6 Multimedia Content Production and Management

6.1 Content for Network-Based Delivery

6.2 Encoding and Transcoding
6.3 Media Production Chains
6.4 Media Asset Management

Literature:

Outline

1. Introduction and Motivation
2. Digital Rights Management
3. Cryptographic Techniques
4. Electronic Payment Systems
5. Multimedia Content Description
6. Multimedia Content Production and Management
7. Streaming Architectures
8. Commercial Streaming Systems: An Overview
9. Communities, the Web and Multimedia
10. Web Radio and Web TV
11. Signaling Protocols for Multimedia Communication
12. Multimedia Conferencing

Part I: Content-Oriented Base Technologies
Part II: Multimedia Distribution Services
Part III: Conversational Multimedia Services
Types of Multimedia Content for Streaming

- Dynamic Text
  - Low bandwidth
  - Proprietary formats (e.g. by RealNetworks)
- Audio
  - Bandwidth unproblematic in modern broadband networks (e.g. DSL access), but still an issue in older network technologies (e.g. 28 kBit/s modem)
  - Proprietary and standard formats (e.g. MP3, MPEG-4 AAC)
- Video
  - Bandwidth is still a serious issue, except very high-speed links (e.g. VDSL2)
  - Proprietary and standard formats (e.g. MPEG-1/2/4, H.264)
- Vector animations
  - Bandwidth mainly unproblematic
  - Proprietary formats (mainly Shockwave Flash SWF)
  - Download & play approach is sufficient

Supporting Good Video Compression

- Already when capturing a video source, good later compression can be prepared:
- High-quality capturing and storing equipment with low *video noise*
  - Noise (grain) may disturb compression algorithms (high frequencies)
- Simple pictures with little detail
  - No total views of large scenes
  - Few objects in large scale
- Little movement
  - Camera movement, object movement, zoom lead to worse compression
- Good lighting
  - Sharp edges and high contrast assist the compression algorithms
- No fine-grain patterns
  - High picture frequencies are problematic in compression
  - E.g. Uni-colour clothes instead of patterned ones
Supporting Low Audio Bandwidth

- Preparing audio content with low bandwidth requirements, a set of tips by Chris Priestman
- Halve your data by using mono unless stereo is really important.
- Halve it again by cutting the sample rate to 22.05 kHz for streams of 28.8 kbps and less.
- Lose the very top of the hearing range (above 10 kHz).
- Preserve detail across the tonal range by keeping resolution at 16 bits (unless for speech at less than 16 kbps)

Video Capturing

- Modern standard computer interfaces (e.g. USB, FireWire)
  - Enable import of video data in compressed raw format with limited quality (e.g. DV standard)
  - Capturing is computing-intensive, CPU has to be shared between capturing and compression (e.g. MPEG)
  - Live streaming only with low-quality input (e.g. WebCam)
- Specialized video capturing hardware for PCs (card or external device)
  - Enable real-time encoding with less CPU load
  - Support various physical video connection standards
  - Limitations in quality (picture size, frame rate)
- Professional video capturing/encoding systems
  - Expensive
  - Able to deal with uncompressed or only mildly compressed video
  - Live compression, often in hardware (e.g. MPEG-4, H.264)
- Possible trend:
  - Video surveillance mass market leads to affordable high-quality encoders
Futuristic Trends in Image & Video Capture

• Automation of cost/ and personal intensive tasks in capturing
  − Direction
  − Quality control
• “Active capture”
  − Intelligent feedback loops
    » Emit stimuli for people to be recorded
    » Check quality and possibly repeat
• Applied image analysis / computer vision
  − Automated checks for quality like contrast, sharpness etc.
  − Recognition of objects, faces, …
  − Layout assessment

6 Multimedia Content Production and Management

6.1 Content for Network-Based Delivery

6.2 Encoding and Transcoding

6.3 Media Production Chains

6.4 Media Asset Management

Literature:
David Austerberry: The Technology of Video & Audio Streaming, Focal Press 2002
Gregory C. Demetriades: Streaming Media, Wiley 2003
Encoding and Transcoding

- Audio and video needs to be converted for streaming delivery
  - Compression, proprietary formats
- Transcoding: Conversion of media files from one format to another
- Repurposing: Using existing content for new purposes
  - e.g. using TV ads as streaming content

Factors Determining Video Bandwidth

- Physical resolution (number of pixels)
  - Determines picture size in standard rendering resolution (e.g. 72 dpi)
  - Dependent on playback device
    - "Set Top Box" for TV set requires full-screen TV signal
    - Video window on PC can be adjusted in size
- Frame rate
  - Desirable: 25 fps
  - Over low-bandwidth links often only smaller rates possible (e.g. 10 fps)
- Colour (sub)sampling
- Audio quality
  - Sampling rate, resolution (e.g. speech vs. CD quality)
  - Mono, stereo, multi-channel
- Degree of compression
  - Determines appearance of compression artefacts
Network Limitations

• Bandwidth towards receiver is limited:
• Effective bandwidths for various access network technologies:
  – 28.8 modem: 20 – 23 Kbps
  – 56.6 modem: 32 – 35 Kbps
  – ISDN: 45 – 55 Kbps
  – Dual-ISDN: 80 – 100 Kbps
  – DSL: 1000 Kbps and more
  – VDSL: 25 Mbps and more
  – Cable modem: 4 – 36 Mbps
  – LAN: 10 – 100 Mbps
• Compromise between bandwidth limitations and quality:
  – Picture format
    » E.g. for 28.8 modem picture format 176 x 144 pixel (QCIF)
    » E.g. for DSL picture format 360 x 288 pixel (CIF)
  – Plus other factors

Example: Multiple Bit Rate Encodings

<table>
<thead>
<tr>
<th>Video source</th>
<th>Broadcast (DVB)</th>
<th>DSL/cable modem</th>
<th>Modem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target data rate</td>
<td>(270 Mbit/s)</td>
<td>4 Mbit/s</td>
<td>500 kbit/s</td>
</tr>
<tr>
<td>Required data reduction</td>
<td>40:1</td>
<td>330:1</td>
<td>4700:1</td>
</tr>
<tr>
<td>Frame size</td>
<td>720 x 480 (CCIR 601)</td>
<td>720 x 480</td>
<td>192 x 144</td>
</tr>
<tr>
<td>Frame rate</td>
<td>30</td>
<td>30</td>
<td>15</td>
</tr>
<tr>
<td>Colour sampling</td>
<td>4:2:2</td>
<td>4:2:0</td>
<td>YUV12</td>
</tr>
<tr>
<td>Uncompressed data rate (Mbit/s)</td>
<td>166</td>
<td>124</td>
<td>5</td>
</tr>
<tr>
<td>Fraction of original data rate</td>
<td>1:1.33</td>
<td>1:33</td>
<td>1:144</td>
</tr>
<tr>
<td>Required compression</td>
<td>30:1</td>
<td>10:1</td>
<td></td>
</tr>
</tbody>
</table>

From: D. Austerberry
### Common Video Image Formats

<table>
<thead>
<tr>
<th>Format</th>
<th>Resolution</th>
<th>Frame Rate</th>
<th>Sub-sampling</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCIR 601 (NTSC)</td>
<td>720 x 480</td>
<td>30</td>
<td>4:2:2</td>
<td>Broadcast (DVB), DVD</td>
</tr>
<tr>
<td>CCIR 601 (PAL)</td>
<td>720 x 576</td>
<td>25</td>
<td>4:2:2</td>
<td></td>
</tr>
<tr>
<td>SIF (NTSC) Standard Interchange Format</td>
<td>352 x 240</td>
<td>30 progressive</td>
<td>4:2:0</td>
<td>CD-ROM</td>
</tr>
<tr>
<td>SIF (PAL)</td>
<td>352 x 288</td>
<td>25</td>
<td>4:2:0</td>
<td></td>
</tr>
<tr>
<td>CIF Common Intc. Format</td>
<td>352 x 288</td>
<td>30 progressive</td>
<td>4:2:0</td>
<td>Videoconference, streaming</td>
</tr>
<tr>
<td>QCIF Quarter CIF</td>
<td>176 x 144</td>
<td>30</td>
<td>4:2:0</td>
<td>Videoconference, streaming</td>
</tr>
</tbody>
</table>

From: D. Austerberry

---

### Multiple Bitrate Encoding

- In general, the same content has to be encoded in several qualities/bitrates
- File allocation:
  - One file multiplexing several qualities, or
  - Several files
- Selection of appropriate quality/bitrate:
  - Dependent on network access technology and dynamic network load
  - Manual selection: Through different alternatives on Web page, or
  - Automatic selection:
    - Using streaming server software and adequate client
    - Often access network type stored in user preferences for client software
Reminder: A Sample SMIL File

```
<smil xmlns="http://www.w3.org/2001/SMIL20/Language">
  <head>
    <layout>
      <root-layout width="356" height="356"
        backgroundcolor="black"/>
      <region id="imgReg" width="256" height="256"
        left="50" top="50"/>
    </layout>
  </head>
  <body>
    <seq>
      <img region="imgReg" src="tiger.jpg" dur="4s"/>
      <img region="imgReg" src="elephant.jpg" dur="4s"/>
      <img region="imgReg" src="butterfly.jpg" dur="4s"/>
    </seq>
  </body>
</smil>
```

From: Digitale Medien

---

SMIL and Bandwidth Selection

- Example:
  ```
  <smil>
    <body>
      <par>
        <switch>
          <audio src="http://www.providerx.com/datei1.rm"
            system-bitrate="250000"/>
          <audio src="http://www.providerx.com/datei2.rm"
            system-bitrate="100000"/>
          <audio src="http://www.providerx.com/datei3.rm"
            system-bitrate="40000"/>
          <audio src="http://www.providerx.com/datei4.rm"
            system-bitrate="1000"/>
        </switch>
      </par>
    </body>
  </smil>
  ```

- RealPlayer supports SMIL
  - Selects first stream which is smaller than bandwidth from user preferences
Network-Based Interactive Video Delivery

- Using the advantages of interactivity and media composition
  - Making a difference between net-based media and traditional media
  - Leveraging the technical potential for better business opportunities
- Example:
  - User is interested in skiing and selects a skiing video from a Video on Demand repository
  - Show him or her in parallel or interleaved an ad for snowboards
  - Provide him or her with an option to enrol in a drawing for a ski-weekend getaway (and collect personal data for later advertising)
  - Allow him or her a one-click opportunity for lift tickets at a nearby resort
  - Provide him or her with links to further video material (e.g. skiing courses, travel guides to skiing regions), which cost a viewing fee

(from: Demetriades 2003, starts to become reality now with advertisement-driven video archives/exchange systems)

Combining Media Elements to Compound Media

- Combining video streams, audio streams, text captions, graphics, links to Web locations
  - In space on the screen (e.g. video with banner advertisement)
  - Temporally (e.g. “pre-roll advertisement” with video streams)
- Enhancing interactivity and flexibility
  - E.g. free navigation
  - E.g. language options
- Technological basis:
  - Spatio-temporally structured compound multimedia documents
  - with high degree of interactivity
  - Example technologies:
    » SMIL in RealPlayer
    » MPEG-4
Automated Media Composition and Repurposing?

- Sufficient metadata annotation enables automated semantic decisions
  - Combination of media elements
  - Repurposing, e.g. creating adequate still picture from video
- Serious open research issue: How to express the overall structure of a composed piece of media?
  - Grammar-like system to express the semantic and aesthetic composition
    - For movies
    - For general multimedia productions
  - “Applied media aesthetics” (H. Zettl)
  - M. Davis: Media Streams - An iconic language for video representation, 1995

6 Multimedia Content Production and Management

6.1 Content for Network-Based Delivery
6.2 Encoding and Transcoding
6.3 Media Production Chains
6.4 Media Asset Management

Literature:
- Gregory C. Demetriades: Streaming Media, Wiley 2003
- Jürgen Mayer (Hrsg.): streaming media - Internet bewegter, bunter, lauter. Markt&Technik 2001
High-Level View of Media Production

- Premeditate
- Capture
- Archive
- Annotate
- Query
- Message Construction
- Organise
- Publish
- Distribute

Lynda Hardman: Canonical Processes of Media Production, CWI Amsterdam, REPORT INS-E0512 SEPTEMBER 2005

Hardware in the Streaming Delivery Chain

- A/V Capture
- Encoding station(s)
  Workstations or automatic servers
- Media archive servers
- Storage
- Content delivery network – possibly many servers
- Streaming
- Web serving
- Possibly many servers with load balancing
- Distribution
- Viewing

Decisions to be taken based on (e.g.)
- expected scale of business
- required reliability
Organisations in the Streaming Delivery Chain

Case Study: “Big Brother” 2000

- Encoding: provided by TV1.de (10 encoding servers)
- Serving:
  - 50 million streaming requests...
  - World Online
    » specially developed load balancing solution
    using global load balancing (based on BGP)
    and local load balancing among streaming servers
  - plus distributed server farm in the network of Deutsche Telekom
- Internet service provider: Nacamar
Automated Transcoding

- Example 1: Publishing Multiple Formats
  - Broadcaster is creating 8 hours of content per day
  - Repurposing into streaming media for Web-based Video-on-Demand
  - Live capturing, encoding (e.g. MPEG)
  - After program end: transcoding to different bitrates, delivery to streaming server

- Example 2: Flipping on Demand
  - Media archive for a cable channel to be made available through Web
  - Media kept in single, high-quality format
  - On demand (request), files are transcoded, watermarked, streamed

- Example 3: Collaboration Distribution
  - Large company working on marketing materials
  - One rough cut of a new commercial to be distributed to 100 clients with varying quality expectations and platforms
  - “Content distribution service” transcodes according to client requirements

- Example product: Telestream FlipFactory (www.telestream.net)

Case Study: Major League Baseball

- Major League Baseball Advanced media (MLBAM)
  - Audio and video indexing techniques
  - Searchable interactive video database (see MLB.com)
6 Multimedia Content Production and Management

6.1 Content for Network-Based Delivery
6.2 Encoding and Transcoding
6.3 Media Production Chains
6.4 Media Asset Management

Literature:

Gregory C. Demetriades: Streaming Media, Wiley 2003
Rosenblatt et al., Chapter 10
Information Progression and Content

- A holistic view according to Demetriades (p. 189) and Virage Inc.:

Content Monetization

- There are several traditional models for gaining a return on investment on content
  - Network-based media enable the integration of all models
Digital Asset Management

- Very similar acronyms:
  - Digital Asset Management DAM
  - Media Asset Management MAM
    - Rich Media Asset Management RMAM
  - Digital Media Management DMM
- Basic idea:
  - To make the right media material (*media assets*) available for each specific use, in the right version and the right format
- Integration technology:
  - Workflow integration
  - Integration with various media processing tools
  - Integration with content management and syndication solutions
- Broad range of product offerings
  - From large IT companies (IBM, EMC) to niche vendors

Example: Artesia TEAMS

- Digital Asset Management product, see www.artesia.com
- Media ingestion:
  - Various import tools, e.g. hot folders, email
- Media file storage, access and delivery
- Complete workflow coverage:
  - Individual activities of team members
  - Group projects
- Individual view:
  - “Inbox” – What are the tasks I am assigned to, which dates, which assets
- Project view:
  - Participants, status, associated assets, events (milestones, new versions)
- Asset management view:
  - Asset-centric, navigation to various projects
  - History: “where used”, “who used”, “how used”
Asset Management, Rights and Metadata

  - “The defining characteristic of a digital asset is that it is an asset.”
  - “There is general agreement that an asset is the asset’s content plus metadata (or data about the content). Metadata include information about … rights and permissions …”

---

The Generality of Digital Asset Management

<table>
<thead>
<tr>
<th></th>
<th>Video</th>
<th>XML</th>
<th>Quark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ingest</td>
<td>Encode and log</td>
<td>Parse and validate</td>
<td>Decompose</td>
</tr>
<tr>
<td>Index</td>
<td>Metadata and closed captions</td>
<td>Structure, attributes and content</td>
<td>Extracted content</td>
</tr>
<tr>
<td>Store</td>
<td>Media-server and HSM support</td>
<td>Fragments</td>
<td>de-binarize</td>
</tr>
<tr>
<td>Metadata</td>
<td>Format and codec dependent</td>
<td>Semantic web, RDF and industry specific DTD support</td>
<td>Extract</td>
</tr>
<tr>
<td>Model</td>
<td>Clips, tracks, key frames and storyboard</td>
<td>DTD or schema</td>
<td>decompose</td>
</tr>
<tr>
<td>Search</td>
<td>Visual search, frame accurate offsets, metadata</td>
<td>Contextual within DTD structure and metadata</td>
<td>Content and metadata</td>
</tr>
<tr>
<td>Navigate</td>
<td>Storyboard, low resolution versions</td>
<td>XLink, Xpointer</td>
<td>component</td>
</tr>
<tr>
<td>Preview</td>
<td>Clip sequencing</td>
<td>XSLT, CSS styles/gfx generation</td>
<td>page preview</td>
</tr>
<tr>
<td>Export</td>
<td>SMIL, play decision lists</td>
<td>Transform via XSLT, DOM, etc.</td>
<td>re-assemble</td>
</tr>
<tr>
<td>Distribute</td>
<td>Transcode, stream</td>
<td>Metadata wrappers</td>
<td>Insert into production workflow</td>
</tr>
</tbody>
</table>

Source: Artesia
Integration of Metadata Extraction and DAM

- Example: NorthPlains TeleScope DAM product integrating Virage VideoLogger, including direct encoding for streaming
- Further example product: Convera ScreeningRoom

Centralcasting

- Production of various programmes on various distribution channels out of a central content server
- Source: Artesia Technologies
Semantic Gap in Content/Media Management (1)


"... An approach that goes beyond representing what is being directly shown in a video or a movie, and aims to understand the semantics of the content portrayed and to harness the emotional, visual appeal of the content seen."

\[ P(a) = \alpha(W(s(n))) + \frac{\beta(m(n) - \mu_m)}{\sigma_m}. \]

Pace flow function

- \( n \) = shot number
- \( s \) = shot length
- \( m \) = motion magnitude
- \( \mu \), \( \sigma \) mean and standard deviation of \( m \)

Semantic Gap in Content/Media Management (2)