

Using Temporal Video Annotation as a Navigational Aid for Video Browsing

Stefanie Müller

Dept. of Automation and Computer Science
University of Applied Sciences Harz
smueller@ece.ubc.ca

Gregor Miller, Sidney Fels

Dept. of Electrical and Computer Engineering
University of British Columbia
{gregor, ssfels}@ece.ubc.ca

ABSTRACT

Video is a complex information space that requires advanced navigational aids for effective browsing. The increasing number of temporal video annotations offers new opportunities to provide video navigation according to a user's needs. We present a novel video browsing interface called TAV (Temporal Annotation Viewing) that provides the user with a visual overview of temporal video annotations. TAV enables the user to quickly determine the general content of a video, the location of scenes of interest and the type of annotations that are displayed while watching the video. An ongoing user study will evaluate our novel approach.

ACM Classification: H5.2 [Information interfaces and presentation]: User Interfaces. - Graphical user interfaces.

General terms: Design, Human Factors

Keywords: Video Browsing, Video Navigation, Video Annotation, Video Search

INTRODUCTION

Although video annotation is a growing phenomenon, the current approach for viewing and finding video annotations is still at a basic level. Even though many annotations refer to a specific subset (or scene) of the video and are therefore temporal in nature, the established approach does not take this characteristic into account. On popular video platforms such as YouTube, video comments (a widely used type of annotation) are displayed in a single entry list that does not change during playback. However, we believe that this static approach is not appropriate for a time-based medium such as video. We propose that video annotations be displayed simultaneously with the scene to which they refer.

Furthermore, enabling the user to add temporal video annotations also provides new opportunities for browsing video. The static approach for visualizing video annotations cannot be easily used to provide navigational cues. However, with a temporal approach video annotations can be used to

support the user with additional information such as the scene's content and locations of interest. We believe that providing the user with a visual overview of temporal video annotations improves video navigation speed.

In addition to the problem of non-temporal video annotation there is also a lack of filtering mechanisms. The increasing number of annotations requires an interface that enables the user to define which annotations are relevant. For example, only the annotations made by a specific person (such as a friend) or the annotations related to a specific event in the video (a goal in a sporting event).

In the following sections we present related work in the field of video browsing with visual cues and introduce our novel approach called TAV.

RELATED WORK

The existing literature contains many examples of research on visualizing the underlying video content either by extracting image-based features such as dominant color or motion [3], using sound volume or by interpreting video content, for example in the form of visualizing emotions [1]. However, none of this work has focused on using temporal video annotation as a means to provide navigational aids for browsing video. On social video sites such as Vidler [4] temporal video annotation is displayed as black dots in the video timeline. However, these navigational cues do not provide the user with information on the underlying video content. We believe that enhanced visual representations of temporal video annotation will provide effective navigational support while browsing video.

Furthermore, current visualization approaches do not support the user with effective filter mechanisms. Costa et al. [2] emphasized how temporal annotations can provide multiple perspectives on a video. We believe that enabling the user to choose the visual cues according to his needs will result in a more efficient browsing experience.

NAVIGATING VIDEO WITH VISUAL CUES

We have developed TAV, a novel video browsing interface that provides the user with a temporal overview of video annotations including visual cues that represent the underlying video content.



Figure 1. TAV User Interface: (1) traditional video player (2) annotation layer (3) annotation filter (4) annotation viewing

Searching Video by Annotations

The TAV user interface (Figure 1) consists of a traditional video player (1) vertically adjacent to our novel browsing interface. Below the video player multiple timelines (2) are displayed. Each timeline refers to a specific type of video annotation, in this case goals or penalties in a hockey game. Each timeline consists of a visual identifier on the left side followed by a scrub bar with annotation icons. Annotations are located according to their temporal position in the video and can be of different types – our example shows icons representing Canada and USA in a hockey game. Based on this, the user cannot only see where goals occurred in the video but also which team made the goal. The user can drag the playhead to advance to a scene of interest.

Viewing Video with Annotations

While the video is playing the video annotations are shown in the viewing table (4). This enables the user to watch annotations simultaneously with the video scene they refer to. Annotations are added to and removed from the viewing table according to the playhead's position on the timeline.

Filtering Annotations

We integrated two filter mechanisms to allow the user to adjust the visual overview according to his needs. A filter toolbar (3) enables the user to determine which types of annotations are displayed on the timelines. For instance, if a user is only interested in goals and penalties referring to Canada, he can deactivate USA annotations by selecting the related filter button. A second filter method can be applied by deactivating a timeline. For instance, if a user is only interested in goals but not in penalties, he can deactivate the second timeline by selecting it on the left side.

Using Annotations to Identify Interesting Scenes

Besides the qualitative use of video annotation as a visual cue for representing the underlying video content, we also provide quantitative navigational aids. A high number of



Figure 2. A thumbnail's size below the playhead indicates the number of video annotations in a scene to identify scenes of interest.

annotations added to a video scene usually indicates that many people found this scene interesting. TAV visualizes the quantitative distribution of video annotations by placing different scaled thumbnails below the video timeline (Figure 2). This enables the user to quickly determine which scenes got the most attention by other viewers and therefore might be of interest. We addressed the limited space problem by providing the user with zooming functionality.

CONCLUSION

We presented a novel video browsing interface called TAV that enables the user to navigate through video by using a visual overview of temporal video annotations. We emphasized the potential of temporal video annotations as a valuable source for navigational cues and demonstrated how simple filter mechanisms can help the user to focus on relevant information. Future work will include the evaluation of our novel user interface concept by conducting a user study to measure retrieval speed of video scenes and to examine if users are distracted by the visual cues. Furthermore, we will elaborate how TAV can be used to add temporal video annotation to a video.

REFERENCES

1. L. Chen, G.C. Chen, C.Z. Xu, J. March, and S. Benford. EmoPlayer: A media player for video clips with affective annotations. *Interacting with Computers*, vol. 20, no. 1, pp. 17-28, 2008
2. M. Costa., N. Correia, N. Guimaraes. Annotations as Multiple Perspectives of Video Content. In *Proceedings of the Tenth ACM International Conference on Multimedia*, pp. 283-286, 2002
3. K. Schoeffmann, M. Lux, M. Taschwer, and L. Boeszoermenyi. Visualization of Video Motion in Context of Video Browsing. In *Proceedings of the IEEE International Conference on Multimedia and Expo*, 2009
4. Viddler, Accessed August 2010, www.viddler.com