

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

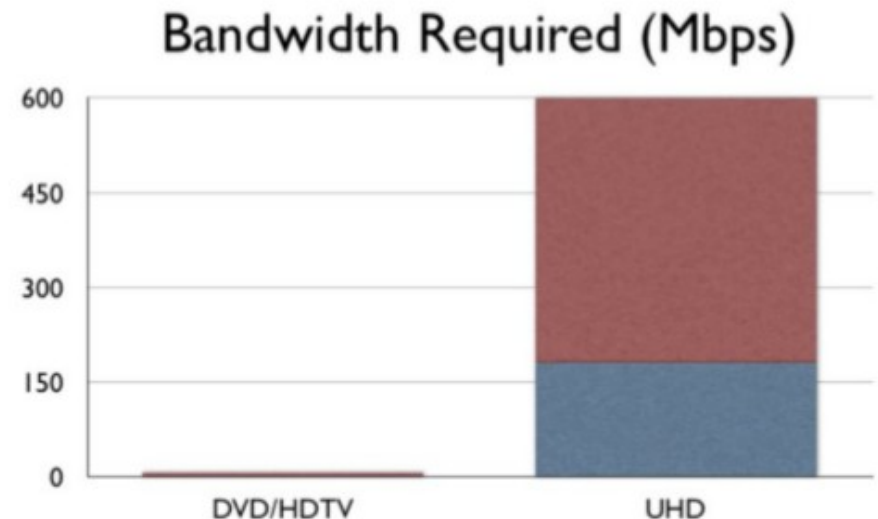
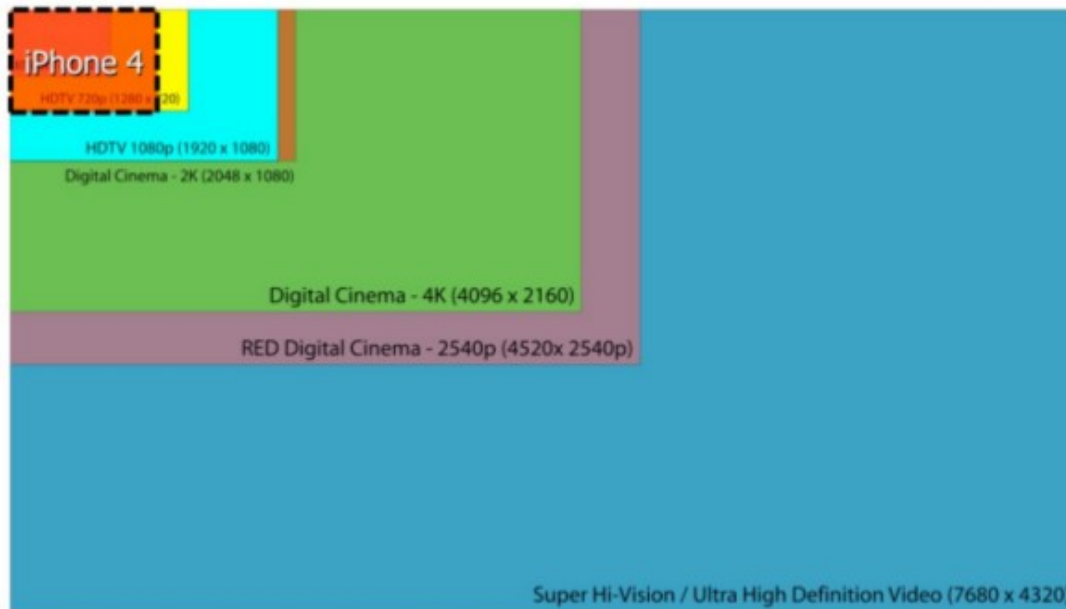
- What is zoomable video ?
- Recommending zoomable regions to users
- Recommendations based on content analysis
- Improving recommendations with users' input

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

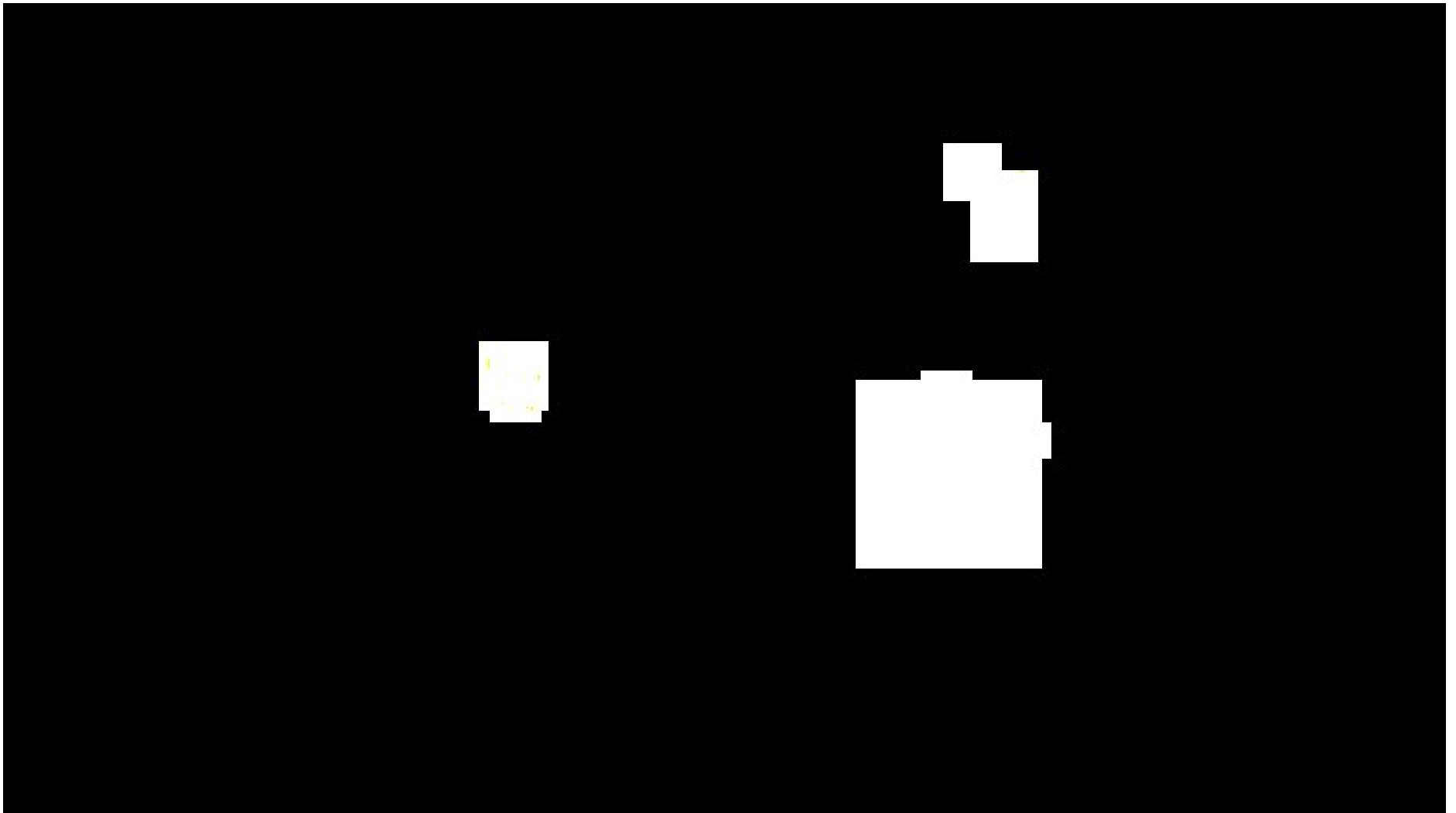
Saliency Detection



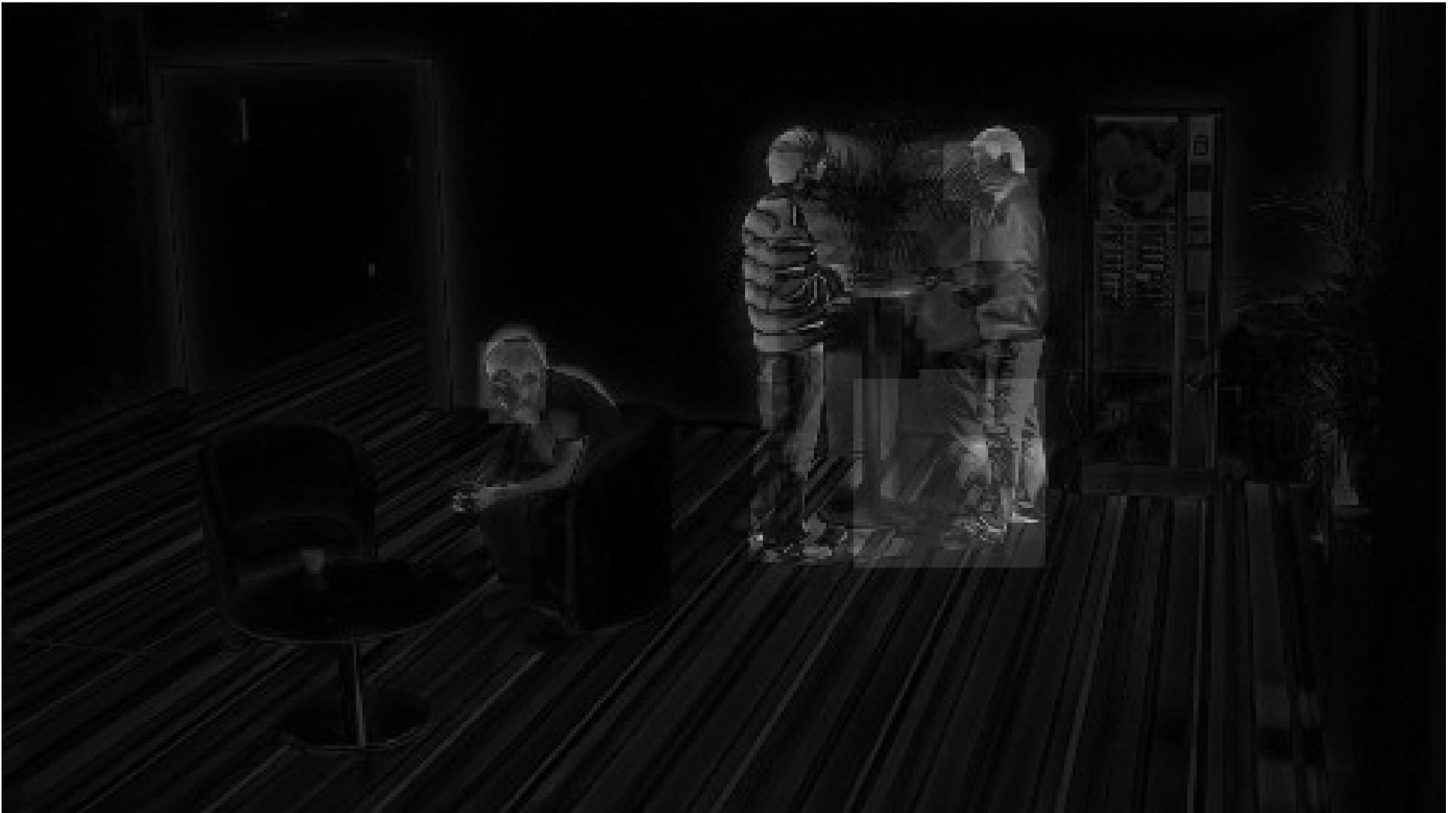
Motion Detection



Important Object Detection (face, body, etc.)

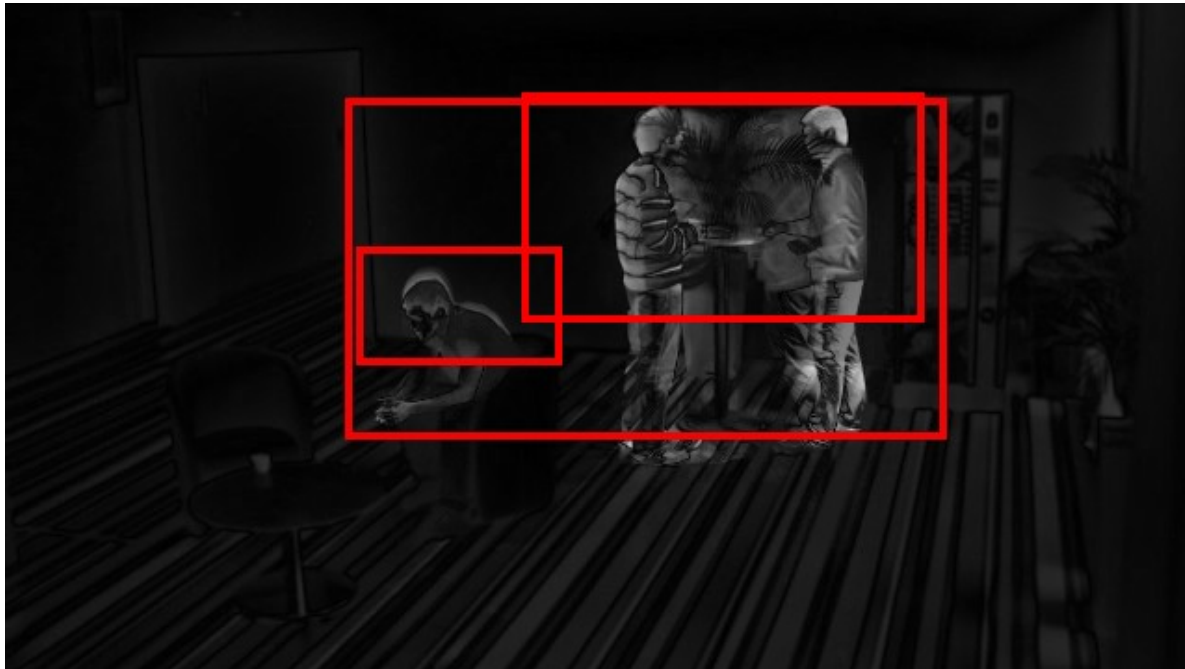


Importance Map



Analyzing Importance Maps

- Bright spots: saliency and/or motion and/or importance maps
 - Interesting regions



Result



User Study

	No Interface	No Recommendation	Recommendations based on Content analysis
Number of Users	10	20	20

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.

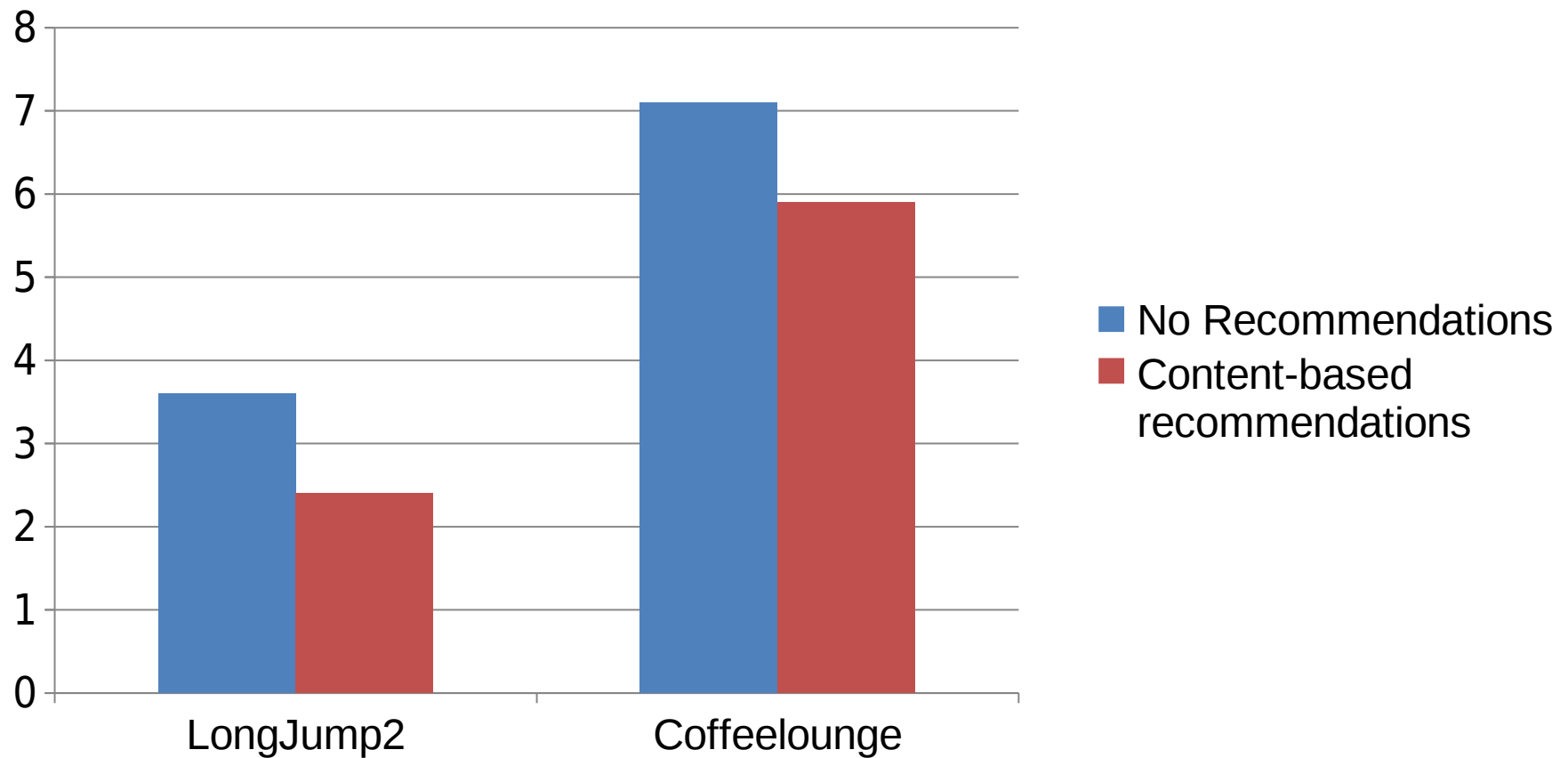
Videos and Protocol



- Tutorial on how to use our interface
- Task

Results – Number of interactions

Number of interactions per
view of the video



Results – Usage of recommendations

	Long jump	Coffee lounge
RC	18%	45%

Percentage of zoom using recommendations over the total number of zooms

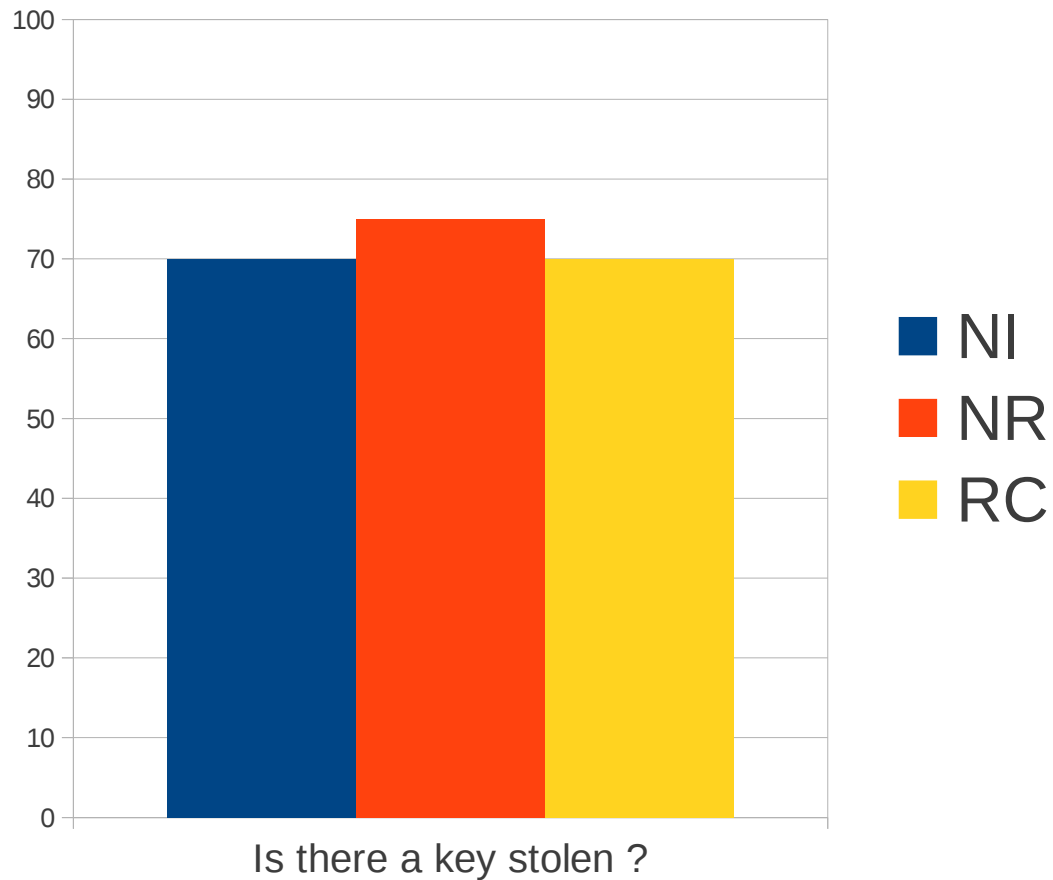
	Zoom Level 1	Zoom Level 2	Zoom Level 3
RC	73.2%	25.6%	1.2%

Size of the recommended viewports clicked by users on **Coffee lounge**



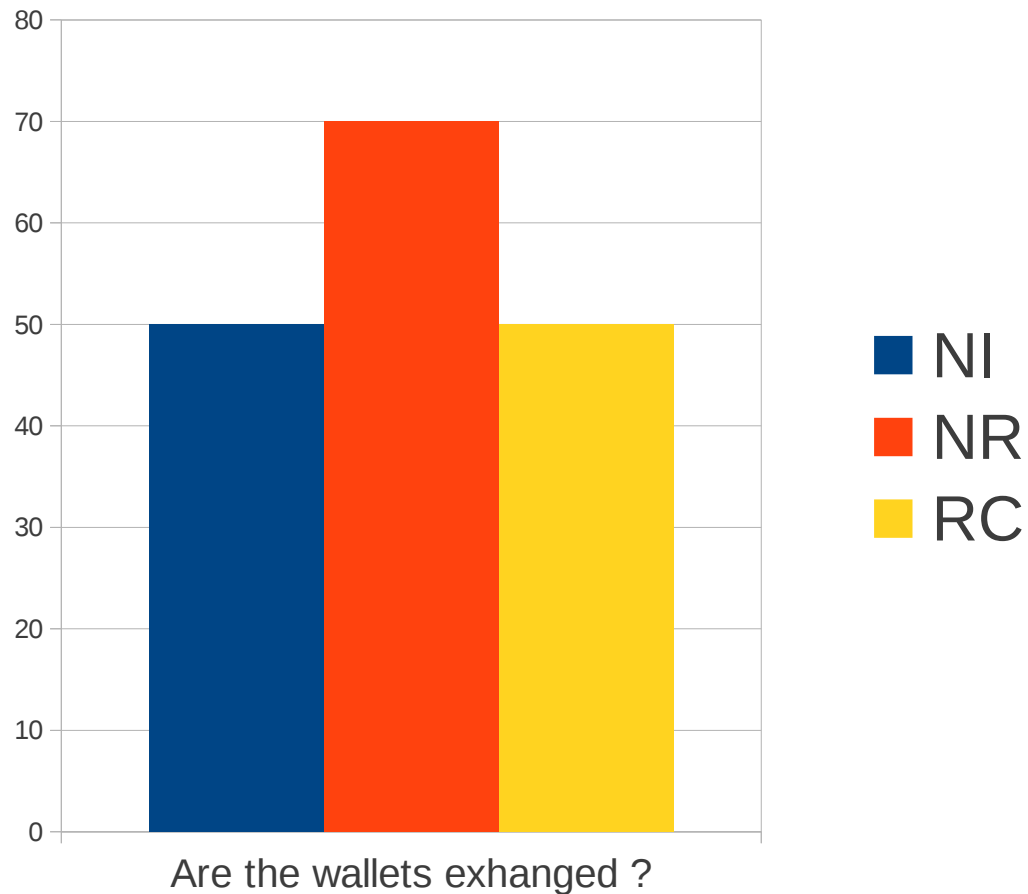
Results – Understanding of the video

Percentage of good answers to the question



Results – Understanding of the video

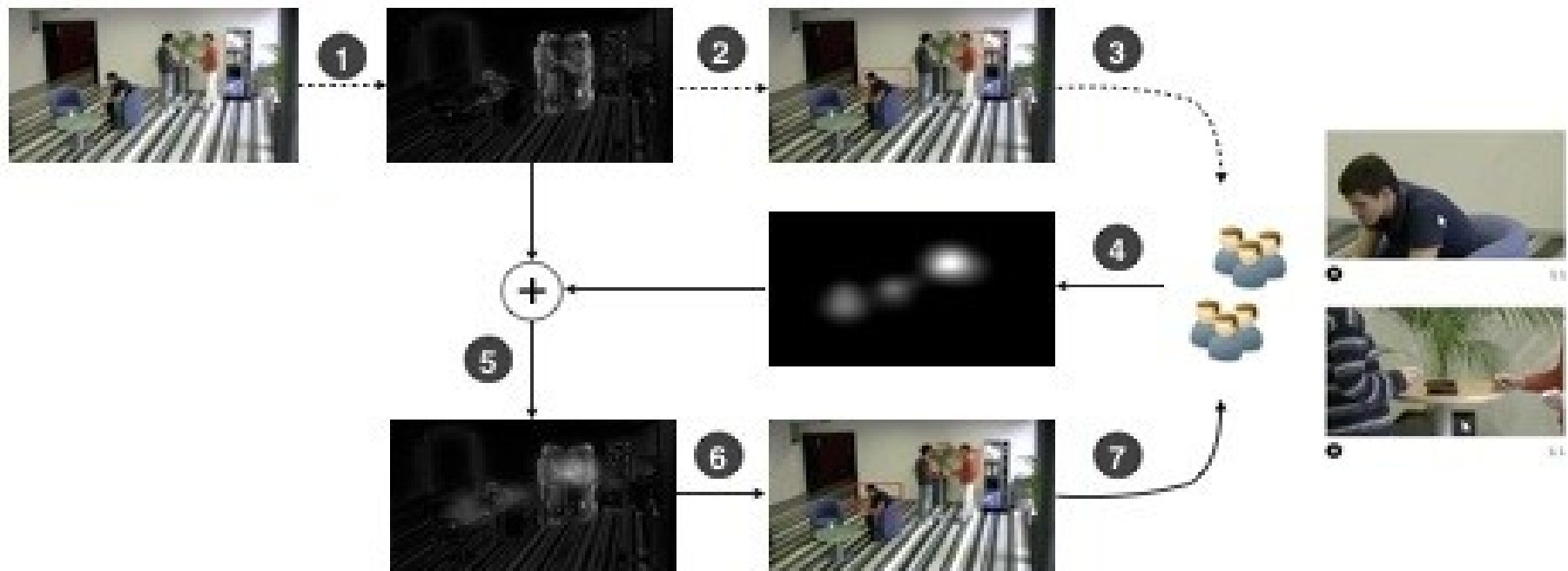
Percentage of good answers to the question



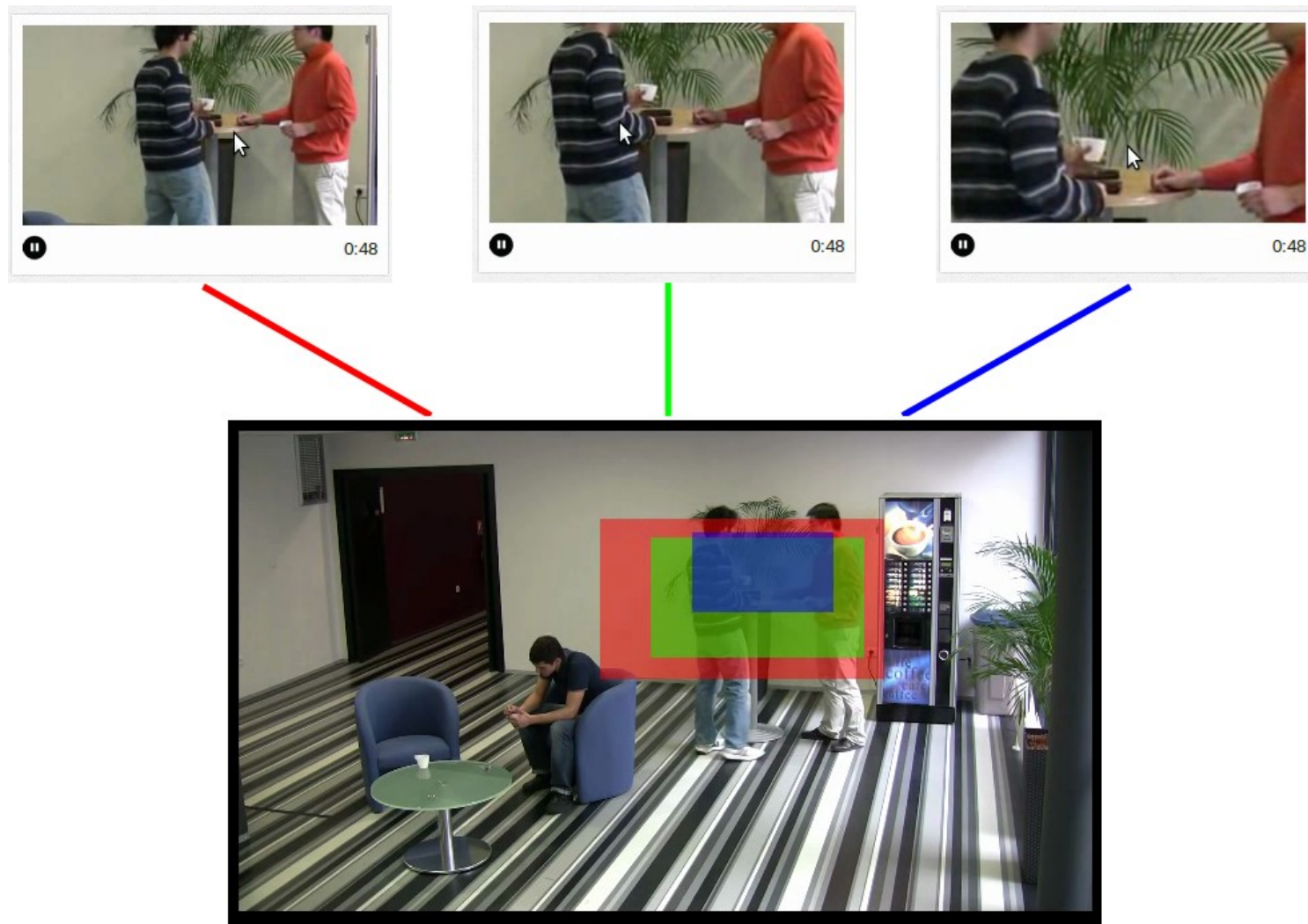
Observations

- 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

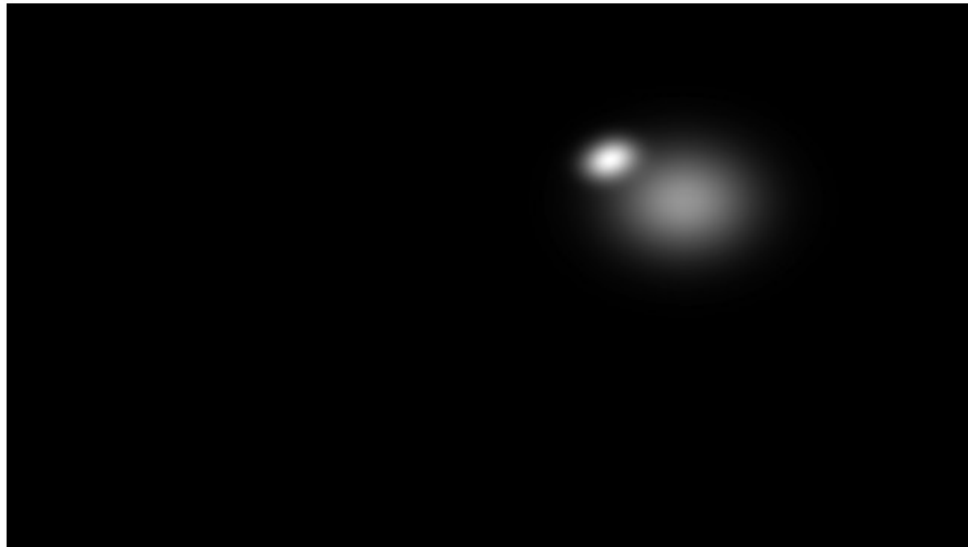
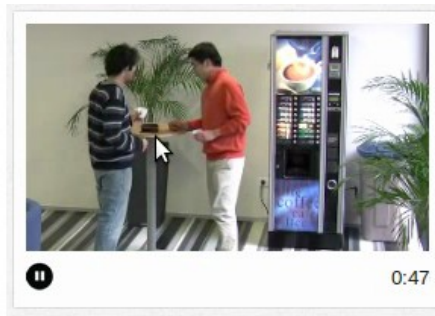
Adding users' input to improve the recommendations



Getting Users' Input



Getting Users' Input



New importance maps



User Study

	No Interface	No Recommendations	Recommendations based on Content Analysis	Recommendations based on Content + Usage analysis
Number of Users	10	20	20	20

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

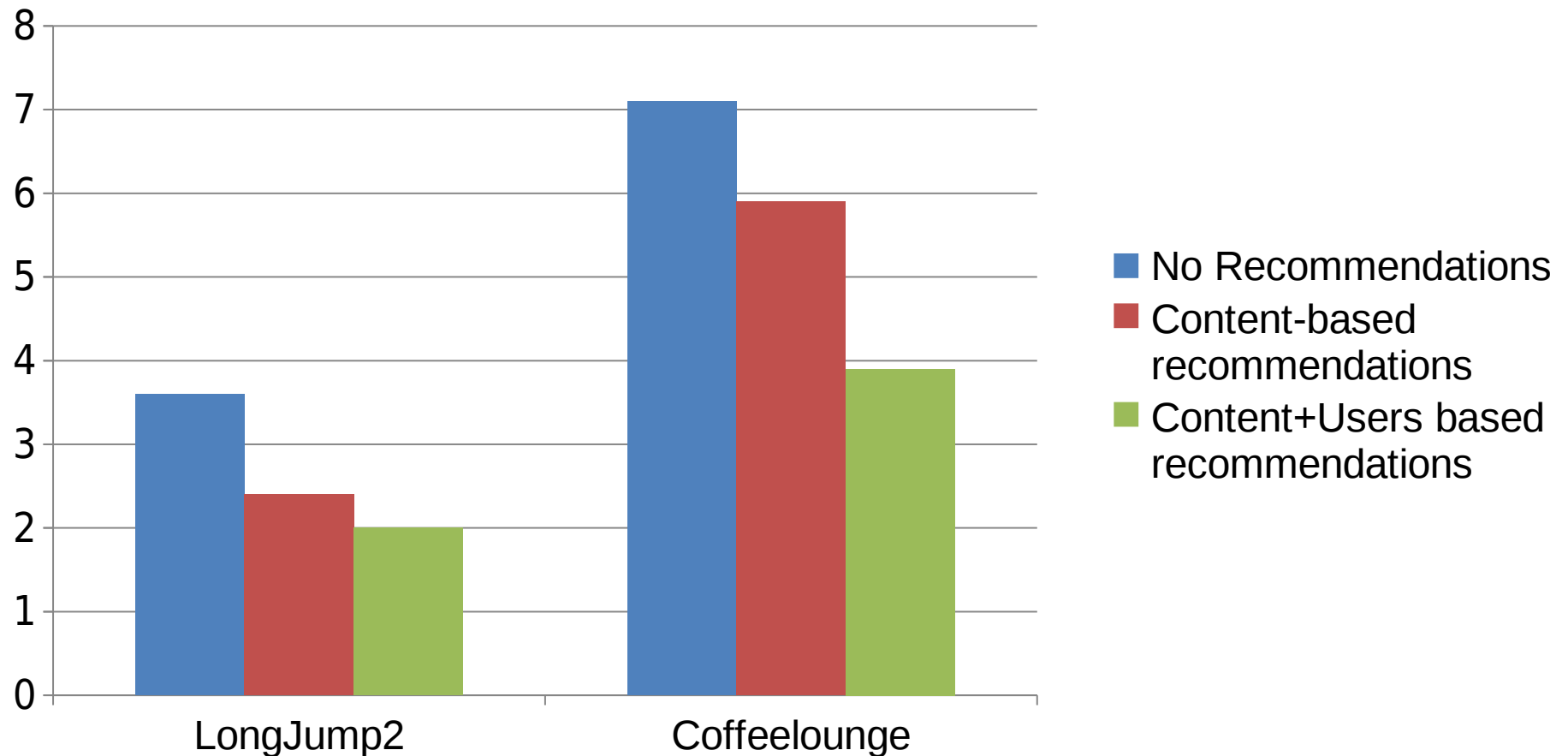
NR – No Recommendation: we provided users with the old version of our interface, without any recommendations.

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



Results – Usage of recommendations

	Long jump	Coffee lounge
RC	18%	45%
RC+U	40%	55%

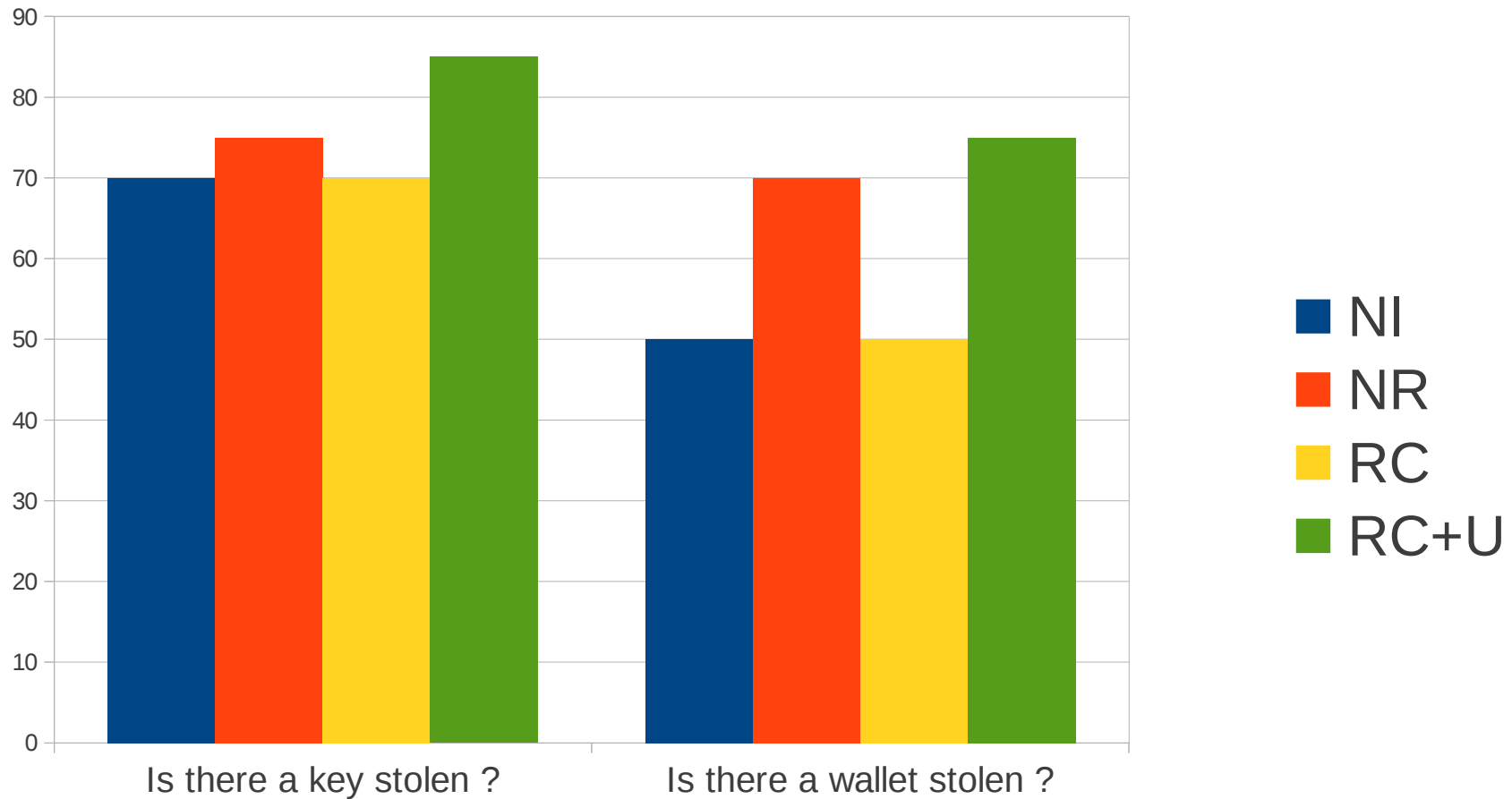
Percentage of zoom using recommendations over the total number of zooms

	Zoom Level 1	Zoom Level 2	Zoom Level 3
RC	73%	26%	1%
RC+U	25%	42%	33%

Size of the recommended viewports clicked by users on **Coffee lounge**

Results – Understanding of the video

Percentage of good answers to the question



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 others problems (e.g. 3D viewpoint determination)
 - With better mathematical models