18-899 Special Topics in Signal Processing

Multimedia Communications:
Coding, Systems, and Networking

Prof. Tsuhan Chen
tsuhan@ece.cmu.edu

Lecture 10

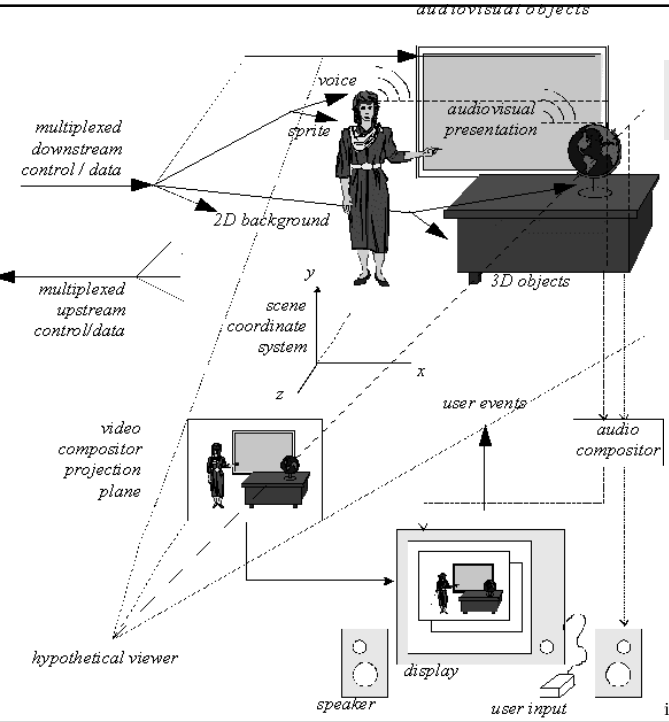
MPEG-4 Overview

MPEG-4

- Originally
  - A standard for very low bit rate coding of limited complexity audio-visual material
- In July 94, the scope was extended to
  - Functionalities not supported by other standards
    - Content-based interactivity
    - Universal access
    - High compression
  - Coding of general material for a wide bit rate range
  - Flexibility and extensibility

Content-Based Interactivity

- A scene is composed of audio-visual objects
  - Not just pixels or moving blocks
- Objects can be of different nature
  - Text or images
  - Rectangular or arbitrary shape
  - 2D or 3D objects
  - Natural of synthetic
- Different coding schemes applied to different objects
- Compositor puts objects back in a scene

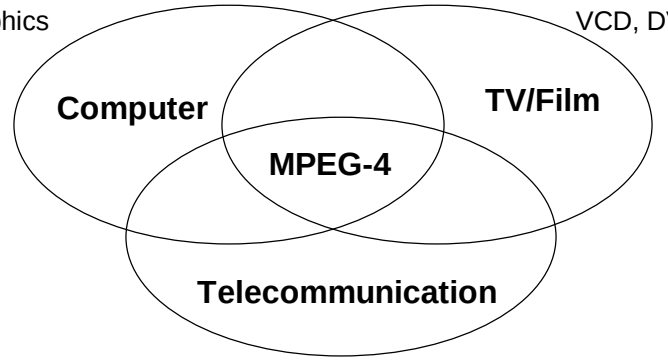


[MPEG-4 N1909, "Overview of the MPEG-4 Version 1 Standard"]

Applications

Human-machine interface
GUI, Virtual environment
Vision, Graphics

Content creation
Digital TV, HDTV
VCD, DVD



Wireless, Internet, ISDN, POTS, Cable

Parts of MPEG-4

- Part 1: Systems
- Part 2: Video
- Part 3: Audio
- Part 4: Conformance testing
- Part 5: Reference software
- Part 6: Delivery multimedia integration framework
- Others
  - Synthetic and Natural Hybrid Coding (SNHC)
  - Requirements and applications
  - Implementation Study
  - Intellectual property rights (IPR)

MPEG-4 Activities

- Competitive phase
  - Proposals and evaluations
- Collaborative phase
  - Verification model and core experiments
    1. Define Verification Model (VM-n)
    2. Define core experiments for improving VM-n
    3. Perform core experiments. Compare with VM-n
    4. n++, go to Step 1

MPEG-4 Time Table

- July 93 Started work
- Nov 95 Subjective tests and tool evaluation
- Jan 96 Define verification model (VM) and core experiments (CE)
- Mar/July/Sept/Nov 96, Feb/Apr/Jul 97
  Update VM and define a new set of CEs
- Oct 97 Committee Draft (CD)
- July 98 Final CD (FCD)
- Nov 98 Draft international standard (DIS)
- Jan 99 International standard (IS)

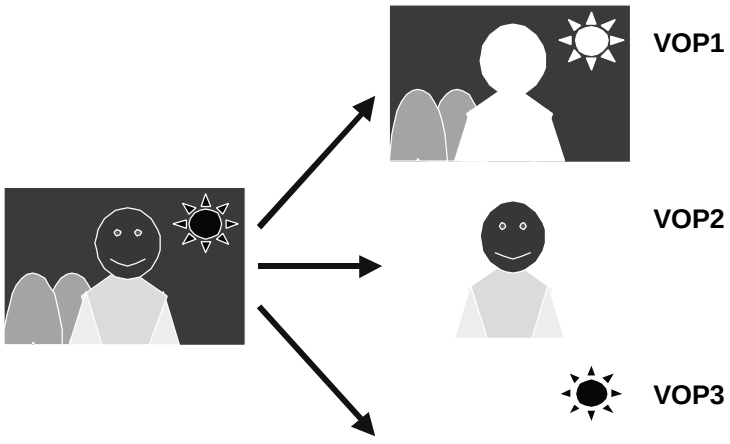
MPEG-4 Video

- General functionalities
  - Coding efficiency
    - For 5 kbit/s ~ 5 Mbit/s
    - From small images to TV resolution
    - Progressive/interlaced
  - Error resilience and robustness
  - Spatial and temporal scalabilities
- Content-based functionalities
  - Shape coding and sprites
  - Content-based scalabilities
  - Error resilience and robustness

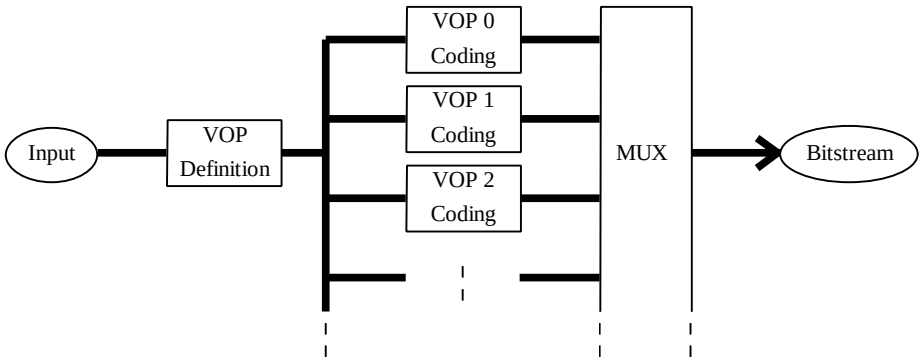
MPEG-4 Video (cont.)

- Tools
  - Motion/texture coding derived from H.263
  - Coding of video object plane (VOP): I, B, P
  - Binary and gray-scale shape coding
  - Scalabilities: temporal/spatial
  - Static sprites
  - Interlaced prediction
  - 12 bit video
  - Computational graceful degradation (CGD)

Video Object Plane (VOP)

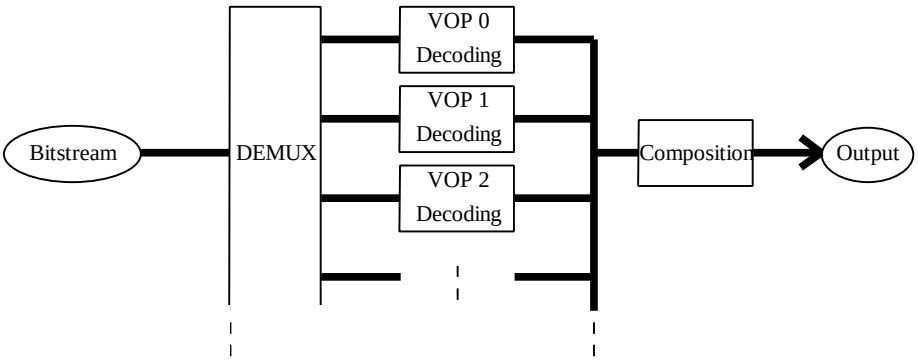


Structure of VOP Encoder



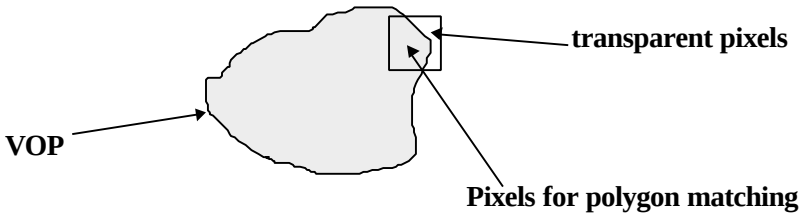
– Note: Segmentation is outside the scope of MPEG-4

Structure of VOP Decoder



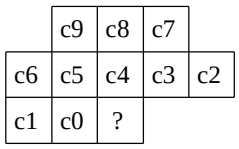
Coding of VOP

- Motion compensation and DCT
  - Similar to H.263
- Polygon matching for motion estimation

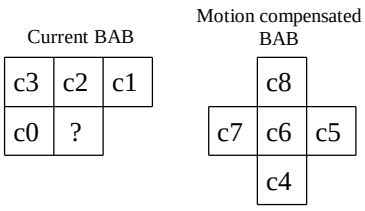


Binary Shape Coding

- Context-based arithmetic encoding (CAE)
  - A binary shape is treated as a binary image
  - Apply CAE to each binary alpha block (BAB)
- The "context"



Intra



Inter

Synthetic & Natural Hybrid Coding (SNHC)

- Efficient representation and composition of synthetically and naturally generated audiovisual data
- To be integrated into MPEG-4 Video and Audio
  - Not a separate part of MPEG-4
- Applications
  - Virtual environment, conferencing, education, entertainment, media production, and real-time, interactive and broadcast media experiences

SNHC Target technologies

- Video
  - Face animation
  - 2D/3D mesh compression
  - Still texture coding: wavelet-based
  - View dependent scalability
- Audio
  - Text-to-speech synthesis, structured audio, environmental auralization, 3D audio, etc.

Face Animation

- Face animation
  - 2D/3D polygon mesh for face rendering
  - Facial Definition Parameter (FDP) Set
    - Controls shape, texture, gender, age, etc.
  - Facial Animation Parameter (FAP) Set
    - Controls expressions and animation

MPEG-4 Version 2

- One year following Version 1
- Adds new profiles with new functionalities
- Video
  - Scalable transmission of arbitrary-shaped objects
  - Tools for additional efficiency improvements
  - Tools for improved error robustness
  - Coding of multiple views
  - Body animation
  - Coding of 3D meshes and scalabilities