浙江大学计算机学院
数字媒体与网络技术

Digital Asset Management
数字媒体资源管理

3. Multimedia Database Technologies

任课老师：张宏鑫
2014-10-14

Outline

- What is Multimedia?
- Why do we need multimedia?
- Classification of Media Types
- Types of Media
- Characteristics of MM Data
- Definition- MMDBMS
- MMDBMS Characteristics

What's Multimedia?

What's Multimedia?

- Multi => Many
- Media:
  - A means to distribute and represent information: Text, graphics, pictures, voice, sound and music..
    - Perception media (how do humans perceive information?)
      - Audio/visual media
    - Representation media (how is information encoded?)
      - ASCII, JPG, MPEG, PAL.
    - Presentation media (medium used for output/input)
      - Input/output media (keyboards, papers)
    - Storage media (Where is information stored?)
      - Magnetic disk, optical disk

What's Multimedia?

- Multimedia:
  - –To distribute and present information coded as
    - Text, Graphics, animation, audio and video..
  - –By
    - Computer, TV, phone, etc.

What is Multimedia?

Multimedia (MM) has pervaded the worlds of entertainment, education and business.

- Different view to define MM:
  - Technologist: MM system
  - Games Player: Not only see the blood but hear the screams
  - Artist: The potential to create life

Definitions are contextual and depend on the perspective of the user.

What is Multimedia?

- Multimedia: **a working definition**
  - A combination of two or more categories of information having different transport signal characteristics
  - Typically, one medium is a continuous medium while another is discrete
  - Image, audio, video and graphics are usually the examples of media

What is Multimedia?

Some serious definitions:

"From a user's perspective - multimedia enables computer *information* to be *represented* through audio, video, text, images, graphics and animation."

"Multimedia is defined as an interactive computer-mediated presentation that included at least two of the following elements: text, sound, still graphic images, motion graphics and animation."

What is Multimedia?

- Most definitions agree that there are a number of key aspects of multimedia:
  - Multimedia involves "Many Media"
    - THE MEDIA DOMAIN
  - Multimedia involves Computers
    - THE SYSTEMS DOMAIN
  - Multimedia enhances the presentation and communication of information
    - THE APPLICATION DOMAIN

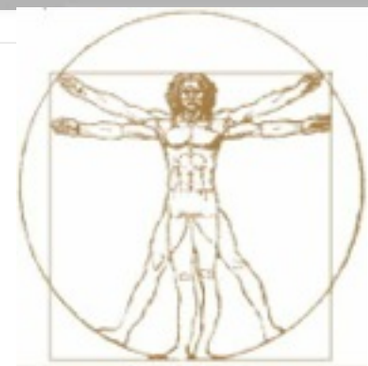
Multimedia: Computers v.s. Humans

Machine-Centred View

- People are:
  - vague (含糊)
  - disorganized（无序）
  - distractible （不专注）
  - emotional （情绪化）
  - illogical （非逻辑）
- Machines are:
  - precise
  - orderly
  - undistractible
  - unemotional
  - logical

VS

Human-Centred View



- People are:
  - multimedia
  - creative
  - attentive to change（着力改变）
  - resourceful （富想象力）
- Machines are:
  - rigid & text oriented
  - dumb
  - insensitive to change
  - unimaginative

Multimedia: Computers → Humans

- Machines cannot do many *mundane* things which can be performed routinely by humans （机器不擅长）
  - Natural language processing
  - Recognizing objects in images, vision

- Humans (untrained) cannot do many *clever* things which can be performed routinely and efficiently by machines
  - Numerical and symbolic computation, theorem proving
  - Searching information backwards
  - Chess playing

Why do we need multimedia?

- A computer database supports following basic data types:
  - numeric (e.g. student number); character string (e.g. surname, course title);
  - alphanumeric (e.g. course code, telephone number);
  - Boolean (e.g. female or male); date (e.g. date of birth);
  - text (e.g. description of a particular course, address).
- Inadequate to hold documents such as letters, application forms or image (student photograph).
- Need a document image processing system including:
  - a scanner or digital camera to capture (input) the images;
  - image database to store and retrieve the image.
- Further enhancement to handle other types of information (voice)
- A short video clip

Classification of Media Types

- Media types can be divided into two groups:
  - **Temporal** (Continuous media)
    - Time or more exactly time-dependency between information items, is part of the information itself.
    - dynamic, time-based, continuous
  - **Non-temporal** (Discrete media)
    - Time is not part of the semantics of the media.
    - static, non-time-based, discrete

Classification of Media Types

- Temporal media types:
  - audio, video, music, animation
- Non-temporal media types:
  - text, graphics, images

Media Type: Text

Media Type: *Text*

Different Representation

- ASCII
- ISO character sets
- Marked-up text
- Structured text
- Hypertext

Different Operations

- Character operations
- String operations
- Editing
- Formatting
- Pattern-matching
  - Searching
- Sorting
- Compression
- Encryption

Media Type: Text

- Plain Text
  - American Standard Code for Information Exchange (ASCII)
  - Each ASCII code uses seven bits for; 8 bits are used to store each character with the extra bit being 0.
  - ***Unicode and UTF-8***

- Structured Text
  - SGML, XML, HTML
  - Latex,
  - Office Document Architecture (ODA)

Media Type: Images

Media Type: Image

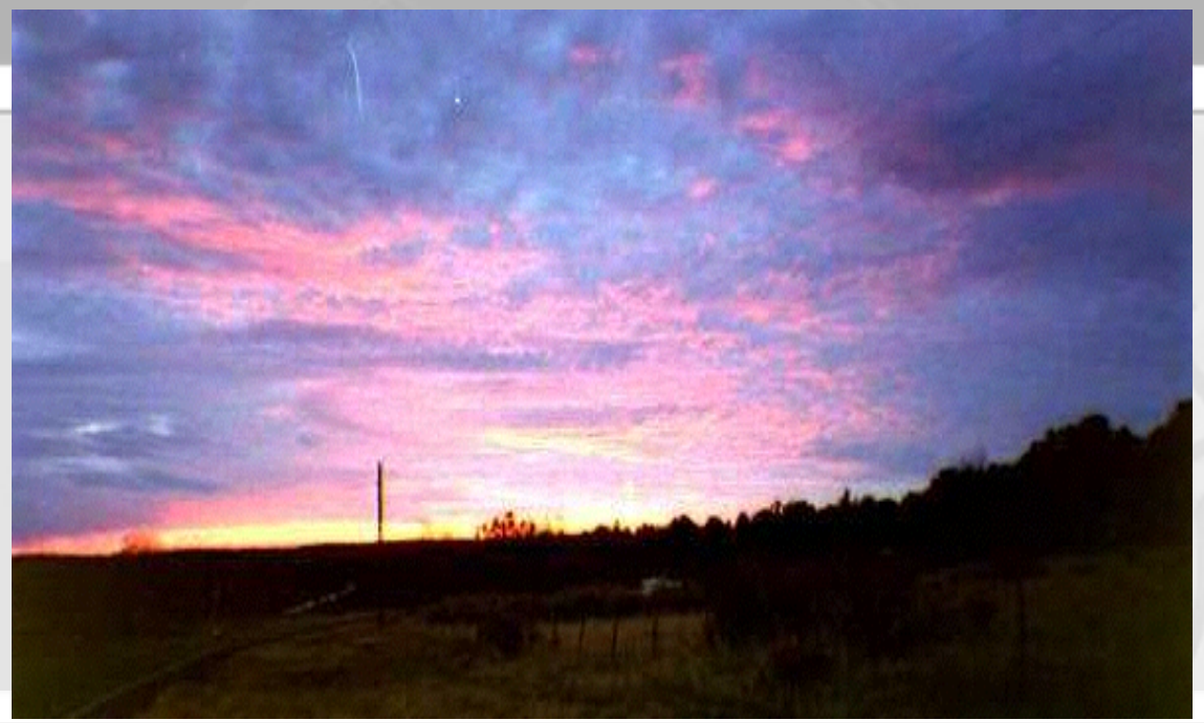
Different Representation

- Color model
- Alpha channels
- Number of channels
- Channel depth
- Interlacing
- Indexing
- Pixel aspect ratio
- Compression

Different Operations

- Editing
- Point operations
- Filtering
- Compositing
- Geometric Transforms
- Conversion

Media Type: Images



- Several Issues
  - Is the image as high-quality as I need? (**Representation**)
  - Is the image efficiently stored and transmitted? (**Compression, transmission**)
  - Can I find the similar images in a database (**Retrieval engine, Internet**)? How? (**Analysis, processing**)

Media Type: Graphics

Media Type: *Graphics*

Different Representation

- Geometric models
- Solid models
- Physically based models
- Empirical models
- Drawing models
- External format for mode

Different Operations

- primitive editing
- Structural editing
- Shading
- Mapping
- Lighting
- Viewing
- Rendering

Media Type: Analog Video

Media Type: Analog Video

Different Representation

- Frame rate
- Number of scan lines
- Aspect ratio
- Interlacing
- Hypertext
- Quality
- Component/Composite

Different Operations

- Storage
- Retrieval
- Synchronization
- Editing
- Mixing
- Conversion

Media Type: Digital Video

Media Type: *Digital Video*

Different Representation

- Sampling
- Sampling rate
- Sampling size and quantization
- Data rate
- Frame rate
- Compression
- Support for interactivity
- Scalability

Different Operations

- Storage
- Retrieval
- Synchronization
- Editing
- Effects
- Conversion

Media Type: Videos

- Issues
  - The format of the video? (Compression)
    - MPEG (.mpg), AVI (.avi), Realplayer (.rm).
  - What do I need to stream video across the network?
    - Watch video online.
  - Are the different media synchronized?
    - Multimedia description and characterization.
  - How can I describe a video retrieval request like a query?
    - Video processing, database management, retrieval

Media Type: Digital Audio

Media Type: *Digital Audio*

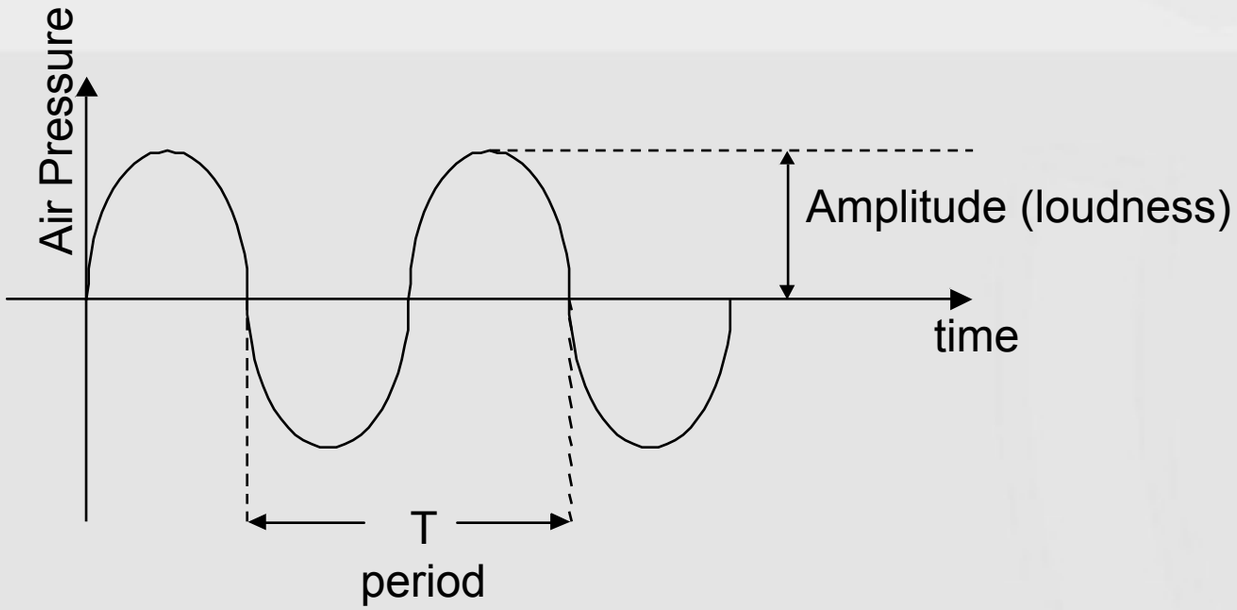
Representation

- Sampling frequency
- Sampling size and quantization
- Number of channels
- Interleaving
- Negative samples
- Encoding

Operations

- Storage
- Retrieval
- Editing
- Effects and filtering
- Conversion

Media Type: Audio



- The frequency range is divided into:
  - Human hearing frequency: 20 Hz to 20 KHz,
  - Infra-sound: 0 to 20 Hz,
  - Ultrasound: 20 KHz to 1 GHz,
  - Hypersound: 1 GHz to 10 THz

- **Sound** consists of pressure waves that move through a compressible medium.
- The ***frequency*** of a sound is the reciprocal value of the period (wavelength).
- Wavelength is the distance between identical points in the adjacent cycles of a waveform signal

  Terahertz (THz) is used; 1 THz = 1,000,000,000,000 cycles per second

Media Type: Music

Media Type: *Music*

Different Representation

Operational/Symbolic

MIDI

(Musical Instrument Digital Interface)

Different Operations

Playback and synthesis

Timing

Editing and composition

Media Type: Animation

Media Type: Animation

Different Representation

- Cel models
- Scene-based models
- Event-based models
- Key frames
- Articulated objects and hierarchical models
- Scripting and procedural models
- Physically based and empirical models

Different Operations

- Graphics operations
- Motion and parameter control
- Rendering
- Playback

Other Media Types

- Surface images
- Volume images
- Digital ink
- Speech encoding
- Speech synthesis
- Non-temporal video and animations

Characteristics of MM Data
-- challenges about managing MM data

- ***Huge Size***
  - multimedia objects are large in size (compared with traditional alphanumerical data) and are not readily accommodated by 'old' DBMSs
  - E.g.,
    - a video – each second contains, say, 30 frames; each frame may require, depending on video quality, several megabytes of storage.
    - a color picture of 1280 x 960 pixels using 32-bit color requires about 5MB of memory
      ≈ **4,915,200 = 1280 x 960 x 4 *bytes***
- Needs fast and powerful processors
- Large storage capacities
- Multimedia data have to move very quickly through the different components of the computer

Characteristics of MM Data
-- challenges about managing MM data

Data Volume

- Data dynamic in time inherently leads to a high data volume for single data elements (cf. Table 1):

MediaType	SampleFormat	DataVolume	TransferRate
Text	ASCII	1 MB / 500 pages	2 KB / page
Color Image	GIF, TIFF;	1.6GB / 500 images	3.2 MB / image
	JPEG	0.2GB / 500 images	0.4 MB / image
CD-music	CD-DA	52.8 MB / 5 min.	176 KB / sec.
Consumer video	PAL	6.6 GB / 5 min.	22 MB / sec.
High quality Video	HDTV	33 GB / 5 min.	110 MB / sec.
Speech	m-law, linear;	2.4 MB / 5 min.	8 KB / sec.
	ADPCM;	0.6 MB / 5 min.	
	MPEG audio	0.2 MB / 5 min.	

Characteristics of MM Data

Data Volume

- Concerning when designing a MMDB system:
  - MM applications: dealing with **huge amount of data** under **real time constraints**
    - Design of hardware, OS, and networks ...
  - Perform the **processing** on **the references to the values**.
    - E.g., video script editing.
  - Dynamic data
    - cannot be performed over references (e.g., copying) on one hand
    - cannot be executed in the standard way as for alphanumeric data because dynamic data exceeds the physical resources.

Characteristics of MM Data

Data Volume

- Specific form of dynamic data management
  - speed the process over time
    - s.t., at each distinct moment, only a limited amount of physical resources are needed.
  - This kind of dynamic operation heavily affects the behavior of a system, hence must not be transparent

Characteristics of MM Data

- The requirements regarding data volumes
  - appropriate referencing mechanisms to refer to MM data units
  - dynamic data management for very large objects

Characteristics of MM Data
-- challenges about managing MM data

- ***Quality of Service (QoS)***
  - multimedia applications differ from traditional DB applications with respect to **performance requirement**
  - multimedia applications, in general, require **high throughput and a constant delivery** of information

- e.g, real-time requirement for audio

note that lost audio samples are perceived much sooner by a human user than lost video frames:

the ear is more "sensitive" than the eye!

Characteristics of MM Data
-- challenges about managing MM data

- ***Quality of Service (QoS)***

- To offer a certain quality of service, the system has to **reserve** an adequate number of resources
- the higher the required service, the more resources the system has to reserve, and the more '**expensive**' the service will be.

Characteristics of MM Data
-- challenges about managing MM data

- ***Synchronization***
  - multimedia objects: composed of several components,
  - require synchronization of them
    - e.g., a film consists of moving pictures, speech, and subtitles
    - *It'll be odd if the lip movements of actors/actresses do not synchronize precisely with their voices and with the text of the subtitles.*

- Research results have shown that
  - **video/audio** or **video/video** synchronization is **less critical**, expressed in time constraints, than **audio/audio** synchronization!

Characteristics of MM Data
-- challenges about managing MM data

- ***Content-Based Retrieval***

- Hard to use traditional "exact-match" queries to retrieve multimedia objects:
  - interpretations of multimedia data is a rather **unexplored area**
  - multiple interpretations are, in general, possible (=> inexact match)
  - multimedia objects may contain multiple attributes of type audio, video, or text
- Obviously, content-based retrieval of multimedia objects is far more complex than Information Retrieval (IR) on text objects. (Note that IR is already more complex than traditional database accesses.)

Characteristics of MM Data

- ***Similarity-based Search***
  - Unlike traditional 'exact search' in relational database, users usually ask for similar objects based on their contents
    - Finding an image with similar color
  - A multimedia object may contain multiple features
    - Video may contain text, image, audio, etc.
    - Similarity search may be based on multiple features, i.e., integrating content and semantic features

Characteristics of MM Data
-- challenges about managing MM data

Temporal Aspects

- Certain multimedia data types like audio and video have a canonical mapping to real-world time (i.e., their time scale is absolute).

- When processing dynamic data, involving dynamic data types like audio or video, parallel tasks can occur. Also, the dynamics of a multimedia document are based on the temporal relationships between the constituent media types.

  E.g., consider an application to play back a video on a screen: it gets the audio simultaneously from a different device and allows for user interaction to control presentation (like, adding annotations to the video at certain points without interrupting the video presentation).

- There is also a need to support for media-specific synchronization

  e.g., playing back the video frames and the sound track of a movie requires fine-grained synchronization, such as lip synchronization.

Characteristics of MM Data
-- challenges about managing MM data

Temporal Aspects

- There is a need to model, store, and process temporal relationships between media components of a multimedia document or presentation, including synchronization mechanism
- Methods available for controlling parallelism of tasks :
  - **relative scheduling (e.g., one task triggers the other)**
  - **absolute scheduling (placing events on an absolute time scale)**
  - **combination of the above**

Characteristics of MM Data
-- challenges about managing MM data

Temporal Aspects

- Requirements
  - incorporation of **time-related** concepts into the data model
  - **non-transparent parallelism** for explicit control of parallel tasks
  - scheduling and (media-specific) synchronization mechanisms to provide for the description of temporal relationships and their enforcement

Characteristics of MM Data
-- challenges about managing MM data

Media Representation

- The basis data types like **alphanumeric ones are not appropriate** to reflect the structure of multimedia data
- **new built-in data types**
  – e.g, bitmaps or audio samples ...

- **type constructors** taking into account the temporal nature of multimedia data will be needed in some form, and **appropriate support for processing** these data types has to be provided

Characteristics of MM Data
-- challenges about managing MM data

Media Representation

- To be able to use **different formats** for the **same multimedia types** is crucial for multimedia data:
  - **different compression techniques**
    - for different applications, for different resources may need different formats, and each format should be convertible to another in principle
  - the internal representation may not be appropriate to be presented to the user, so special representations for different users to provide different views of the same data may be needed (either generated on the fly or stored persistently)

Characteristics of MM Data

Media Representation

- In SQL3 (or, SQL:1999), the repertoire of built-in types are extended: in addition to the types known from SQL2, there are
  - **Boolean type**
  - **Large character objects**
  - **Large binary objects**
- The latter two allow for character strings and bitstrings in the Gigabyte ranges, thus matching the storage requirements of certain MM applications.

Characteristics of MM Data

E.g., in Oracle 9i…

Name	Data type	Size	characteristics
BLOB	*Binary*	*4Kb in table space* *4GB in ext. space*	*random access;* *transaction support*
CLOB	*Character*	*4GB*	*random access;* *transaction support*
NCLOB	*National character sets*	*4GB*	*random access;* *transaction support*
BFILE	*Binary*		*Read only;* *external file*

Characteristics of MM Data

e.g., in Oracle 9i, after the following:

```
CREATE  TABLE  Grape
        (grape_name VARCHAR2(25) primary key, picture BFILE);
CREATE DIRECTORY “PHOTO_DIR” AS ‘C:\PICTURES’;
```

The user can then do insert:

```
INSERT INTO Grape(grape_name, picture)
    VALUES (‘chardonnay’, BFILENAME(‘PHOTO_DIR’, ‘chardonnay.jpeg’))
```

Characteristics of MM Data

Media Representation

- Requirements:
  - new built-in **data types** and **operations** for MM data
  - modular and efficient representation of **various formats**
  - data **representation** should be **transparent** to users
  - **different views** of the same data should be possible

Characteristics of MM Data
-- challenges about modeling MM data

- High level content abstraction is natural to the way humans think
- Effective modeling of MM data is critical
  - support for semantically rich conceptual contents
  - ability to represent diverse aspects of the data to be modeled
  - facilities for dynamic concept enrichment and expansion
  - incorporation of knowledge of low level data
  - isolation of the user from the low level representation and storage levels

Characteristics of MM Data
-- challenges about modeling MM data

- Low level representation of multimedia data, which encodes the physical reality, leads to the problem of huge amounts of data.
  - references to (as *abstractions* of) MM data are needed to process it efficiently to avoid copies;
  - more complex abstractions, like references enriched with more information than that needed for identification, can be used to index data to provide fast access.
- Another reason for introducing abstractions (provided by the user/system) is to allow the user to refer to the data in terms of abstractions which make up his model of the application domain.

Characteristics of MM Data
-- challenges about modeling MM data

- However, traditional (relational/object-oriented) data models fall short in representing MM data adequately

Data model features	Relational model	Object oriented model	"Multimedia" model
Structural model	Relations (tables)	Objects attributes	Continuous data streams
Behavioral model	SQL	Methods	Time-dependent operations
Constraints	Primary keys, foreign keys etc.	Referential integrity, pre-/post-conditions, etc.	QoS parameters

Characteristics of MM Data
-- challenges about modeling MM data

- MM Data Abstraction for Time-Dependent Data
  - to reflect the time-related characteristics
  - and requirements of continuous data should address the description, the processing of temporal relationships, and synchronization constraints.

Time-dependent data description:

- sequential-composition,
- parallel-composition

Two approaches for describing MM data compositions are:

- time-critical medium (e.g., audio) as the reference for the others
- an abstract temporal dimension ref. *common time line*

Characteristics of MM Data
-- challenges about modeling MM data

MM Data Abstraction for Time-Dependent Data

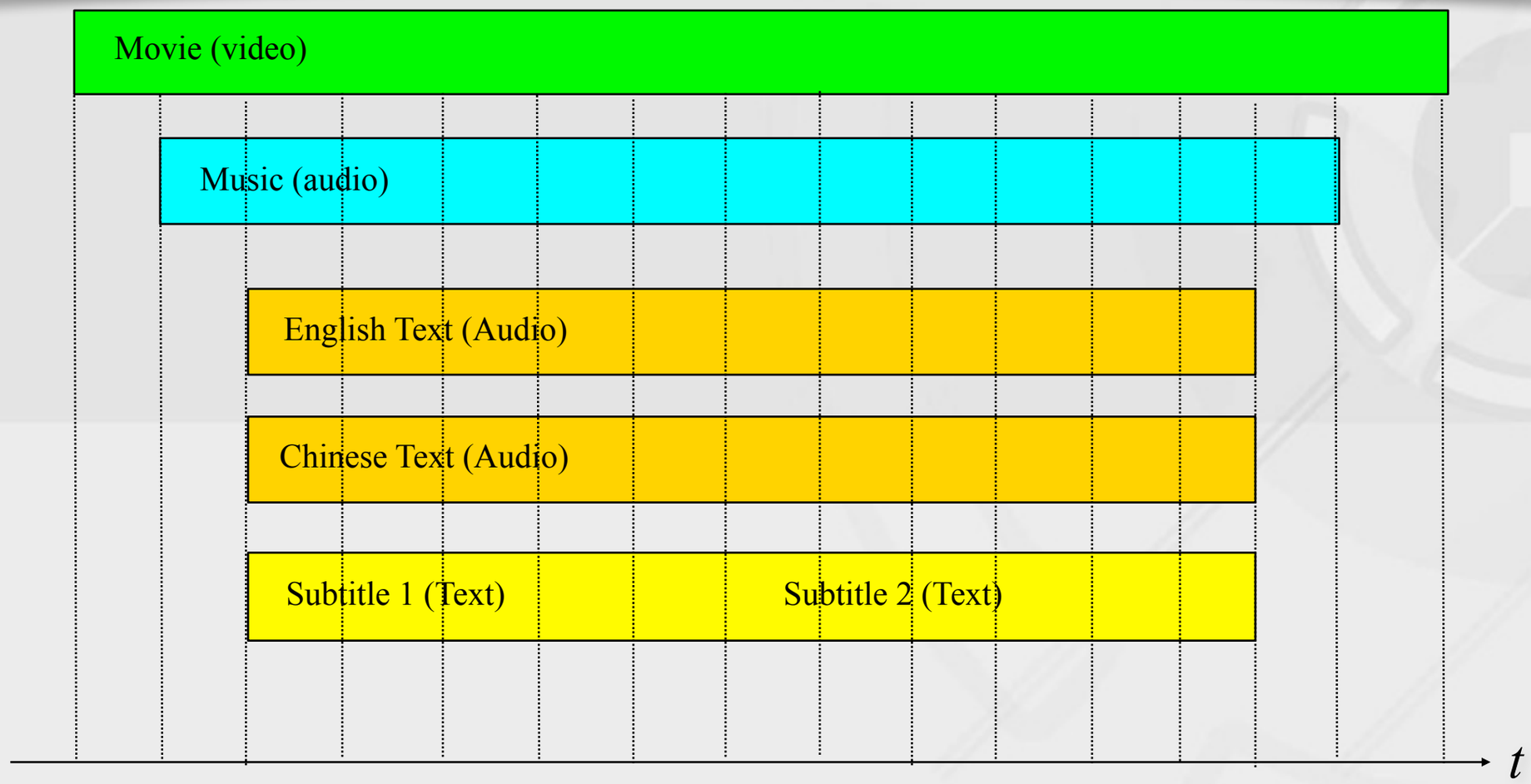


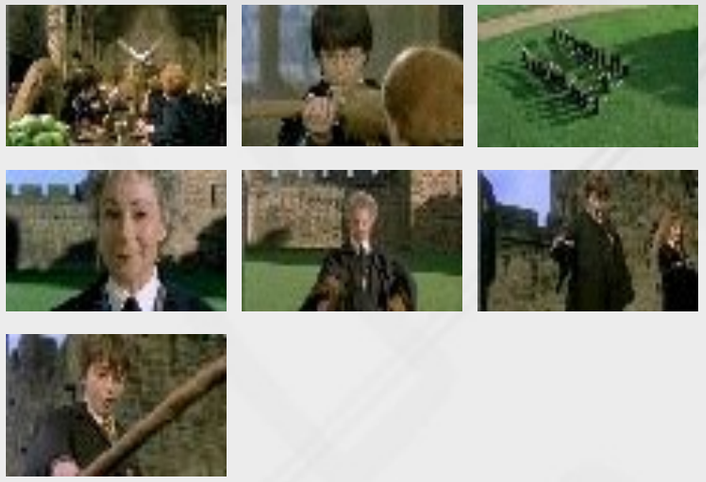
Fig. Different Media Bound to a Common Time Line

Characteristics of MM Data
-- challenges about modeling MM data

- **Semantic nature** of MM data is more complex...
  - A picture often means different things to different people => even manual annotation is hard, not to mention tedious and expensive;
  - Difficulties of cross-media search

E.g., an image query (a) and semantically relevant results (b)

image



video

Harry Potter has never been the star of a Quidditch team, scoring points while riding a broom far above the ground. He knows no spells, has never helped to hatch a dragon, and has never worn a cloak of invisibility.

text

Characteristics of MM Data
-- challenges about modeling MM data

- Operations pertinent to individual media are needed
- Data objects in MMDB can be divided into "**static**" and "**time-dependent**" ones
  - Two groups of operations:
    1. Operations on static data (eg, image)
    2. Operations on dynamic data (eg, audio/video)

  - **static**:

    op: P x DB => R x DB **(P: parameters; R: domain)**

    eg, *op(p,db)*: a compressed image file is passed to a viewer provided/ incorporated by the MMDBMS. Zoom and rotate are examples of operations for images.

Characteristics of MM Data
-- challenges about modeling MM data

- **time-dependent**:

  op: P x DB x T => R x DB x T

  eg, *op(p,db, t):* when presenting a video on a video viewer (or delivering it to any other applications), operations such as play, pause, fast forward, reverse etc. would be applicable and need to be devised.

Characteristics of MM Data
-- challenges about modeling MM data

- Providing users with a semantically effective and performance **efficient query language** is critical
  - **exact match** query of conventional databases not suitable
  - specific properties of MM data, esp.
    - the temporal and spatial aspects, the concepts of vagueness or unsharpness

  - Additional concepts are required/useful to query a MMDBMS:
    1. **content-based access**, including query-by-example (QBE)
    2. **domain-specific knowledge** (subject to the high costs on acquisition and the expensive maintenance of the domain knowledge)
    3. **speech recognition and retrieval**

Characteristics of MM Data
-- challenges about modeling MM data

- **Multi-layers of abstraction** for the retrieval and organization of the MM data
  - Example: a database of video
    - first layer: identify single scenes in videos
    - second layer: identify geometric objects in these scenes
    - third layer: geometric objects could be related to real world entities in the scenes.

A user may search for such entities based on attribute values stored elsewhere in the DB, and so, may access a video in which this entity occurs using indices at each layer.

Characteristics of MM Data
-- challenges about modeling MM data

- In order to have these features available in a MMDB system, appropriate solutions have to be found for
  - **indexing mechanisms**
  - **semantic and consistent modeling of abstractions**

- abstraction and modeling facilities (hardware resources, e.g., compression chips, presentation devices like speakers, monitors, windows, and software resources)
  - device transparency
  - resource sharing
  - device classification and modeling
  - distribution of data should be transparent to applications and users.

Definition- MMDBMS

- MMDBMS
  - *a framework that manages different types of data potentially represented in a wide diversity of formats on a wide array of media sources*

MMDBMS Characteristics

- **Uniformly query** data represented in different formats
- Query data represented in **diverse media**
- Retrieve media objects from a local storage devices in a smooth, jitter-free manner
- Provide audio visual presentation of a query result
- Deliver presentation to satisfy quality of service requirements

MMDBMS Characteristics

- Uniformly query data represented in different formats
  - Integrate data from different relational databases with different schemas
  - Query flat files, OO and spatial databases, legacy data sources
  - Ability to handle fuzzy searching

MMDBMS Characteristics

- Query data represented in **diverse (多变的) media**
  - –Queries spanning multiple media types
  - –→ ability to merge them together

MMDBMS Characteristics

- Retrieve media objects from a local storage devices in a smooth, jitter-free manner
  - Media stored in up to 10 gigabytes
  - Stored on mix of storage devices
- → variable performance characteristics

MMDBMS Characteristics

- Provide audio visual presentation of a query result
  - The user should have the ability to specify the form and content of the presentation he/she would like to obtain

MMDBMS Characteristics

- Deliver presentation to satisfy quality of service requirements
  - Buffer availability and bandwidth need to be taken into account

Multimedia equation

Multimedia equation

Multimedia = presentation + context

Multimedia equation

Multimedia = presentation + context

presentation: sensory, aesthetic part （美学）

Multimedia equation

Multimedia = presentation + context

presentation: sensory, aesthetic part （美学）

context = convergence + information + architecture

- convergence = data +platform + distribution
- information = storage and retrieval
- architecture = compression + components + connectivity

http://www.notre-dam.org

an open source DAMS

Q&A