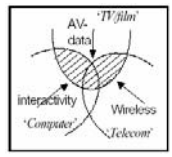
C.Sc. 366
Video Compression

The MPEG-4 Standard…
(ISO/IEC 14496)

Motivation…

- In '93 main goal was to achieve very low bit rate audio-visual coding
- In '94 this changes to 'coding of audio-visual objects'
  - Go beyond 'frame based' audio and video
- Support variety of transmission forms
  - Broadcast, network based (including low-bandwidth wireless) or local storage



The Content-based approach…

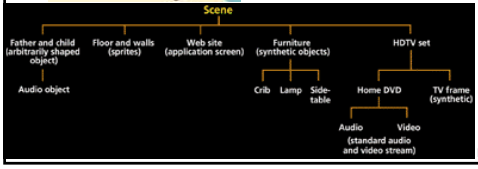
- Completely different approach to compression…

1. Support coded representation of both Natural and Synthetic audio and visual objects
2. Supports the description and synchronization of streaming data for media objects
3. Supports user interaction with media objects
4. Supports the management and identification of intellectual property

Example MPEG-4 scene (1)…

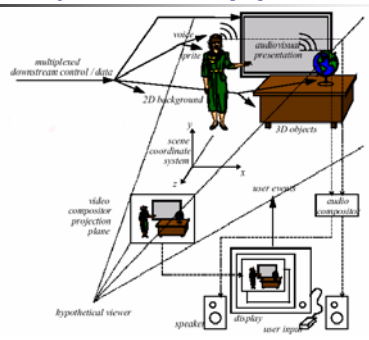
Scene Description…

- Hierarchical structure…
  - Each scene defined as comprising:
    - A background & *n* Audio-Visual Objects
  - An AVO comprises:
    - n video objects and/or audio objects
  - Video/audio objects
    - Defined as comprising *n* sub-objects

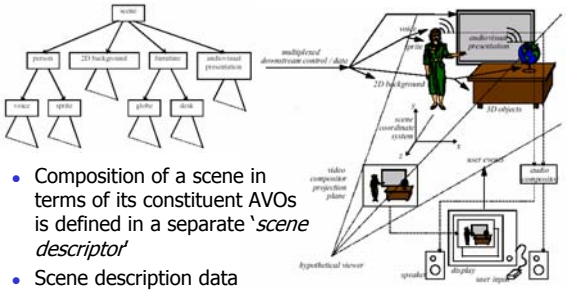


Hierarchical Scene Description

Example scene (2)…



Example scene (2)…

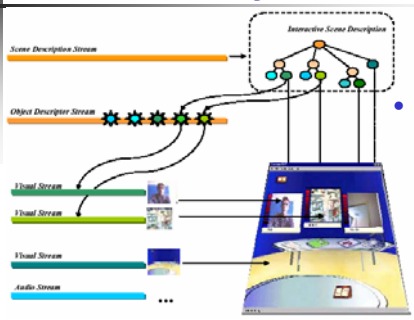


- Composition of a scene in terms of its constituent AVOs is defined in a separate '*scene descriptor*'
- Scene description data represented using BIFS…

BIFS

- Binary Format for Scenes
- Language for describing and dynamically changing a scene
- Based on VRML but…
  - Binary format
  - Supports streaming
- Example commands
  - Add/delete objects
    - "Delete node" (DN)
    - – delete nodes by identify the ID of the node
  - Change visual/acoustic properties of an object
    - "Changing the field value" (CV)
    - – colour, position,

Scene and Object Streams…



- Object descriptors
  - associated with AVOs
  - enables manipulation of the AVO

User Interaction…

- MPEG-4 provides tools for enabling dynamic scene behaviour and user interaction
  - Client-side interaction involves content manipulation that is handled locally at the end-user's terminal.
  - It consists of the modification of attributes of scene objects according to specified user actions.
  - EXAMPLE: A user could click on a scene to start an animation or video sequence.
  - Viewer can also change vantage point and calculations for update can be carried out locally.
    - Cf needs no return channel, e.g. broadcast

The Visual Part…

Coding of visual components etc…

Video Coding...

- Each Video frame is segmented into a number of video object planes (VOPs)
- Each VOP corresponds to an AVO
- Video object planes are divided further into MBs of size 16x16
- Each VOP comprising minimum number of MBs

VOP Extraction (1)…

VOP 1

VOP 2

VOP 3

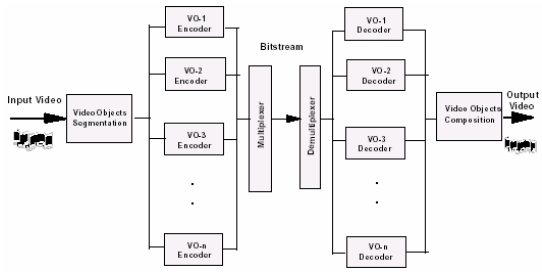
I, P and B VOPS are supported

VOP Extraction (2)… Sprites

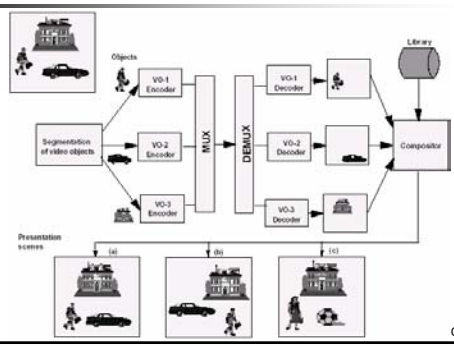
Can achieve significant bandwidth conservation if assume that the background of a scene is fixed…

- Code/Send once
  - Sprite Buffer
  - Send transforms
  - Recompose

VOP Mux/Demux…

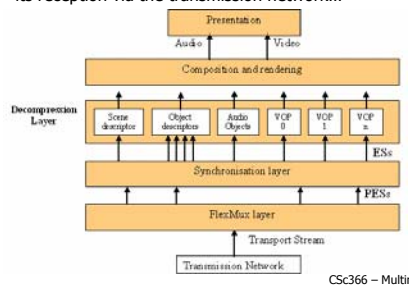


VOP Mux/Demux…



Steps for decoding a transmission…

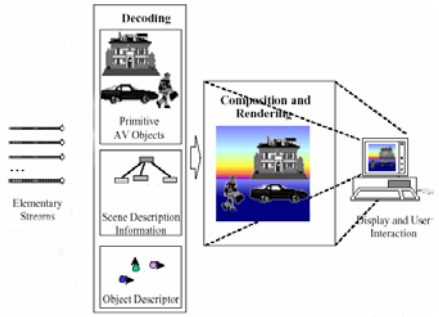
- Steps required to present an mpeg-4 encoded movie following its reception via the transmission network…



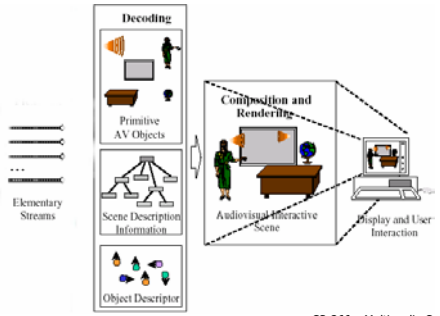
Steps for decoding…

- Elementary Streams (ESs)
  - Contain the compressed audio and video information relating to each AVO
- Packetised Elementary Streams (PESs)
  - Carry ESs in their payload
  - Each PES packet
    - routed to the appropriate synchronisation block in the sync layer
    - contains a time-stamp to ensure that information reaches decoder blocks in a timely manner
- Composition and rendering block
  - Receives decoded AVO + scene descriptor info

Decoding Example (1)…



Decoding Example (2)…



Synthetic and Natural Hybrid Coding (SNHC)…

Efficient representation and composition of synthetically and naturally generated audiovisual data…

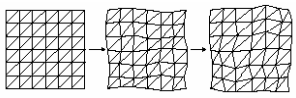
Applications include Virtual environments, conferencing, education, entertainment, media production and broadcast media experiences etc.

Facilities…

- 2D/3D Mesh
- Face Animation
- Body Animation
  - *"A body model is a representation of a virtual human or human-like character that allows portraying body movements adequate to achieve nonverbal communication and general actions."*
    - From ISO/IEC 14496-2

2D and 3D meshes…

- Map image onto mesh…
  - deformable nested polygons

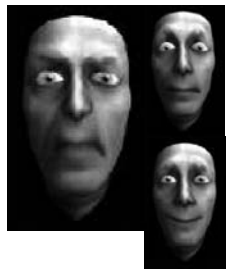


Compression potential…

Face Animation…

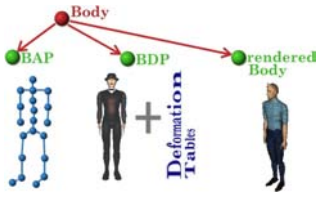
- Facial Definition Parameter (FDP) Set
  - Controls shape, gender, age, etc.
- Facial Animation Parameter (FAPs) Set
  - 68 parameters for animation/expressions

```
class FaceAnimationParams {
  public:
    int move_h_l_eyeball;
    int move_h_r_eyeball;
    int move_v_l_eyeball;
    int move_v_r_eyeball;
    int enlarge_l_pupil;
    int enlarge_r_pupil;
    ....;
}
```

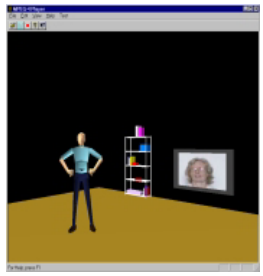


Body Animation…

- Body Animation Parameters…
  - manipulate independent degrees of freedom in the skeleton model of the body



As part of the bigger picture….



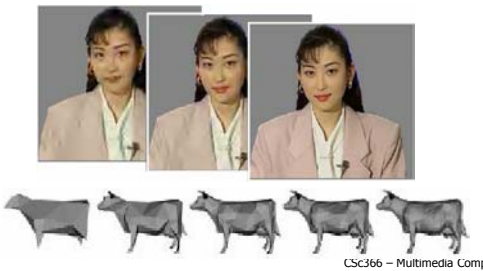
Coping with wireless…

Meeting bandwidth Reqs…

- Use terminal capabilities in order to reduce bandwidth…
  - *cf* image warping, sprites etc.
- Audio defined down to 2 kbit/s
- Because of the object based approach …
  - Priorities can be associated with objects in order to maximise available bandwidth
    - Can have greater forward error correction with high priority objects
    - If bandwidth falls below a given threshold then don't decode and display certain objects
  - OR… have multiple streams associated with an objects
    - Scalability

Scalability…

- Encode different levels of detail for a given object …
  - use lower bandwidth streams based on bandwidth available



IPRM…

- Intellectual Property Rights Management
- Intellectual property identification contained within ES descriptors
- Interface also defined in standard
  - Enables advances in cryptography standards etc. to be decoupled from MPEG-4 developments

MPEG-4 file format etc.

- Unlike approach taken by mpeg1-2 mpeg-4 specified its own file format
  - *.mp4
- Like QUICKTIME it is a file wrapper and so can be used to ship numerous tracks.
- MPEG-4 is poised for great success
  - Once licensing/patent issues sorted
  - QUICKTIME 6.0
- Try for yourself…
  - http://www.mpeg-4.philips.com/
  - Also DivX strongly based on MPEG-4

Summary…

- Supports object-based (AVOs) Natural and synthetic coding
- Potential for significant client side interactivity
- Supports management of Intellectual Property
- Wide range of access conditions
  - e.g. Low bandwidth, broadcast, local retrieval etc.
- Not intended as replacement of MPEG-2
- Specifies its own file format (mp4)