Recent research projects

James M Turner
professeur titulaire
Université de Montréal

Outline

• Research interests
• Audio description projects
• Variable media projects
• Localisation projects
• Next steps

Research interests

• Indexing moving images
• Storage and retrieval of still pictures
• Metadata for networked moving images
• Preservation of digital moving images
• Translingual storage and retrieval (a/v, mm)
• Audio description
• Localisation

Moving images

• General research problem: indexing shots by recycling text produced for other purposes:
  • subtitles
  • audio description
  • production notes
  • scripts broken down for the shoot
  • editors’ logs
  • camera reports
  • many others

Other aspects

web environment
W3C standards
Indexing moving images
multiple languages
open access
minimal human intervention

Previous work

• Determining the subject of still & moving images (PhD thesis, 1994)
• Comparing user-assigned terms with indexers’ terms for the same shots (1995)
• Audio description as a tool for indexing moving images (1998)
• Using shooting scripts for indexing moving images (2005)
• Using ancillary text to index cultural objects on the web (2006)
Audio description projects 1

- Audio description: voice track for blind and visually impaired (film, tv, theatre)
- Audio description as a tool... (1996-1999)
- The first ½ hour of 3 productions (This Old Pyramid, Poirot episode, Jurassic Park)
  - ½ of shots are described
  - description “episodes” mostly occur within 2 shots before or after
  - typology of things described (costume, action, etc.)

Audio description projects 2

- Sources of metadata... (2000-2003)
- Two feature length productions (Cher Olivier 1 & 2)
  - Production script, subtitles, audio description
  - Refined methodology for where shots occur
  - Refined typology

The E-Inclusion Research Network

- [http://e-inclusion.crim.ca](http://e-inclusion.crim.ca)
- Funded mainly by Canadian Heritage
- Partnerships with ÉTS, McGill, UdeM, ULaval
- Also CNIB, NFB, AudioVision, a dozen others
- 2005-2007:
  - access to audiovisual material for deaf/hearing loss
  - access to audiovisual material for blind/vision loss
- 2007-2009:
  - quality of description, testing video player

Research interests • Audio description • Variable media • Localisation • Next steps

Mostly low-level approaches

- shot detection
- scene detection
- face recognition
- voice recognition
- voice synthesis
- gait and gesture
- High-level approach: text manipulation

Research interests • Audio description • Variable media • Localisation • Next steps

Our sub-project

- Audio description for films and television
- Background: new CRTC regulations for digital channels
- As a pilot project, the NFB describing about 200 films (many already online)

Research interests • Audio description • Variable media • Localisation • Next steps

Audio description projects 3

- First E-Inclusion project (2006-2007)
  - analysing 11 films shot by shot (two feature length + NFB shorts)
  - some experimental (Norman McLaren)
  - refining typology (stable now)
  - viewing and discussion with blind and visually impaired people

Research interests • Audio description • Variable media • Localisation • Next steps
Objectives

- Validate typology of information elements
- Compare types of information in description to user needs
- Compare English text with French
- Recommendations/guidelines for describers

Method

- Analysis of 11 productions using method developed in previous projects
  - identify individual shots
  - transcribe audio description text
  - relate it to corresponding shots
  - write shot descriptions before viewing with sound

Productions analysed

<table>
<thead>
<tr>
<th>Productions analysed</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Production 1</td>
<td>Description 1</td>
</tr>
<tr>
<td>Production 2</td>
<td>Description 2</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

Shots and audio description

Typology with examples

Types of information given

- Action
- Information about attitudes of characters
- Decor
- Lighting
- Spatial relationships between characters
- Facial and corporal expressions
- Clothing
- Weather
- Movement of characters
- Physical description of characters
- Indicators of proportions
- Occupation, roles of characters
- Setting
- Description of sound
- Temporal indicators
- Text information in image
- Appearance of titles
- Credits
- Audio description

Research interests • Audio description • Variable media • Localisation • Next steps
Audio description & image description

- Here we want to compare:
  - keywords found in the audio description text with
  - keywords in a written description of each shot
- Objective: find out if one type of text is more fruitful than the other for generating indexing terms

Orange keywords appear in both

Analysis for 6 chapters

- Total 238 shots
- 50 shots (21%) have no keywords in common
- For the 188 remaining shots, 463 keywords appear in both types of text

Observations

- Essential keywords appear in both, other useful keywords in one or the other
- For indexing: one or the other text source ok, but both better

Improve this performance

- If queries got filtered through a thesaurus, synonyms could also be searched in the text
- This should improve performance
- However, performance is already quite good

Why this is good indexing

- Inter indexer consistency studies: the success rate is only about 80%
- Our own studies: art pictures require special knowledge to index, but everyday pictures do not
- User indexing, tagging à la flickr, YouTube, MySpace is widespread
- Some information science studies on this
Other aspects

- Since indexing by humans is so expensive, shot-level indexing will only happen if it is automated
- Exceptions: stockshot libraries, TV newsrooms
- Where we need to invest: automatically identify indexing terms, tag them, attach them to shots

Audio description projects 4

- Second E-Inclusion project (2008-2009)
  - testing video player (<videodescription.crim.ca/app/>)
  - viewing and discussion with blind and visually impaired people
  - 2 groups (completely blind / visually impaired)
- Completed 2009.03.31, writing reports now

Variable media projects

- Touched on this question in InterPARES2
- DOCAM (Documentation and conservation of the media arts heritage / SSHRC, Daniel-Langlois Foundation, partners, 2005-2009)
  - <www.docam.ca>, lots of information available

Variable media art

- Art that relies heavily on technology, changes over time
- Push computers past their limits, hack chips, write weird programmes, interact with physical objects, humans...
- How to buy, preserve, re-exhibit a big headache for museum people
- The key: copious metadata

DOCAM seminar

- People from Queens (master in art conservation) worked on DOCAM, but no courses offered in their programme
- DOCAM seminar (over 5 years, alternately English and French, McGill and UQÀM [=Université du Québec à Montréal])
- New home: UQÀM/UdeM museology programme (joint master’s)
- Clientèle: largely information studies students

Terminology committee

- Provide definitions for variable media terms (English & French)
  - Glossary
  - Thesaurus
  - Ontology?
- Using SKOS (Simple Knowledge Organization System)
- Glossary/thesaurus online, research reports 2009
Localisation projects

- MultiCulture (link from JT web site): documentation, glossary, resources in English, French, Arabic, Spanish
- IFLA/MIC project (IFLA AVMS with LC/AMIA/MIC)

What’s localisation?

- Translating is not enough, you need to include social, cultural aspects
- Make resources (computer software, web sites, etc.) usable to others
- Transcreation: not direct translation, rather adapting text to read naturally

MIC project (2006-2007)

- Moving Image Collections (mic.loc.gov) aims to be “a window to the world’s moving images”
- But the site is only available in English
- Funding from IFLA FAIFE (Committee on Free Access to Information and Freedom of Expression)

Localising the MIC site

- Objective: build a kit for localising to any language/ethnic group
- Test the kit with French, Arabic, Spanish
- AVMS partners in Montréal, Alexandria, Montevideo

Method

- We chose key pages on MIC site, hired research assistants to build the pages using the kit
- They got source code from MIC site, changed info between tags, found new pictures, cataloguing examples, identified issues
- Because of problems at MIC, localised pages not yet installed

Next steps

- Identify levels of description, types of information with XML tags, so audio description users can choose
- Assess to what degree audio description could be automated using existing text, voice recognition, gesture analysis, facial expressions, and so on
- Variable media programmes in universities
- Better identify cultural markers for localisation
- Build systems using hybrid approach (low-level and high-level indexing)
Video clip

- An interesting example of what's coming
- Good example of hybrid approach
- Propagating indexing terms: a new approach nobody would have dreamed of

The End
james.turner@umontreal.ca