

CS 414 – Multimedia Systems Design  
Lecture 11 – MPEG-4 Video  
(Part 6)

Klara Nahrstedt  
Spring 2008



# Administrative

- MP1 deadline Friday, February 8
- Demonstrations, Friday, February 8, 5-7pm
- Please sign up for a demonstration slot
- MP2 will be posted on Friday, February 8 and deadline will be Friday, 22 (please, start early – we will have two discussion sections for MP2)

# Interact With Visual Content

The image is a screenshot of a web page from IBM, dated 12 April 2001. The page features a dark background with a glowing blue circular graphic on the left side. The main content is organized into several sections:

- IBM tv**: Located at the top right, with the date "12 April 2001" below it.
- Nasdaq**: A section showing market data: "2,840.39", "Adv ▲ 117", and "▲ 82.48". Below this, it shows "Dec ▼ 73".
- DJA**: A section showing market data: "10,676.06" and "10,574.25".
- Company profile**: A green button located at the bottom of the main content area.
- NYSE**: A section showing "IBM 108.56 ▲ 2.59" and "Some".
- AMEX**: A section showing "Hawthorne 12.5 unch" and "Watson Labs 33.5 ▲ 4.7".
- News Headline**: "IBM's MPEG-4 player drives stock price up".
- News Text**: "IBM 108.56 ▲ 2.59: at current price stock's still a steal".




# MPEG-1/MPEG-2

- Pixel-based representations of content
  - takes place at the encoder
- Lack support for content manipulation
  - e.g., remove a date stamp from a video
  - turn off “current score” visual in a live game
- Support manipulation and interaction if the video is aware of its own content



# Original MPEG-4

- Conceived in 1992 to address very low bit rate audio and video (64 Kbps)
- Required quantum leaps in compression
  - beyond statistical- and DCT-based techniques
  - committee felt it was possible within 5 years
- Quantum leap did not happen



# The “New” MPEG-4

- Support object-based features for content
- Enable dynamic rendering of content
  - defer composition until decoding
- Support convergence among digital video, synthetic environments, and the Internet



# MPEG-4 Components

- Systems – defines architecture, multiplexing structure, syntax
- Video – defines video coding algorithms for animation of synthetic and natural hybrid video (Synthetic/Natural Hybrid Coding)
- Audio – defines audio/speech coding, Synthetic/Natural Hybrid Coding such as MIDI and text-to-speech synthesis integration
- Conformance Testing – defines compliance requirements for MPEG-4 bitstream and decoders
- Technical report
- DSM-CC Multimedia Integration Framework – defines Digital Storage Media – Command&Control Multimedia Integration Format; specifies merging of broadcast, interactive and conversational multimedia for set-top-boxes and mobile stations

# MPEG-4 Example

Custom Interface with Personalization

Full-screen MPEG-4 video

**Interactivity**

Multiple Video Windows

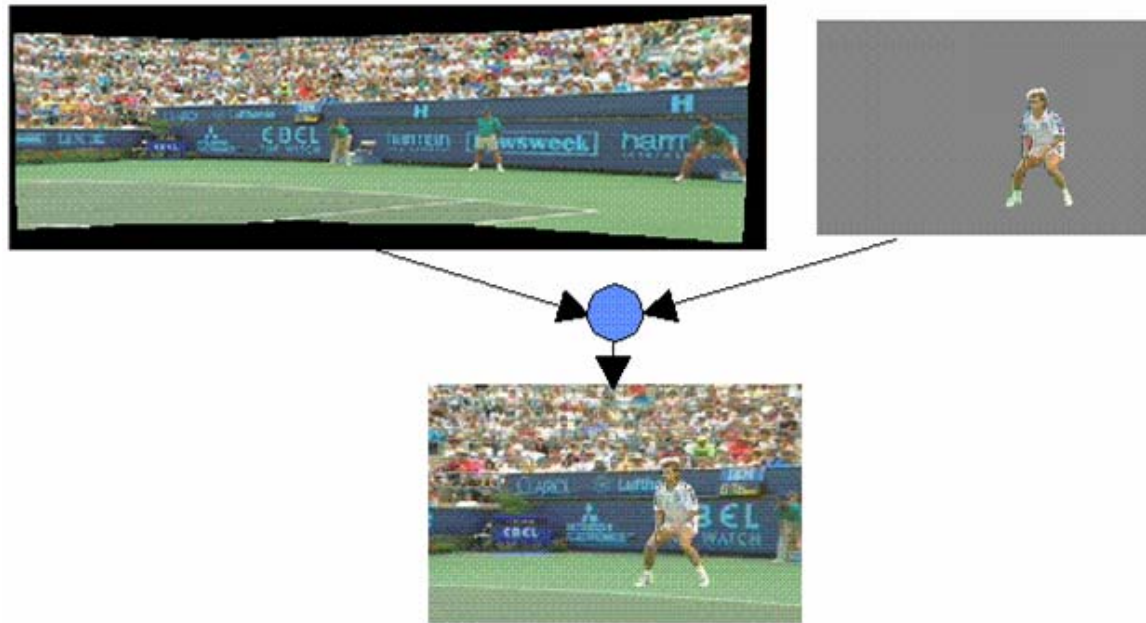
2D Graphics alpha-blended

CD-Quality Audio

Custom Branding

The screenshot displays a rich media player interface for 'Yukon Adventure'. It includes a main video window, a 'FACT BOARD' with a menu, a 'Geography' window with a map, and a vertical stack of three video thumbnails. The interface is branded with 'mediascentre' and 'VIEW 1'.

# MPEG-4 Example



# MPEG-4 Example

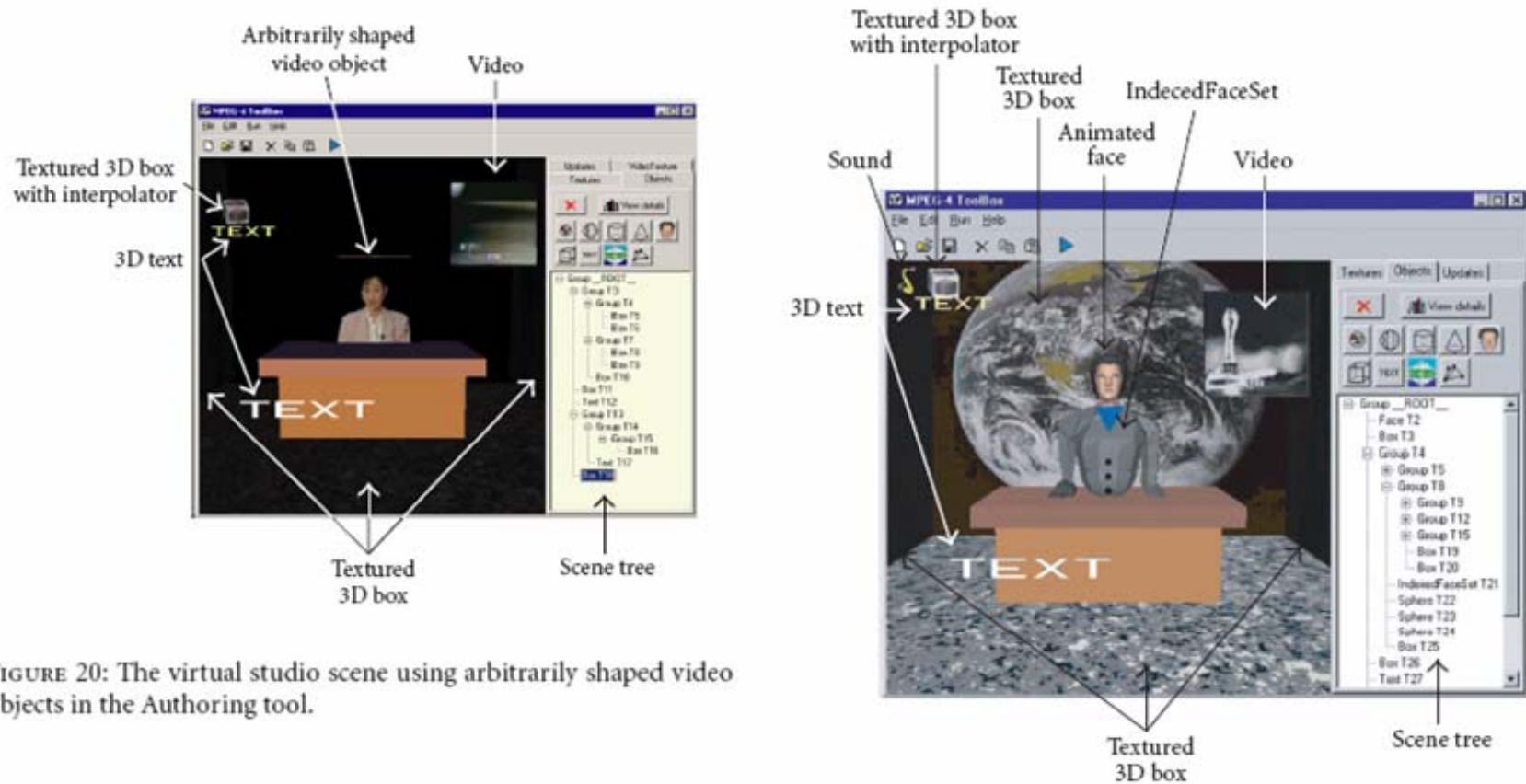


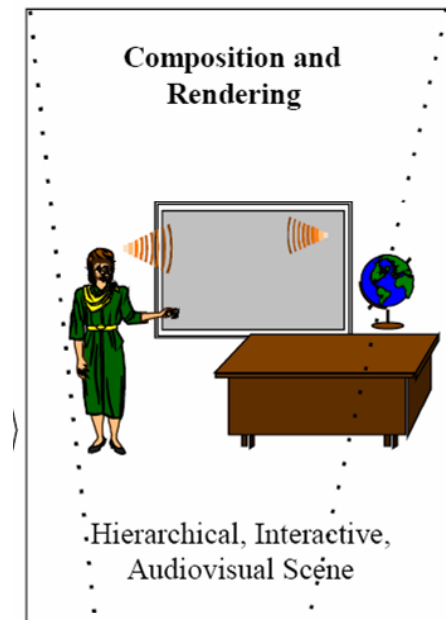
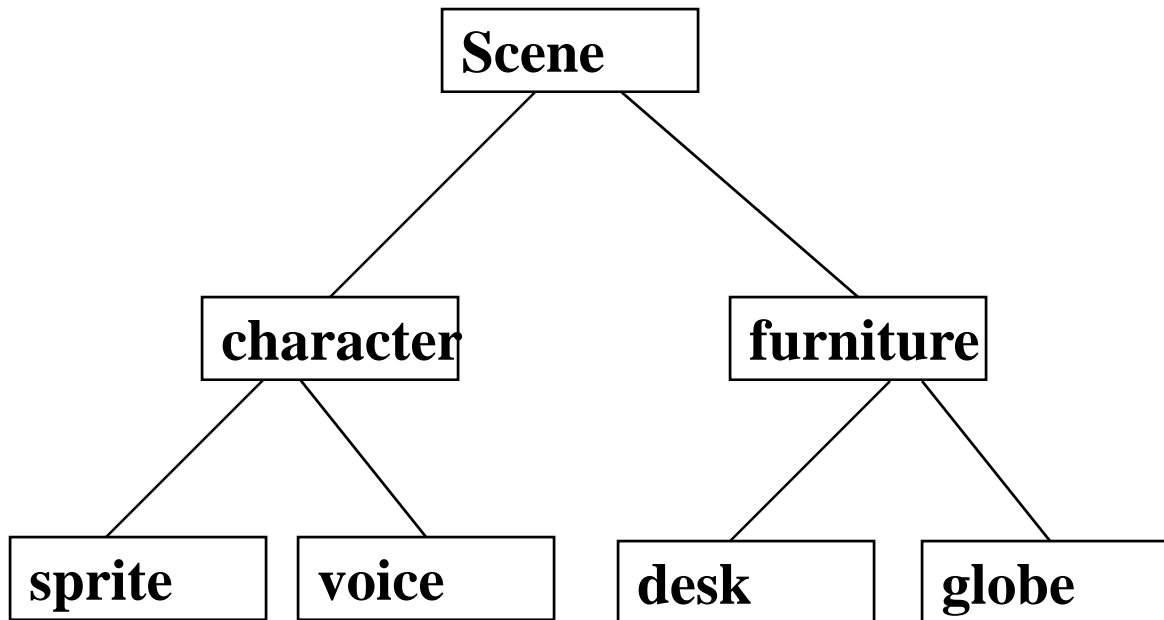
FIGURE 20: The virtual studio scene using arbitrarily shaped video objects in the Authoring tool.



# Media Objects

- An object is called a *media object*
  - real and synthetic images; analog and synthetic audio; animated faces; interaction
- Compose media objects into a hierarchical representation
  - form compound, dynamic scenes

# Composition

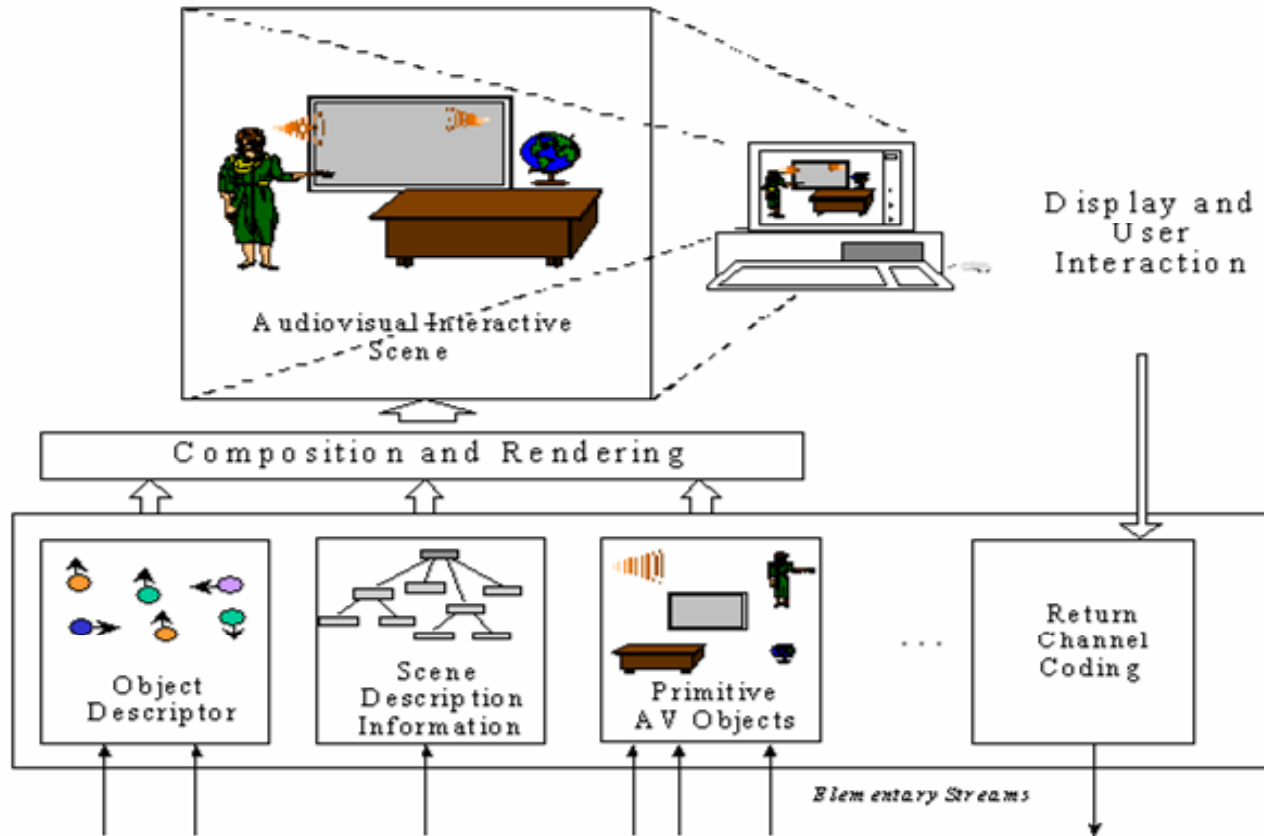




# Composition (cont.)

- Encode objects in separate channels
  - encode using most efficient mechanism
  - transmit each object in a separate *stream*
- Composition takes place at the decoder, rather than at the encoder
  - requires a binary scene description (BIFS)
- BIFS is low-level language for describing:
  - hierarchical, spatial, and temporal relations

# MPEG-4 Rendering





# Interaction as Objects

- Change colors of objects
- Toggle visibility of objects
- Navigate to different content sections
- Select from multiple camera views
  - change current camera angle
- Standardizes content and interaction
  - e.g., broadcast HDTV and stored DVD



# Hierarchical Model

- Each MP4 movie composed of tracks
  - each track composed of media elements (one reserved for BIFS information)
  - each media element is an object
  - each object is a audio, video, sprite, etc.
- Each object specifies its:
  - spatial information relative to a parent
  - temporal information relative to global timeline



# Synchronization

- Global timeline (high-resolution units)
  - e.g., 600 units/sec
- Each continuous track specifies relation
  - e.g., if a video is 30 fps, then a frame should be displayed every 20 units
- Others specify start/end time



# MPEG-4 Audio

- Bit-rate 2-64kbps
- Scalable for variable rates
- MPEG-4 defines set of coders
  - Parametric Coding Techniques: low bit-rate 2-6kbps, 8kHz sampling frequency
  - Code Excited Linear Prediction: medium bit-rates 6-24 kbps, 8 and 16 kHz sampling rate
  - Time Frequency Techniques: high quality audio 16 kbps and higher bit-rates, sampling rate  $> 7$  kHz




# Conclusion


- A lot of MPEG-4 examples with interactive capabilities
- Content-based Interactivity
  - Scalability
    - Spatial Scalability
    - Temporal Scalability
  - Sprite Coding
- Improved Compression Efficiency (Improved Quantization)
- Universal Accessibility
  - re-synchronization
  - data recovery
  - error concealment

# Another Interaction


Zoo **WILD GUIDE**




Crested Tinamou



Cuban Amazon



Fairy Bluebird



Concave Casqued Hornbill

00:00:26 00:03:32

The image shows a digital interface for a 'Zoo Wild Guide'. It features a yellow background with a white title bar at the top containing the word 'Zoo' and 'WILD GUIDE' in large, bold, white letters. Below the title bar, there are four bird species displayed in a 2x2 grid. Each species is represented by a small photograph and a text label below it. The species are: Crested Tinamou (top-left), Cuban Amazon (top-right), Fairy Bluebird (bottom-left), and Concave Casqued Hornbill (bottom-right). At the bottom of the interface, there is a white bar containing a video player control interface with a play/pause button, a progress bar, and a volume icon. The progress bar shows a current time of 00:00:26 and a total duration of 00:03:32.

# Interactive Drama

