

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

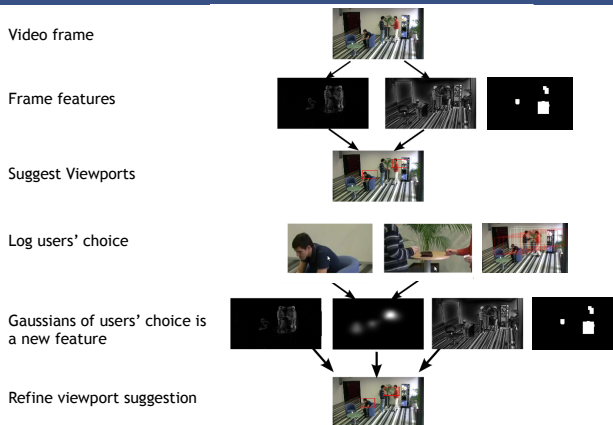
Axel Carlier* Guntur Ravindra⁺ Vincent Charvillat* Wei Tsang Ooi⁺

ENSEEHT, University of Toulouse*, School of Computing, National University of Singapore⁺

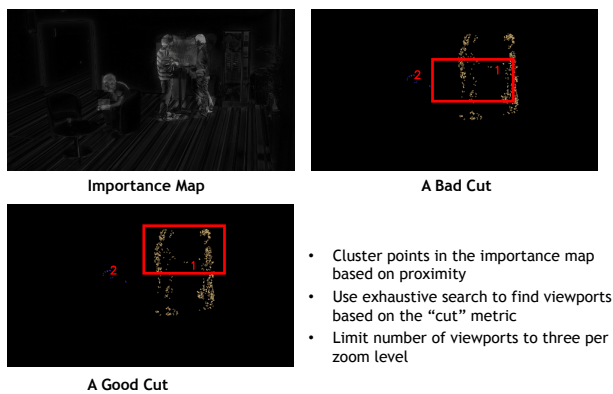
ABSTRACT

- Zoomable video allows users to zoom and pan in streamed video
- A dynamic scene with moving objects motivates users to frequently interact with the video
 - Frequent interaction can be irritating to the users
 - Users may not be fast enough to follow objects of interest in a rapidly changing scene
 - Frequent user interaction can be supported only if the end-to-end system has a low reaction time
- Content-based analysis of the video scene can be used to identify interesting regions in the video
 - Users are presented with automatically generated viewports at different zoom levels
 - The user interface automatically zooms and pans within the scene when a user clicks on a recommended viewport
- The user interface allows for overriding suggested viewports
- Users' viewport selection is combined with machine generated viewports to refine the interesting region

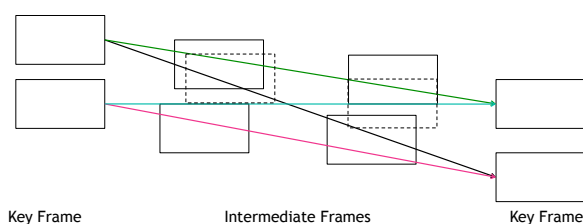
OVERVIEW OF THE SYSTEM



CUTS AND VIEWPORT GENERATION



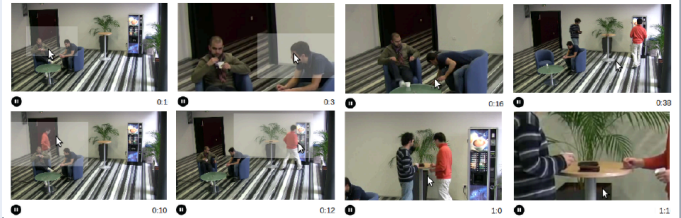
TUBING AND PANNING



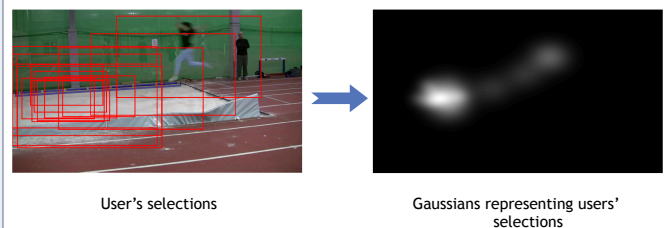
For all pairs of paths between viewports in two consecutive key frames

- Find a path that minimizes a "tubing metric"
- The tubing metrics are "cluster cuts", "heat", "regularity" and "coherence"
- Repeat for each zoom level

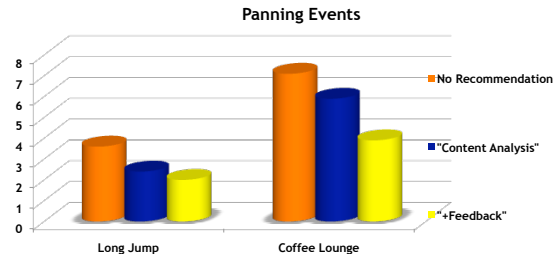
RENDERING



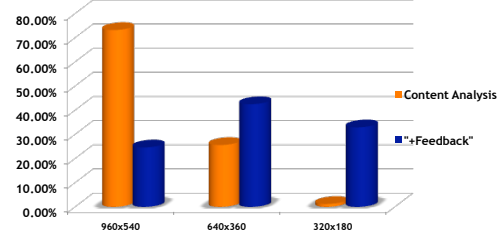
COMBINING IMPLICIT FEEDBACK WITH CONTENT ANALYSIS



RESULTS



Distribution of Recommended Viewport Selection



User Selection of Viewports	Long Jump	Coffee Lounge
Content Analysis	18%	45%
+Feedback	40%	55%

Interface	No Zoom	No Recommendation	Content Analysis	+Feedback
Key Stolen?	70%	75%	70%	85%
Wallet stolen?	50%	70%	50%	75%

ACKNOWLEDGEMENTS

Sincere thanks to our colleagues Arash Shafiei, Julien Badie, Salim Bensiali