A Touch-panel based User Interface and Utilization of User’s Memories for Known-item Search (KIS) Task in TRECVID 2011

Yuan Zhou and Takashi Yukawa
Nagaoka University of Technology
2011-12-05
Interactive KIS Task

**Goal**

**Output**
- 1 single video

**Evaluate**
- Mean Elapsed Time (<5mins)
- Mean Average Precision
- User Satisfaction (1-7)

**Data Collection:**

**Training Data** (with XML)
- ~3200 videos
- 47.2 GB
- 200 hours
- 3.6 – 4.1mins/video

**Test Data** (with XML)
- ~8000 videos
- 50 GB
- 200 hours
- 10s – 4.1mins/video

**Query:**

**391 Queries in test:**

Find the video of a man in white T-shirt, blond hair, talking about FBI

Find the video of a baby in stroller being pushed by a woman in blue pants in a house with hardwood floors.

Find the video showing a hand next to a computer keyboard and a cat's tail and paw
Feature(1): User Interface

• Description:

• Solution:
  – iPad as UI
  – Intuitive, Easy to use

• Goal:
  – Better mean elapsed time and user satisfaction.
Feature(2): Visualizing Weighted Terms

• Description:
  – Salient Term, Computer Only, NOT Effective.

• Solution:
  – Let user decide query terms’ weight, In Visual.

• Goal:
  – Better Engine Performance & User Satisfaction.
Visual Level

Salient Words Selection

How Long is the Video?

Long  Medium  Short
Feature(3): Length-check

• Description:
  – User remember video’s length roughly.

• Solution:
  – Add “short”, “medium”, “long” functions.

• Goal:
  – Better Engine Performance.
Visual Level

Short: less than 20 seconds
Medium: 21-59 seconds
Long: more than 1 min
Index Process

XML

- identifier
- title
- description

Lucene

Index

```xml
<?xml version="1.0" encoding="UTF-8"?>
<metadata>
  <identifier>2ndSHEDUnderwaterHockeyTournamentFinals</identifier>
  <title>2nd SHED Underwater Hockey Tournament Finals</title>
  <creator>www.usauwh.com</creator>
  <mediatype>movies</mediatype>
  <collection>sports</collection>
  <description>2nd SHED Underwater Hockey Tournament Finals, Located in San Jose CA - 2005</description>
  <date>2005</date>
  <year>2005</year>
  <year>2006</year>
  <subject>Underwater Hockey; H2ockey; Octopush; San Jose; SHED; Tournament; Finals</subject>
  <licenseurl>http://creativecommons.org/licenses/by/2.5/</licenseurl>
  <publicdate>2006-07-07 05:15:07</publicdate>
  <adder>duereg@hotmail.com</adder>
  <uploader>duereg@hotmail.com</uploader>
  <publisher>www.usauwh.com</publisher>
  <sound>sound</sound>
  <color>color</color>
  <resource>movies</resource>
  <license>http://creativecommons.org/licenses/by/2.5/</license>
  <format>Video</format>
  <keywords>SHED, Underwater Hockey, Tournament, San Jose, H2ockey, uwh, Octopush</keywords>
  <holder>Matt Blair</holder>
  <updatedate>2006-09-05 16:00:47</updatedate>
  <updater>duereg</updater>
</metadata>
```
Search Process

query

Lucene Search Engine (server)

Salient Word

Salient Word (refined)

Video Length info

Long index

mid index

Short index

video data

Server

Client (iPad)
Network

Client → Tomcat (Java Servlet) → Server

Http(Post)

Video Data → Index
iPad User Interface...
Video
Interesting Issue

10/25 have keywords

8/10 by Default Lucene

Salient word: Children
Video Length → long

Video length → short

<table>
<thead>
<tr>
<th>Rank change</th>
<th>Topic No.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>501</td>
</tr>
<tr>
<td>Pre.</td>
<td>2</td>
</tr>
<tr>
<td>Aft.</td>
<td>1</td>
</tr>
</tbody>
</table>
## Result:

<table>
<thead>
<tr>
<th>Index &amp; Methods combination</th>
<th>25 Test Topics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean Precision</td>
</tr>
<tr>
<td>XML</td>
<td>0.320</td>
</tr>
<tr>
<td>+ Salient Word Refining &amp; Length Verification</td>
<td>0.400</td>
</tr>
</tbody>
</table>

### MP & MIR comparison:

<table>
<thead>
<tr>
<th>Index &amp; Methods combination</th>
<th>25 Test Topics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean Precision</td>
</tr>
<tr>
<td>XML</td>
<td>0.320</td>
</tr>
<tr>
<td>+ Salient Word Refining &amp; Length Verification</td>
<td>0.400</td>
</tr>
</tbody>
</table>
Summary

• TRECVID 2011:
  – Interactive KIS task
• Index data: only metadata
• 3 issues:
  – iPad
  – Salient word refining
  – Video Length Check
• Index and Search process
• Results.
Future Work

• Performance enhancement:
  – OCR, ASR index implementation.
  – Voice-aid Query Input.
Thank You

Question?