DCU at VideoClef 2008
Vid2RSS Pilot Task

Eamonn Newman
Gareth Jones
• Vid2RSS task and data

• DCU @ VideoCLEF
  – System description
  – Results and Conclusions
  – Future Work
Goals of VideoCLEF benchmark evaluation

- Promote research on intelligent access to multilingual video collections
- Encourage exploitation of speech transcripts
- Encourage exploitation of video metadata

Develop and evaluate multilingual video analysis tasks

Extend the recent Cross-Language Speech Retrieval track with a new track for CLEF 2008
Data for Vid2RSS 2008

- **50 dual language videos** (30 hours) from The Netherlands Institute for Sound and Vision (Beeld en Geluid)
- Videos are episodes of Dutch television shows, mostly documentaries
- Dutch is the main language; English is an embedded language.
- Embedded language is spoken mainly by interviewees
- Videos are accompanied by Dutch-language archival metadata records
- Metadata includes series title, episode title, description, date of broadcast and other production information
- Speech recognition transcripts and Shot-level keyframes
Task requirements

Classification Task (Main Task)
- Assign videos to thematic classes using speech recognition transcripts only (required)
- Use combination of metadata and speech recognition transcripts to perform classification

Translation Task
- Translate output RSS-feeds (e.g., into English)

Keyframe Extraction Task
- Select a keyframe to provide a semantic representation of the entire video to be used to depict the video in the feed.
System Overview

- Information Retrieval system classifies videos into particular topic-based RSS feeds
- Based on Lucene Search and Indexing technology
- Designed to provide a baseline for the Vid2RSS classification task
System Functionality

- Free text index built using ASR transcripts and video metadata

- Feed labels were used as queries (e.g., “Architecture”, “Visual Arts”)

- Retrieved items added to that feed

- Some of the runs allowed an item to appear in one feed only, while others allowed items to appear in multiple feeds
Submitted Runs

- **Run 1** Dutch ASR transcripts; query only
- **Run 2** English ASR transcripts; query only
- **Run 3** Dutch ASR transcripts; relevance feedback on query
- **Run 4** English ASR transcripts; relevance feedback on query
- **Run 5** Catalogue metadata in Dutch; query only
## Results

<table>
<thead>
<tr>
<th>metric</th>
<th>Run 1</th>
<th>Run 2</th>
<th>Run 3</th>
<th>Run 4</th>
<th>Run 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>micro-average precision</td>
<td>0.50</td>
<td>0.32</td>
<td>0.16</td>
<td>0.17</td>
<td>0.83</td>
</tr>
<tr>
<td>micro-average recall</td>
<td>0.35</td>
<td>0.21</td>
<td>0.91</td>
<td>0.72</td>
<td>0.18</td>
</tr>
<tr>
<td>f-score micro-average</td>
<td>0.41</td>
<td>0.25</td>
<td>0.28</td>
<td>0.28</td>
<td>0.29</td>
</tr>
<tr>
<td>macro-average precision</td>
<td>0.54</td>
<td>0.62</td>
<td>0.42</td>
<td>0.50</td>
<td>0.93</td>
</tr>
<tr>
<td>macro-average recall</td>
<td>0.55</td>
<td>0.38</td>
<td>0.90</td>
<td>0.70</td>
<td>0.28</td>
</tr>
<tr>
<td>f-score macro-average</td>
<td>0.54</td>
<td>0.47</td>
<td>0.58</td>
<td>0.59</td>
<td>0.43</td>
</tr>
</tbody>
</table>
Results

• **Runs 1 and 2** Dutch ASR transcripts are more useful than English.

• **Runs 3 and 4** Relevance Feedback improves recall but downgrades precision.

• **Run 5** The catalogue metadata achieves high precision scores.
Future Work

• Refine relevance feedback techniques to maximise precision and recall.

• Examine system performance at query-level

• Investigate combination of ASR and catalogue metadata.
Class Labels

- Archaeology
- Architecture
- Chemistry
- Dance
- Film
- History
- Music
- Paintings
- Scientific Research
- Visual Arts