Combining Content-based Analysis and Crowdsourcing to Improve User Interaction with Zoomable Video

Axel Carlier, Vincent Charvillat
University of Toulouse (France)

Ravindra Guntur, Wei Tsang Ooi
National University of Singapore
Outline

• What is zoomable video?
• Recommending zoomable regions to users
• Recommendations based on content analysis
• Improving recommendations with users' input
What is Zoomable Video?
Interest of zoomable video

- Video resolutions: up to 7680 x 4320 pixels
  - Bandwidth
  - Display

[Khiem10] Supporting zoomable video streams with dynamic region-of-interest cropping
Issues with our previous interface

- Incomplete understanding of the video
- Too much interaction required from users
- Moving regions
The problem with moving regions

Interest

Zoom in
Recommendations of zoomable regions

- Idea: give users recommendations
  - To guide them
  - Reduce the number of interactions
  - Provide a better experience to users
Recommendations

- To present useful recommendation, we need to:
  - Detect regions of probable interest to the user
  - Present the recommendations in an aesthetic way

Use content analysis
Saliency Detection
Motion Detection
Important Object Detection
(face, body, etc.)
Importance Map
Analyzing Importance Maps

- Bright spots: saliency and/or motion and/or importance maps
  - Interesting regions
Result
# User Study

<table>
<thead>
<tr>
<th></th>
<th>No Interface</th>
<th>No Recommendation</th>
<th>Recommendations based on Content analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Users</td>
<td>10</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

**NI - No Interface:** we only provided users with a scaled down version of the videos.

**NR – No Recommendation:** we provided users with the old version of our interface, without any recommendations.

**RC – Recommendations based on Content analysis:** we provided users with our new interface including recommendations.
Videos and Protocol

- Tutorial on how to use our interface
- Task
Results – Number of interactions

Number of interactions per view of the video

- **LongJump2**
  - No Recommendations: 3
  - Content-based recommendations: 2

- **Coffeelounge**
  - No Recommendations: 8
  - Content-based recommendations: 6
Results – Usage of recommendations

<table>
<thead>
<tr>
<th></th>
<th>Long jump</th>
<th>Coffee lounge</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC</td>
<td>18%</td>
<td>45%</td>
</tr>
</tbody>
</table>

Percentage of zoom using recommendations over the total number of zooms

<table>
<thead>
<tr>
<th></th>
<th>Zoom Level 1</th>
<th>Zoom Level 2</th>
<th>Zoom Level 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC</td>
<td>73.2%</td>
<td>25.6%</td>
<td>1.2%</td>
</tr>
</tbody>
</table>

Size of the recommended viewports clicked by users on **Coffee lounge**
Results – Understanding of the video

Percentage of good answers to the question

Is there a key stolen?

- NI
- NR
- RC
Results – Understanding of the video

Percentage of good answers to the question

- NI
- NR
- RC

Are the wallets exchanged?
Observations

• Pretty bad results:
  • Understanding of the video less good
  • Regions not very informative (not very clicked)...
  • ... And preventing users from looking at the interesting regions!
  • Less interactions
Adding users' input to improve the recommendations
Getting Users' Input
Getting Users' Input
New importance maps
### User Study

<table>
<thead>
<tr>
<th>Number of Users</th>
<th>No Interface</th>
<th>No Recommendations</th>
<th>Recommendations based on Content Analysis</th>
<th>Recommendations based on Content + Usage analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

**NI - No Interface:** we only provided users with a scaled down version of the videos.

**NR – No Recommendation:** we provided users with the old version of our interface, without any recommendations.

**RC – Recommendations based on Content analysis:** we provided users with our new interface including recommendations.

**RC+U – Recommendations based on a combination of Content and Usage analysis:** the final version of our interface.
Results – Number of interactions

Number of interactions per view of the video

<table>
<thead>
<tr>
<th></th>
<th>No Recommendations</th>
<th>Content-based recommendations</th>
<th>Content+Users based recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td>LongJump2</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Coffeelounge</td>
<td>7</td>
<td>6</td>
<td>4</td>
</tr>
</tbody>
</table>
Results – Usage of recommendations

<table>
<thead>
<tr>
<th></th>
<th>Long jump</th>
<th>Coffee lounge</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC</td>
<td>18%</td>
<td>45%</td>
</tr>
<tr>
<td>RC+U</td>
<td>40%</td>
<td>55%</td>
</tr>
</tbody>
</table>

Percentage of zoom using recommendations over the total number of zooms

<table>
<thead>
<tr>
<th></th>
<th>Zoom Level 1</th>
<th>Zoom Level 2</th>
<th>Zoom Level 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC</td>
<td>73%</td>
<td>26%</td>
<td>1%</td>
</tr>
<tr>
<td>RC+U</td>
<td>25%</td>
<td>42%</td>
<td>33%</td>
</tr>
</tbody>
</table>

Size of the recommended viewports clicked by users on **Coffee lounge**
Results – Understanding of the video

Percentage of good answers to the question

- **Is there a key stolen?**
  - NI: 70%
  - NR: 80%
  - RC: 60%
  - RC+U: 90%

- **Is there a wallet stolen?**
  - NI: 50%
  - NR: 70%
  - RC: 50%
  - RC+U: 80%
Summary of the results

• Recommendations based on content and usage analysis:
  • Improve users' understanding of the video
  • Decrease the number of users' interactions
• Bad recommendations can get a negative effect on users' understanding of the video
Conclusions

• We improved a zoomable video interface by proposing recommendations to users, leading to:
  • Better ergonomics
  • Better understanding of the video
• Combination of users' inputs with content analysis brings interesting results
Future Work

• Further exploration of this combination of content and usage analysis:
  • In other problems (e.g. 3D viewpoint determination)
  • With better mathematical models