Analysis and Interface for Instructional Video

Alexander Haubold and John R. Kender
Department of Computer Science
Columbia University, New York
{ah297,jrk}@columbia.edu
Overview

- Motivation
- Segmentation by *Media Type*
- Clustering by *Content-based Matching*
- Visualization and Indexing of segmented data
- Demo
- Future Investigations
Motivation

• Video-taped courses currently made available to students with “keyframe” index
• “Keyframe” = snapshot of video at points of substantial change every 20-25 seconds
• Typical course length: 75 min per lecture, 26 lectures per semester = 32.5 hr of video data (3.5 megaframes!) => 5000 snapshots
• Need a more compact, content-directed indexing method for keyframes
Segmentation by Media Type

- Snapshots belong to six Media Types
- Media Type can be easily visualized

<table>
<thead>
<tr>
<th>Color</th>
<th>Media Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green</td>
<td>Board</td>
</tr>
<tr>
<td>Blue</td>
<td>Podium/Instructor</td>
</tr>
<tr>
<td>Red</td>
<td>Computer</td>
</tr>
<tr>
<td>Yellow</td>
<td>Hand-drawn Sheets</td>
</tr>
<tr>
<td>Orange</td>
<td>Printed Media (Illustrations)</td>
</tr>
<tr>
<td>Cyan</td>
<td>Class</td>
</tr>
</tbody>
</table>

(see paper for decision tree classifier)
Clustering by Content Matching

- Snapshot content evolves slowly
- Hand-drawn slides grow monotonically
- Blackboard panels varies dynamically

(see paper for content pixel filters, pan and zoom detectors, global match consistency heuristics)
Interface and Visualization
Interface and Visualization: Index

• Through analysis: temporal topic model
• Through user study: perception is topologic
• So: Topological View with relative time scale, each topic represented by media icon, interrupted topics reunited by tapered lines
• Verified: users quickly identify key concepts!
Interface and Visualization: Details

- Through user study: need content in context
- So: Key Frame View with full index of key frames, organized by *Media Type* and *Topic*
- Verified: users quickly access full-size images and the video!
Demo
Numbers

- First semantic clustering of extended collection of videos
- 17 videos
- 40 hours of lecture
- 4479 keyframes
- 334 content-matched topics
Future Investigations

- Indexing over all videos in a course and all courses in a semester
- Automatic textual annotations for topics
- More robust segmentation and faster content matching
- User studies: patterns of retrieval -> automatic content queuing