The Open Video Digital Library: The Challenge of Transition from Test Bed to Sustainable Library

Digital Humanities Lecture Series
Texas A & M
February 1, 2008

Gary Marchionini
march@ils.unc.edu
University of North Carolina at Chapel Hill
School of Information and Library Science
Interaction Design Lab
Theoretical View

- Digital video is crafted expression
  - Multiple channels (analog and digital)
  - Visceral as well as intellectual effects (analog and digital)
  - A descendant of film but with potential dynamics/behavior (digital)—changes over time, every time

- Digital Libraries are journeys (learning environments) rather than destinations for patrons and librarians
  - Beyond libraries as repositories to sharium

- Open Video deals with reusable (open) video objects
  - A journey toward new forms of expression and reflections on history
  - What do you do with 24/7 feeds of video from every street corner in Manhattan?
Digital Video Status

- Digital video a burgeoning DL challenge: YouTube phenomenon (fall 2007: 65K new videos/day; 20TB/mo; 100M views/day)
- Substantial research activity on storage, retrieval from engineering perspective (see IEEE, ACM MM)
- Many large-scale DLs and services
  - InforMedia, Fischlar, ECHO, Internet Archive, Open Video, public.tv, researchchannel
- Most attention on system/collection building rather than services
- Commercial attention on system and management
  - IBM, MERL, Microsoft, Artesia, Virage
- NIST TREC Video Track for retrieval evaluation
- Nice advances on capture, critical need for reuse tools
- Crucial need for evaluation that includes human factors
- Crucial need for sustainability models
  - Public service, advertising, pay-per-view, institutional repository?
Open Video Vision/Contributions

• An open repository of video files that can be reused in a variety of ways by the education and research communities
  – Open and encourage contributions, reuse
• An easy to use DL based upon the agile views interface design framework
  – Multiple, cascading, easy to control views (pre, over, re, shared, peripheral)
  – Views based upon empirically validated surrogates
  – An environment for building theory of human information interaction
• A set of methods and metrics to study how people understand digital video through surrogates
• Research agenda coupled to production service

Gary Marchionini, UNC
Background & Status

- Begun 1995 with colleagues at UMD & BCPS; current instance at UNC initiated in 1999
- Collaborators/Contributors: I2-DSI, ibiblio, CMU, UMD, Prelinger Archive, Internet Archive, NASA, ACM
- ~4000+ video segments
- ~40000 unique visitors per month
- ~1.8M hits/month
- MPEG-1, MPEG-2, MPEG-4, QT
- OAI provider
- Ongoing user studies
Backend Tools and Services

- Workstations, servers, disk arrays
- Tape players (VHS, Beta SP, PAL), boxes of tapes, DVDs (uggg)
- Final Cut Pro, Media Cleaner, Adobe Premier
- Bandwidth (UNC-CH switched ethernet)
- Linux OS, PHP scripting language, MySQL DBMS, Apache server
- SILS backup service, SRB preservation

Gary Marchionini, UNC
Backend Tools and Services (cont’)

- Merit (UMCP UMIACS), ported to Linux to extract candidate keyframes
- Speech to text (e.g., Sphinx at CMU)
- VAST keyframe/posterframe extraction, selection, and management
- Transaction logs and scripts (for evaluation and for recommenders)
- Peer to peer exchange
- ISEE (shared remote video use, e.g., DE)
- Indexer workstation (VIVO)
Tools and Services for User Studies

- Database driven web pages for user interaction
- Usability workstation (multiple camera, mixer, VCR)
- Eye tracking system; BioPac
- Speech synthesis (Cepstral ‘Amy’ license for audio keywords)
- Java and Perl scripts for managing, moving files, managing server (security, upgrades, etc.)
Video Browsing Interfaces for the Open Video Project (2001)
Overview of The Open Video Project.

Genre: Educational
Keywords: Open source; Digital video; Interface; Browse;
Duration: 04:52
Popularity (downloads): 341

Talking to the Ceiling: An Interface for Bed-Ridden Manually Impaired Users (1999)
Video demonstration from the 1999 CHI conference.

Genre: Educational
Keywords: Interface; Impaired; User; CHI;
Duration: 07:04
Popularity (downloads): 193

Manipulative User Interfaces: Exploring Physically Embodied User Interfaces (1999)
Video demonstration from the 1999 CHI conference.

Genre: Educational
Keywords: Interface; Manipulative; Explore; Embodied; User; CHI;
Duration: 03:31
Popularity (downloads): 62

DLGA - A Multimodal Interactive Information Assistant (1998)
Video demonstration from the 1998 CHI conference.

Genre: Educational
Keywords: DLGA; Interactive; Information; Assist; CHI;
<table>
<thead>
<tr>
<th>Title</th>
<th>Year</th>
<th>Duration</th>
<th>Genre</th>
<th>Popularity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Video Browsing Interfaces for the Open Video Project</td>
<td>2001</td>
<td>04:52</td>
<td>Educational</td>
<td>341</td>
</tr>
<tr>
<td>Talking to the Ceiling: An Interface for Bed-Ridden Manually</td>
<td>1999</td>
<td>07:04</td>
<td>Educational</td>
<td>193</td>
</tr>
<tr>
<td>Impaired Users</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manipulative User Interfaces: Exploring Physically Embodied</td>
<td>1999</td>
<td>03:31</td>
<td>Educational</td>
<td>62</td>
</tr>
<tr>
<td>User Interfaces</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OLGA - A Multimodal Interactive Information Assistant</td>
<td>1998</td>
<td>06:48</td>
<td>Educational</td>
<td>219</td>
</tr>
<tr>
<td>ambientROOM: Integrating Ambient Media with Architectural Space</td>
<td>1998</td>
<td>05:30</td>
<td>Educational</td>
<td>143</td>
</tr>
<tr>
<td>An Animated Direct-Manipulation Interface to Digital Library Services</td>
<td>1997</td>
<td>06:58</td>
<td>Educational</td>
<td>77</td>
</tr>
<tr>
<td>Technology at Home: A Digital Personal Scale</td>
<td>1997</td>
<td>01:40</td>
<td>Educational</td>
<td>199</td>
</tr>
<tr>
<td>A GUI Paradigm Using Tables, Two Hands and Transparency</td>
<td>1997</td>
<td>10:18</td>
<td>Educational</td>
<td>156</td>
</tr>
<tr>
<td>A Tour of Teamrooms</td>
<td>1997</td>
<td>08:40</td>
<td>Educational</td>
<td>159</td>
</tr>
</tbody>
</table>
Progress was made on structures appurtenant to the dam.
Hoover Dam Construction, segment 12 of 17

Progress was made on structures appurtenant to the dam.
Video Information

Year: 1996
Genre: Documentary
Keywords: 
Duration: 00:01:09
Color: Yes
Sound: Yes
Amount of Motion: Medium
Language: English
Sponsor: Bureau of Reclamation
Contributing Organization: Carnegie Mellon University, Informedia Project
Transcript Available: Yes

Digitization Information

Digitization Date: 1996
Digitizing Organization: Carnegie Mellon University Informedia Project
Agile Views Interface Research

- Provide a variety of access representations (e.g., indexes) and control mechanisms
  - Overview, preview, review, peripheral view, shared view
  - Text search plus view browsing (facets)
- Leverage both visual and linguistic cues to create surrogates for gisting
- Create and test surrogates in different views
Digital Video Surrogates

• Classes
  – Textual
  – Visual
  – Audio
  – Combined

• Cost benefit analysis: maximize ‘meaning’ per unit time (gist and vist)
  – Transmission time
  – Compaction rate
  – Cognitive processing time

• Performance vs. Preference

Gary Marchionini, UNC
User Interaction Research Framework

**GOALS**
Learning, work, entertainment

**TIME**
time spent searching and viewing results

**MENTAL LOAD**
perceptual load cognitive load

**PHYSICAL LOAD**
amount of muscle movement

**EFFORT**

**TASKS**
- select video for viewing
- select scene for viewing
- copy and use scenes
- copy and use frames
- other tasks?

**VIDEO CHARACTERISTICS**
genre: documentary, narrative
style: visual, audio, textual, place

**INDIVIDUAL CHARACTERISTICS**
domain experience
video experience
cultural experience
computer experience
info seeking experience
metacognitive abilities
demographics

**SURROGATES, AGILE VIEWS**
display controls
keywords
storyboard w/text, audio
slide show w/text, audio
fast forward w/audio
poster frames

**PERFORMANCE**
- retrieval (precision, recall)
- recognition (objects, action)
- gist comprehension (linguistic, visual)

**SATISFACTION**
- perceived usefulness
- perceived ease of use
- flow
- user satisfaction

**OUTCOMES**
domain experience
video experience
cultural experience
computer experience
info seeking experience
metacognitive abilities
demographics

**Gary Marchionini, UNC**
Surrogates Examined

- Storyboard with text keywords (20-36 per board@ 500 ms)
- Storyboard with audio keywords
- Slide show with text keywords (250ms repeated once)
- Slide show with audio keywords
- Fast forwards 32X, 64X, 128X, 256X
- Poster frames (1-3)
- Real time clips/excerpts (7 sec)
- Text (title, keywords, etc.)
- Visual features (e.g., in/out, people, etc.)
- Spoken descriptions
- Spoken keywords
- Combined visual (storyboard, fast forward) and spoken (descriptions, keywords)
## Tasks

<table>
<thead>
<tr>
<th>Recognition/Selection</th>
<th>Text</th>
<th>Still Image</th>
<th>Moving Image</th>
<th>Audio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Object selection</td>
<td></td>
<td>Object selection</td>
<td>Excerpt selection</td>
<td>Select Spoken Description, Select Spoken Keyword</td>
</tr>
<tr>
<td>(text)</td>
<td></td>
<td>(graphical)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Keyword selection</td>
<td></td>
<td>Keyframe selection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Description selection</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Title selection</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Generative Inference</th>
<th>Text</th>
<th>Still Image</th>
<th>Moving Image</th>
<th>Audio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gist writing</td>
<td></td>
<td>Visual gist determination</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(free text)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## Metrics

- Accuracy
- Confidence
- Time to complete
- Usefulness, usability, engagement, enjoyment, preferences

Gary Marchionini, UNC
User Studies

- Qualitative Comparison of Surrogates (Spring 02, ECDL 02)
- Fast Forwards (Fall 02, JCDL 03)
- Text or Pictures (Spring 03, CIVR 03)
- Narrativity (CHI 02, ASIST 03)
- Shared views and History Views (Geisler dissertation)
- TREC evaluation (Spring/summer 03; 05)
- ViSOR (Gruss Master’s paper)
- Look vs Read (Hughes Master’s paper)
- Video relevance (CHI 05; ASIST04; Yang dissertation)
- Cognitive load (Mu dissertation)
- Teachers using video (Brown dissertation)
- Spoken Audio and Storyboards (CHI 07)
- Spoken Audio and Fast Forwards (current)
Take Away Summary

• User studies inform design
• Give people multiple views and easy control mechanisms
• No silver bullets (many factors determine performance and preference)
• Words are powerful for gisting, visuals are liked and offer distinct value
• Video offers new kinds of potentials for learning and communication
Two Related Spin Offs
(aka how to sustain when funding ends)

• Sustain the OV Vision
  – Support student assistants
  – New research directions
  – Open information services

• What about the OV production system?
  – OV 2.0
  – Transition to Library?

Gary Marchionini, UNC
Video Preservation Project

- What kind and how much context to preserve?
- National Digital Information Infrastructure Preservation Program (NDIIPP) funding via NSF and LoC.
- Focus on specific topics
  - 2008 Presidential campaign (15K May 07-present)
  - Energy, truth commissions, health, pandemics
- Harvest video, metadata, and activity from YouTube; use API to query rather than crawl
- Create Curator’s tools and services
- **Fundamental DL issue of content/metadata/context boundaries in WWW objects**
Content, Metadata, & Context: Boundaries?

The World (Base context)

The World Time N context

Release Time

Time N
Information In Life Series for UNC YouTube Channel

www.youtube.com/uncchapelhill

• Google video hosting of subset of OV (2005-06)

• YouTube Education Channels
  – Information In Life series: lectures, interviews
  – Campus policies and logistics
Sustainability?

- Personal → Local → Global
  - How do people make sense of video?
  - How does video get integrated into interpersonal communication (e.g., mobile/ubiquitous)?
  - OV production system sustainability and preservation?

- The research & practice mobius strip works for a person, unclear whether this scales to institution or global levels, thus need for partnerships (TAMU Cervantes, industrial labs, etc.)

Gary Marchionini, UNC
Q&A

• Primary Investigators: Gary Marchionini, Barbara M. Wildemuth
  • Lead Developer: Gary Geisler
  • Ron Brown, April Edlin, Rich Gruss, Brenn Hill, Anthony Hughes, Xiangming Mu, Terrell Russell, Yaxiao Song, Larry Taylor, Tom Tolleson, Curtis Webster, Todd Wilkens, Meng Yang

Gary Marchionini, UNC