Video Summarization via Crowdsourcing

Shao-Yu Wu¹, Ruck Thawonmas², Kuan-Ta Chen¹
¹Academia Sinica, ²Ritsumeikan University

Goals

• Generate satisfactory summaries for different types of video clips through a lightweight method
• Generate different abstraction levels of video summaries according to viewers’ preferences in real time

Concepts

• Leverage the wisdom of the crowd to generate video summaries with a low workload for workers
• Viewers click the SPACE key whenever they feel that a particular part of the video should be included in a video summary

Experiment Instructions

1. Watch and Click
   During this study, you need to complete 3 tasks. In each task, a video clip will be played 2 times. Please watch each video carefully from the beginning to the end. Whenever the video reaches its highlights, please press SPACE key on your keyboard.

2. How to determine Video Highlights?
   Imagine that you are a director and editing a trailer for a video clip, the video frames or segments that you perceive worth to be included in the trailer are called Video Highlights.

3. How to determine Video Highlights? (Cont.)
   Each Video Highlight can be instantaneous or spanning a period of time.
   - If it is instantaneous, click SPACE key and release it immediately.
   - If it spans a period of time, press SPACE key and only release it until the end of the highlight.

4. Watch and Click - Example
   When SPACE key is pressed, the color of the border of the video player will be changed from white to red. It will be back to normal when SPACE key is released.

Summarization Algorithm

• video clip: \( n \) non-overlapping video shots
• \( t \): timestamp
• collected \( k \) samples: \( t_1, t_2, \ldots, t_k \)
• viewer-preferred summary length: \( T \)

1: compute the histogram, \( H \), of \( t_1, t_2, \ldots, t_k \)
2: \( summary = \emptyset \)
3: while video length (summary) < \( T \) do
4: \( t_{old} = \arg \max H(t), \forall t \text{ not contained in } summary \)
5: \( shot_{old} = \text{the video shot containing } t_{old} \)
6: \( summary = summary \cup shot_{old} \)
7: end while

• output: \( summary \) (set of video shots that comprise the video summarization result )

Generated Summary Segments

• Summary segments highlighted in blue were generated from the collected samples (i.e., the timestamps when the participants pressed the SPACE key) by using our algorithm

Evaluation

• Within-subjects user study (N=81)
• Three genres of video source: News, Sport, Commercial
• Participants were asked to rate the summary segments generated from our method (crowdsourcing), experts, and subsampling with the following instruction:
“Please rate the summary segments that you just watched in terms of how much they helped you make sense of the full video.”

Result

Future Work

• Prepare for a wider-scale deployment of this framework
• Increase the data sets of video source for more general evaluation
• Investigate how to encourage users to participate in video summarization for non-monetary rewards