who are interested in multimedia systems.

[Key points]
Multimedia information systems are the profusion of text, graphics, animation, audio, still and full-motion video and interactivity on the computer.

Research challenges include real-time multimedia data transfer and synchronisation, virtual reality, large storage devices, multimedia operating systems, object-oriented tools and multimedia databases.

Research and development efforts in multimedia information systems fall mainly into two groups, standalone multimedia and networked multimedia. Networked multimedia information systems are computer based, real-time and interactive IS which combine text, image, audio and video over a networked infrastructure.

There seem to be two chief partner technologies that are being implemented in the superhighways of present, Broadband ISDN and Asynchronous Transfer Mode (ATM).

Authoring systems tend to emphasise interactive navigation, database access, and preparation of productions for mastering and/or distribution. ‘Spatial’ framework and ‘procedures’ and ‘constraints’ techniques are powerful in developing authoring systems.

Fluckiger, F. (1995) *Understanding Networked Multimedia: Applications and Technology*, Prentice Hall, London.

[Overview] it provides a comprehensive overview of networked multimedia applications and their underlying technology, including sets the scene, existing and future multimedia applications, requirements placed by remote applications, data communication technologies and data compression and coding. The book contains solid treatments of asynchronous transfer mode, buffering, traffic analysis, traffic shaping, and scheduling. It also discusses system software trends.

[Key points]
The financial and technical future of the information superhighway initiatives is not clear. In the meantime, the Internet provides a laboratory for the future information society.

Lines have progressively become blurred between conventional circuit- and packet- based videoconference systems.

Multicasting is a key network feature required by many multimedia applications. Bi-directional connections that allow interactivity are another key requirement of most multimedia applications.

LAN switching, fast Ethernet, and ATM are competing technologies to give high-speed support to local-area multimedia applications. The cost of the host interface will be the key in this competition. ATM – when available end to end and with its complete range of services – is in theory the technology of choice for multimedia applications.

The Internet Protocol or an equivalent network protocol will keep the role of unifying layer in end-systems for at least a decade, regardless of the underlying transport technology.

Feldman, T. (1994) *Multimedia*, Blueprint, London

[Overview] Tony Feldman addresses the significance of multimedia in a general sense and examines the impact of multimedia on education, training, business and professional sectors, leisure and entertainment, publishing, bookselling and library services.

[Key points]
Multimedia with interactivity and bearing different media on the issue of clarifying, communicating and informing is an obvious candidate for both training and educational applications.

The basis of the linking is to insert pointers within searchable text fields which effect the image retrieval once the text field retrieval has taken place. This approach using linkages between text fields in the database records looks like being the most promising model for multimedia database design.

Four areas selected for the multimedia future are: high definition television, networked multimedia, handheld multimedia and virtual reality.

Harry W. Agius and Marios C. Angelides. (1997) "Integrating logical video and audio segments with content-related information in instructional multimedia systems", *Information and Software Technology*, 39, 679-694.

[Overview] an architecture for instructional multimedia systems that are interactive and structured is provided to reduce the information overload and disorientation through the learning process. A content-based multimedia application is built in the development of MAT, Multimedia Animal Tutor, to illustrate the benefits of semantic modelling approach.

[Key points]
In interactive-structured systems the student-user is actively involved in the teaching-learning interaction, and appropriately learns where and why they are going right or wrong.

The concept of multimedia frames is developed so that logical video and audio sequences could be indexed and integrated with content-related information.

Multimedia frames integrate content-related information about logical video and audio segments that is pertinent to the pedagogy of the instructional multimedia system.

An architecture in which domain, tutor, and student modules together make up the knowledge-based multimedia support environment of an instructional multimedia system is the central point to build this content-based multimedia application.

Wen Jiin, T. and Suh Yin, L. (1998) "Dynamic Buffer Management for Near Video-On-Demand Systems", *Multimedia Tools and Applications*, 6, 61-83.

[Overview] the number of concurrent on-demand services supported by the video server is often limited by the I/O bandwidth of the storage systems. A discrete buffer sharing model is suggested to tackle with the problem, which uses batching and buffer sharing techniques in video servers to support a large number of VOD services.

[Key points]
Two operations, splitting and merging, can be used to enable a video server to fully utilise system resources such as buffers and disk bandwidths.

Imprecise video viewing means that certain degree of quality loss is allowed during the video playback the quality loss can be resulting from inserting advertisements or skipping some video contents during the playback. Three shrinking strategies based on this which include backward shrinking, forward shrinking and two way shrinking, are explored to reduce buffer requirements in the system.

Wanjiun, Liao and Victor O.K. Li. (1997) "The Split and Merge Protocol for Interactive Video-on-Demand, *IEEE MultiMedia*, Oct-Dec, 51-62.

[Overview] a new protocol, Split and Merge (SAM), is introduced to reduce the per-user video delivery cost in VoD system. It allows true VoD and multiple users may be batched and share the same video stream.

[Key points]
A true video-on-demand (VoD) system should let users view any video program, at any time, and perform any VCR-like user interactions.

Sharing the same video should be transparent to users while allowing true user interactivity.

VoD will be one of the most important commercial applications of distributed multimedia systems. It provides an electronic video rental service, which gives users the ultimate flexiblity in selecting any video programs, at any time, and in performing any VCR-like user interactions.

To achieve commercial success, VoD must be priced competitively with existing video rental services.

SAM refers to the split and merge operations incurred when each user performs user interactions. These operations enable any kind of user interactions. SAM starts by serving customers in a batch. When a user in a batch initiates a user interaction, the protocol splits off the interactive user from the original batch and temporarily assigns that user to a new video stream. With a dedicated video stream, the user can perform any interactions desired. As soon as the user interaction terminates, the system merges this user back to the nearest ongoing video system.

Siu-Wah Lau, John C.S. Lui, Leana Golubchik (1998) "Merging video streams in a multimedia storage server: complexity and heuristics", *Multimedia Systems*, 6, 29-42.

[Overview] the stream-merging approach is proposed to reduce the I/O demand to the VoD server through data- and resource-sharing techniques. In the paper, the author formalises a static version of the stream-merging problem, derive an upper bound on the I/O demand of static stream merging, and propose efficient heuristic algorithms for both static and dynamic versions of the stream-merging problem.

[Key points]
The cost/benefit trade-off considered is the balance between the reduction in I/O bandwidth demand and the amount of storage overhead required for each video, i.e., we should only apply the stream-merging approach to a request for a given video when the benefit due to the I/O bandwidth demand reduction is greater than the cost of the storage overhead.

Dustdar, S. and Huber, R. (1998) "Group Decision Making on Urban Planning Using Desktop Multimedia Conferencing", *Multimedia Tools and Applications*, 6, 33-46.

[Overview] to use desktop multimedia conferencing for group decision making on Internet is possible now. The paper review the design, hardware and software requirements and organisational issues in a desktop multimedia conferencing system through a case study on urban planning using desktop multimedia conferencing on the Internet.

[Key points]
Preparation and realisation of desktop multimedia conferencing has two aspects, the technical set-up procedure and organisational issues.

Research on desktop multimedia conferencing and its application in decision-making processes is interlinked with other information processing, communication and co-ordination activities. The tools need to be integrated into organisational information systems such as word processing, project management software and spreadsheet applications.

David De Roure and Wendy Hall (1997) “Distributed Multimedia Information Systems”, *IEEE MultiMedia*, Oct-Dec.

[Overview] University of Southampton is well known for its work in open hypermedia systems and for developing the Microcosm hypermedia system. The early work on hypermedia systems is extended to distributed information systems and digital libraries. Agent technology is used to create, manage, and customise links in a distributed link service environment that emerged from the Microcosm architecture.

[Key points]
To extend streaming work to HyperRadio is built on the idea that a program is actually a tour through available resources, where users can interactively follow links to other resources that interest them.

Cristina A., Andrew T. C., and Linda Hauw (1998) “A survey of QoS architectures”, *Multimedia Systems*, 6, 138-151.

[Overview] the paper examines the state-of-the art in the development of QoS architectures. Multimedia systems designers should adopt an end-to-end approach to meet application-level QoS requirements.

[Key points] To date, most of the work has been within the context of individual architectural layers in the area of quality-of-service (QoS), such as the distributed system platform, operating system. Much less progress has been made in addressing the issue of overall end-to-end support for multimedia communications.

All architectures of QoS provide services should be based on both hard (guaranteed service) and soft (best effort) QoS guarantees.

A generalised QoS framework should be motivated by five design principles: that is, the principles of transparency, integration, separation, multiple time scales and performance.