Casual Authoring using a Video Navigation History

Matthew Fong, Abir Al Hajri, Gregor Miller and Sidney Fels*

Human Communication Technologies Laboratory, University of British Columbia, Vancouver, BC, Canada

Figure 1: Our prototype history-based interface for casual video authoring. A history of personal video viewing provides a searchable set of intervals (*left, red*) which can be filtered to identify intervals viewed more than once (*right, red*). The video is linearly searchable using the filmstrip (*blue*). Any interval in the history can be added to the playlist (*green*), which can then be exported to a video and shared with friends.[1]

ABSTRACT

We propose the use of a personal video navigation history, which records a user's viewing behaviour, as a basis for casual video editing and sharing. Our novel interaction supports users' navigation of previously-viewed intervals to construct new videos via simple playlists. The intervals in the history can be individually previewed and searched, filtered to identify frequently-viewed sections, and added to a playlist from which they can be refined and re-ordered to create new videos. Interval selection and playlist creation using a history-based interaction is compared to a more conventional filmstrip-based technique. Using our novel interaction participants took at most two-thirds the time taken by the conventional method, and we found users gravitated towards using a history-based mechanism to find previously-viewed intervals compared to a state-of-the-art video interval selection method. Our study concludes that users are comfortable using a video history, and are happy to re-watch interesting parts of video to utilize the history's advantages in an authoring context.

Index Terms: H.5.2. [Information Interfaces and Presentation]: User Interfaces—; H.1.2. [Models and Principles]: User/Machine Systems—

1 INTRODUCTION

Casual video authoring is becoming popular with the rise of video sharing sites, such as YouTube[TM]. However, video authoring continues to be a challenging task for consumers due to a lack of tools for efficient and simple editing [14]. Part of the problem is that

---

*{mfong, abira, gregor, ssfels}@ece.ubc.ca

the 3D spatio-temporal representation of video complicates relatively simple actions such as cropping and selection. Further, video editing often taxes human memory by requiring memorization of a large quantity of digital assets. In particular, finding and selecting interesting parts has poor navigation and search support. The situation is such that recent studies have shown that unless given a significant incentive, users will avoid video editing [20]. We propose that the addition of a video history mechanism into casual video authoring will overcome some of these difficulties.

We investigate the usefulness of video history through a casual video authoring interface that supports users employing their own video navigation history to quickly retrieve previously-viewed intervals and assemble them into shareable movies. Our approach is in contrast to the commercial state-of-the-art, which requires directly linking to a start time (e.g. YouTube[TM]), with no end time given and only one clip linked at a time. Placement of temporal comments can aid navigation (e.g. SoundCloud[TM]) however this requires explicit annotation. Social navigation methods using crowd-sourced viewing patterns [21] can aid the selection of popular clips, although this does not help with home/mobile video or editing video before upload.

Current video navigation methods mostly use the familiar VCR-like controls (play, pause, seek, fast forward, rewind) and sometimes include chapter systems (e.g. DVDs) to skip to specific sections of a movie. Commercial video viewing systems have not progressed much further, despite significant research into new methods. Girgensohn et al. improved video thumbnails, allowing users to directly manipulate the preview frame by continuously moving a cursor along a timeline to find the desired clip [7]. Kimber et al. encouraged users to directly manipulate within-video content along its natural movement path to explore video [13]. These both require users to remember the temporal location of an event within the timeline (or chapters). In an attempt to summarize and show users video summaries, Christel et al. [5] developed video skims that would summarize the most important contents of a video.
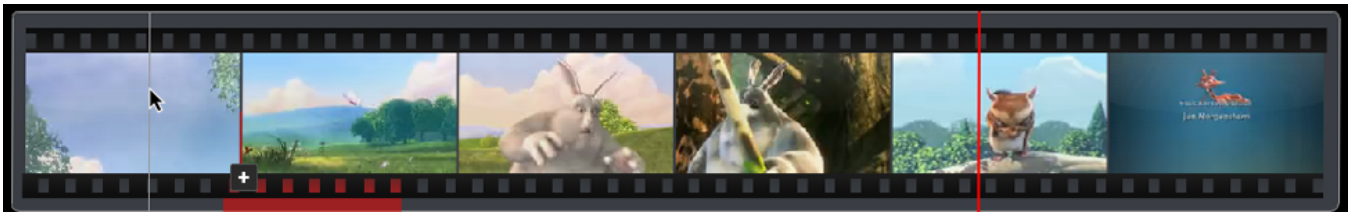
Figure 2: The filmstrip provides easy access to the entire video at once. Each thumbnail shown represents the frame of the video at the temporal location of the start of the thumbnail's interval. The red bar represents the current time being shown in the main player. Moving the cursor across each thumbnail will preview the frame that the cursor position represents. Users can drag and select video intervals, and click the "+" to add them to the playlist.

We believe that the revisitation can be a useful tool for search and selection. Zipf's Law [22] and Pareto's Principle [12] both investigate the idea of natural repetition in human behaviour, under the assumption that people do not invoke more effort than needed to accomplish a task. Alexander et al. [3] use this as the basis for their work on revisitation in documents. It was found that marking user footprints on the document scrollbars yielded better performance in search. An investigation into web histories by Cockburn et al. [6] showed that over half of web page visits were previously-seen web pages. Greenberg et al. [9] investigated the reuse of commands in a command line interface, and built an interface to facilitate repeated usage. Li et al. developed smart bookmarks that can perform macro-like operations on web pages [16]; Bergman et al. created a tutorial generation system which recorded user actions and replayed it as a form of documentation for new users to follow [4]; Grossman et al. developed *Chronicle*, a system to explore workflow histories of graphical content and allow users to indicate items of interest, from which they can observe how to replicate this part of the workflow [10]. We apply a similar methodology in our work, assuming there is an element of repetition in human behaviour and that tools such as online video are supporting this. The use of a history mechanism can keep track of what is watched (and which parts have been watched more than once, or not seen at all); this leads to a new perspective on video editing which can be done directly from the viewing space, without the need for a separate tool. Our novel interface demonstrates the ease of authoring when using user-based viewing history.

Existing work that applies a history methodology to video consumption is sparse, although all demonstrate high utility for navigation, search and summarization. Yu et al. used low-level feature extraction and video interval view counts, to rank and suggest interesting scenes [21]. Mertens et al. visualized users' footprints on the video timeline and allowed users to quickly navigate to the corresponding scenes in the video [17]. Yu et al. and Syeda-Mahmood and Ponceleon used a browsing history to rank particular scenes in a video to generate video summaries [21, 19]. Although these use some form of video history, they do not provide direct access to discrete video intervals watched: the history is used internally to support the system. The first video viewing framework to provide users with direct access to a full history of their personal video navigation was recently introduced by Al Hajri et al. : they demonstrate users can find events in a video more quickly using a history than with traditional controls [1]. Al Hajri et al. later demonstrated an interface for multiple-video history with the same properties i.e. finding previously-seen content was significantly faster [2]. While their prototype demonstrated the utility of a video history for event search, we show that it can be applied successfully in an authoring context to provide a simpler mechanism than existing interactions.

The video history is more effective when users watch segments of video more than once, or use seek and search to avoid viewing portions of the video completely [8, 11, 15, 17, 18, 21]. Evidence of this can also be shown by YouTube and Vimeo's sharing by 'link to start time'. Audience retention statistics (e.g. from YouTube) show that users do not linearly watch video (graphs have peaks and troughs, indicating different view counts across the video). Users do not currently have widespread access to video histories, and we are interested in how users react when given a history which takes advantage of the information gained from re-viewing video. We report on this aspect in the results of our user study.

In the following section we describe the methods for utilizing history to for video interval search and selection, as well as the supporting interface we designed around them; we then outline the study we performed to evaluate its effectiveness against a state-of-the-art selection method, and report on the benefits of using a history as part of a video viewing platform.

2 INTERACTION DESIGN

We envision users will watch videos differently when they have a history collection mechanism so that they may easily save clips to use later. We consider two contexts when this might be done. The first occurs when the the video in question is longer than necessary and only contains a few entertaining intervals worth sharing. This is typical in YouTube videos, for example. Using the history feature, the user watches the video and re-views interesting parts on the fly (knowing it is saved in their history): the user can then quickly share the relevant intervals with other people. Sports video provides our second example context. Here, videos are generally long, with relatively few exciting parts that are typically viewed again frequently (e.g. instant replay, highlights, etc.). Watching sports on a history-enabled viewer, the user can save the exciting parts by re-watching them. They can then add them to a playlist once the game has finished to create a personalized highlights video, and share it with friends.

The major focus of our design is to allow users to reduce video into smaller intervals using video browsing interactions; specifically, seeking (which includes re-watching video). There is one interval for each navigational action a user has taken in the history: seeking and changing videos. As soon as a user seeks within the video, the current video interval ends, and a new one is created. Within the interface, video intervals are represented by seekable thumbnails. By finding the intersections of all video intervals in the history, we can help users find intervals that have been watched more than once.

2.1 Design Guidelines

In designing our interface there were some guidelines that we tried to keep consistent throughout the interface. These guidelines include:

- To represent a video, we use thumbnails of the video itself. Each thumbnail represents an interval of video, with a distinct starting and ending time. Placing the cursor over it forces it to display the frame that corresponds to the cursor position
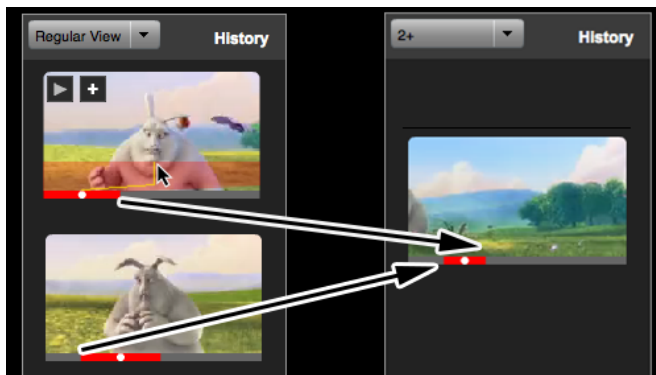
Figure 3: The history is a log of intervals the user views: in this case, the user watched a third of the video (*left, top thumbnail*), rewound via seek to approximately one quarter in, and resumed watching until stopping close to half-way through (*left, bottom thumbnail*). Using the filter (*right*), the interval the user watched twice is extracted, providing the user with a quick and simple way to store and later find intervals of interest.

on the thumbnail. Removing the cursor from the thumbnail reverts the thumbnail's frame to its original position.

- Arranging these thumbnails is also taken into consideration. We have two types of timelines: a video timeline (Figure , *highlighted in blue*), and a user timeline (Figure , *highlighted in red*). Video timelines represent how the video is playing, and is arranged horizontally, from left to right. User timelines represent the actions a user has taken and are shown vertically, from top to bottom. This is employed by the history visualization.

- To represent the entire video, video timelines stretch across the video that they are representing.

2.2 Supporting Interface

To explore how users' navigation behaviour changes in an authoring context when given access to their personal history, we designed an interface with familiar playback controls and a visualization of their history. The interface has four components: a large player to view video; a thumbnail-visualized timeline for in-video navigation (the filmstrip); a navigation history visualization; and a playlist to organize intervals.

2.2.1 Main Player

The main player, shown at the centre of Figure , displays the currently playing video. Interactions with it are limited to clicking to toggle pausing and playing the video.

2.2.2 Filmstrip

The filmstrip, shown in Figure (*highlighted in blue*), is a tool for navigating the entire video. Hovering the cursor over any point shows a preview of the corresponding frame in the video; clicking at this point will seek playback to this position. The initial preview in the thumbnail is the first frame the thumbnail represents in the interval. Intervals can be selected by click-dragging across the filmstrip. A ⊞ appears to let users add the selected interval to a playlist. An example of the filmstrip can be seen in Figure 2. As noted before, the previews are distributed horizontally across the screen, to represent video time. Furthermore, the filmstrip's width is the same as the main player, indicating the full length of the filmstrip is visible to the viewer as the length of the video.

2.2.3 History

The history is a collection of how a user watched a video, and it is a sequential log of all navigational actions taken by users. It is represented by a number of thumbnails which are created as a user is watching a video. There is one thumbnail for each navigational action a user has taken: video seeking and changing videos. These thumbnails represent video intervals, where each interval is the time within the video between actions. As soon as a user seeks within the video, the current video interval ends, and a new one is created. If the video is paused, this new interval represents a single frame and is essentially useless. To combat this, we check for the last interval's length and replace it if the interval is less than one second. These history interval thumbnails allow users to visually see which parts of the video they have watched.

Each thumbnail is initially a small screenshot of the frame of video that is halfway between the start and end time of the interval it represents. On the bottom of the thumbnail is a small timeline that represents the entire video, with the history interval being highlighted in red. Upon placing the cursor over the thumbnail, a popup timeline appears. Moving the cursor across the thumbnail allows users to seek across the video interval, allowing them to check the contents of the thumbnail. Again, in keeping steady with our design guidelines, the entire width of the thumbnail represents the entire length of the interval being represented. Additionally, two new buttons appear on the upper left corner of the thumbnail when the cursor is over the thumbnail. Clicking on the ▶ button will allow the user to play the video interval in the main player, and highlight the corresponding video interval in the filmstrip in blue. The same action can also be performed by clicking and dragging the thumbnail over to the main player. The ⊞ button allows the user to insert the video interval into the playlist. To indicate successful insertion, a ghost thumbnail is animated to fly over from the history to the playlist.

Across the top of the history, there is a drop-down selection box that allows users to filter their history, facilitating users' search for the appropriate video interval. Upon selecting one, the history widget will find the intersections of existing history intervals and use that as a basis for showing the user intervals of video that they have seen more than once. Figure 3 shows a regular history, and the results of filtering for video clips that have been seen two or more times. We would like to refer back to the sports context mentioned earlier. Watching the game, a user would like to re-watch the goals as they happen. Once they do, the history that is created would reflect the particular interval that has been re-watched. After finishing the viewing, it would be useful to be able to quickly find that particular interval to review. By using the filter, it is easy to distinguish parts of the game that were more exciting.

2.2.4 Playlist

The playlist, shown in Figure 4, stores a collection of clips, much like the history, for future viewing. It is a more explicit method for saving video intervals and requires slightly more user interaction. As noted by the other interface elements, inserting clips into the playlist can be accomplished by either selecting an interval in the filmstrip and clicking the ⊞ button, or by adding an interval from the history.

When the interface is started, the playlist's position is held slightly off screen to prevent distraction, as shown in Figure (*left*). It can be easily brought into view by clicking on the arrow on its left side, and subsequently hidden from view by clicking the ⊠ button on the right. The playlist supports saving by clicking the 🆄 button, as well as opening previously made playlists by clicking the 🗁 button. These playlists are saved as XML files for easy transport. The playlist can also be previewed by clicking the ▶ button.

Figure 4: The playlist houses a collection of manually added video clips. A hovering preview is presented for easy viewing, and changes according to the cursor position. These can be adjusted by dragging the edges of each thumbnail, and both previews will change to reflect the edge being dragged. Users can also remove individual clips, and preview the entire playlist.

3 EVALUATION

We ran a comparative user study to: 1) investigate whether using a video navigation history as an interaction tool would make creating clips and adding them to playlists more efficient compared to the state-of-the-art selection method 'filmstrip'; 2) explore any differences on user behaviour when given access to their personal history; and 3) to discover if users have a positive experience with a navigation history, since this is a relatively unknown tool for video.

3.1 Participants

Eighteen volunteers participated in the experiment: thirteen male and five female, ranging in age from 19 to 50. They were monetarily compensated for their time. Each participant worked on the task individually. All were experienced computer users and watched videos at least three to five times a week, but rarely created videos.

3.2 Design and Procedure

Each subject was exposed to two interaction methods, *history* 3 and *filmstrip* 2, and six videos: each video was used once only per subject. For each video the subject created a themed playlist using one of the interactions. We divided the participants into two groups: the first group used the first set of videos with *history* and the second set with the *filmstrip*; reversed for the other group. Each set contained three two-minute videos; a variety of content was used, including sport, news, documentary and comedy.

The evaluation began with a familiarity phase: participants were shown how to use the interface and how the history is created. In the experiment, a trial involved asking the participants to watch a video and re-watch intervals they think fit one of the two themes provided (e.g. "police officers are visible" and "firefighters are visible") for each video. While participants re-watched clips for two themes, once viewing was complete, the task was to collect clips for a single theme. Using two themes ensured the history had distractors (controlling for non-relevant content, and attempting to replicate conditions of a real personal history). Participants were then asked to add clips to the playlist that fit only one theme, using the current interaction technique (either *history* or *filmstrip*); the selection method not being tested was hidden. The participants were asked to complete the task as quickly as possible and the total time taken to add the clips to the playlist was recorded. Timing began when the user clicked the 'Start' button, and stopped when the 'Share' button was clicked. Submitted clips were checked for correctness

Table 1: The mean time taken (with standard deviation, σ) to construct a new video based on the given theme, for each video and method. In all cases the History is significantly faster, with a worst-case of two-thirds the time taken by Filmstrip (comparing means).

Video	Method	Mean (seconds)	σ (seconds)
1	Filmstrip	27.61	9.25
	History	17.04	10.55
2	Filmstrip	23.04	6.92
	History	8.05	7.50
3	Filmstrip	14.55	5.54
	History	8.91	4.32
4	Filmstrip	25.70	7.91
	History	8.08	2.80
5	Filmstrip	18.59	6.89
	History	9.04	6.77
6	Filmstrip	15.24	6.13
	History	9.79	2.69

to ensure a proper evaluation among all participants. Incorrect clips were considered as errors and discarded. Once completed, participants progressed to the next video (and changed method after three video tasks were complete). Upon the completion of the six videos, participants were given two more videos, and extra time to freely experiment with both techniques at the same time (in the same interface). They then answered a questionnaire asking for their reactions to the interface. The total time taken was approximately one hour.

3.3 Results and Discussion

We would first like to note that, as explained in the design, this was a within participant study that was conducted with between elements for the method and videos. Using this design, we would be able to get more informed qualitative data from the participants who would be able to give a more accurate opinion of both techniques. The analysis was done between so that each participant was exposed to each video only once with a single method.

A two-way ANOVA analysis was used to evaluate the performance, and to explore the effect of method and video on the time needed to create a playlist. The analysis showed significant main effects of method: $F(1,96) = 66.69, p < 0.0001$; and of video:

$F(5,96) = 5.86, p < 0.0001$. However, these main effects are qualified by a significant interaction effect - $F(5,96) = 2.37, p = 0.045$. Simple main effects of the video at method show that for every video the time needed to create the playlist was significantly lower using the history. Participants took at most two-thirds the time of *filmstrip* when using *history*, as shown in Table 1. The interaction effect was caused by the different types of videos. Because these were real world videos, chosen to simulate real world scenarios, the location of the themes in the videos were not linearly distributed, causing some video clips to be easier to find than others.

We chose not to include viewing time into our quantitative measurements. Given the video viewing behaviour mentioned earlier in the paper, it does not fit the premise if we included the viewing time, as we do not expect users to find the video, find and rewatch the interval and share it; we expect them to find the interval in their history and share it (or combine it with others in a playlist). Our experiment uses a controlled laboratory setting to investigate this use case and the utility of a history for this purpose. Participants were required to view the videos with the same behaviour with both methods, so the viewing time is similar for both methods.

The table also shows that the video type (e.g. sports footage looks very similar) and the different strategies employed by users (described below) led to a slightly higher variation in times.

Of the six videos, we found that the history worked better for some videos than others, and the same effect can be observed in the filmstrip. These traits can generally be attributed to the content of the videos, which is why the analysis is separated for each video.

1. The first video consisted of a short hockey video of a power-play. The themes for the video were two goals, and two intervals where the defensive team had possession of the hockey puck. The participants were asked to find the two goals. In this video, it was difficult to see what was going on within the thumbnails. This was particularly difficult with the filmstrip because many participants forgot when the first goal happened, and were forced to look through the entire video in the filmstrip, whereas the history reduced the search space required, down to four short clips.

2. The second video consisted of six hockey players, three wearing blue jerseys, and three wearing white jerseys, skating around a rink and competing for the fastest time. The theme in this video is the hockey players turning the last corner. The participants were asked to find the white jerseys to insert into the playlist. Again, like the previous video, participants had to search through the entire filmstrip to find the correct clips, while the history only required participants to look through six different clips, all of which were easily distinguishable by the screenshot within the thumbnail. Furthermore, because one of the players skates in the opposite direction of the other five, using the filmstrip became slightly more difficult because the participant had to distinguish the different turns within the small thumbnail.

3. The third video was a short clip of a comedy sketch where a little girl runs around, makes smart remarks, and pulls pranks on various people. The themes to this video were the reactions to her remarks, as well as the result of the two pranks she pulled. The participants were asked to find the clips with the two pranks. This video had a slight advantage given to the filmstrip because the scene of one of the pranks was visible in the filmstrip, and the second prank was right at the end of the video, making both mental and physical retrieval relatively easy. The data reflects this condition as this was the fastest video for filmstrip.

4. The fourth video was a comedic instructional video of someone taking apart a camera. Along the way, the instructor

Table 2: The aggregated results of our questionnaire (Mean on a 1-7 Likert Scale, with standard deviation σ). Participants found the interface with the history to be useful for the creation of new videos, and their overall reaction to our mechanism was highly positive.

Question	Mean	σ
Overall usefulness	5.77	1.06
Overall ease of use	5.61	1.29
Overall reaction	5.47	1.19
History is useful	6.11	1.23
Creating usable history is easy	5.56	1.29
History filtering is useful	5.89	1.07
History filtering is intuitive	5.28	1.56
Finding history video clips is easy	5.44	1.89
Finding filmstrip video clips is easy	4.50	1.89
Inserting history video clips is easy	6.44	0.70
Inserting filmstrip video clips is easy	5.61	1.50

would make mistakes that were very obvious, and he would disconnect things within the camera. The chosen theme for this video were the three mistakes made throughout the video. Like the second video, the clips were fairly spread out along the video, however they were somewhat hard to see within the filmstrip, which made it slightly more difficult using that method. Additionally, some participants forgot the location of the clips and were forced to look through the entire movie, like the first video.

5. The fifth video was a news interval on an Olympic hopeful wanting to enter the snowshoeing competition. The themes of the video were the six talking heads within the video. Of those, the participant was asked to find two specific ones. This ended up being very easy for history, as it was not visually tasking to find clips of a specific person. Again however, one of the clips appeared in the filmstrip, and the second clip was right after.

6. The last video was of a riot, and the themes of the video were of shots of fire fighters and police officers. The clips to be selected were the fire fighters. These clips were mostly within the middle 30 seconds of the video. For the filmstrip, some participants started searching from the beginning of the video, as they did not remember when the first clip appeared. They were, however, very close together within the video and allowed for quick selection once they found the first clip. They were visually distinctive from the police clips, and were also easy to find within the history.

Errors present in the experiment occurred in the selection of incorrect clips due to misinterpretation of some more complex themes. These errors occurred in both methods and were counted as mistrials and we removed them from the results. In total, there were four errors, with three made using filmstrip on videos 1 and 6, and one mistake made using history on video 2.

The questionnaire results shown in Table 2 demonstrate the positive reaction of participants to our history-based authoring interface. The overall scores for the history method were all above 5 (on a 1-7 Likert Scale). The filmstrip method scored well but participants did not find it as easy to find video clips; this is supported by the quantitative data. We performed a Wilcoxon Signed rank test on the last four questions in Table 2 but found no significant results.

Observation of users in the free play section of the study showed a strongly positive reaction to using the history. 88% of the participants changed their viewing behaviour with a history creation model in mind, and relied on this behaviour to select and add intervals to their playlists. Eleven participants deliberately watched

events twice to insert them in their history, and used the view count filter as intended. Three participants created a playlist by adding intervals from the full history and refining them in the playlist. Two participants applied a combination of *history* and *filmstrip*: they used the *history* for as much as possible, then used the *filmstrip* to find additional clips they did not re-watch. Finally, two participants were unreceptive to creating and using the *history*, and exclusively used the *filmstrip* to create the playlist. They stated that the viewing pattern required by the interface did not fit with their current viewing behaviour, and that changing their behaviour would disrupt their video viewing experience.

While these participants did not use the history, it still recorded the video as they watched: In this case, the history contained a single item for the entire video. If we extend the system to multiple videos, it would be easy to see which videos they liked, if the video was viewed more than once, and when they watched them.

While only two participants used both techniques to create their own video in the free play, it is worth noting the history works well in conjunction with the filmstrip. Enhanced integration between the two, such as placing markers in the filmstrip indicating history intervals, may further improve performance.

It may also be interesting to note that this is a fully functioning system, and the first and second videos were actually created using the interface, being pulled from a video that was two hours long. It took approximately one minute per video to create using the history.

4 CONCLUSIONS AND FUTURE WORK

Our work defines a new way to casually string video intervals together by incorporating a personal video navigation history. We demonstrated that users can simply watch videos naturally, seek or search, re-view intervals of interest and use the history to quickly share or combine them later. Our comparative user study quantitatively proved that using the history lets users more quickly find intervals and create playlists, and qualitatively demonstrated the usefulness and usability of a history-based interaction. Users were generally happy to change their video viewing behaviour in our authoring context when given an accessible history of their video watching experience. The presented method of authoring can be extended to provide access to multiple videos, which will be useful for editing collections of home movies. This also brings up the problem of scalability, since as video collections grow the history will grow as well. We also plan to observe users current viewing habits on video sites such as YouTube to establish which content types lead to specific classes of viewing behaviour (such as re-watching). This will inform on future designs of the authoring tool and the history based on the current content being viewed.

REFERENCES

[1] A. Al-Hajri, G. Miller, S. Fels, and M. Fong. Video navigation with a personal viewing history. In *Human-Computer Interaction – INTERACT 2013*, volume 8119 of *LNCS*, pages 352–369. Springer, 2013.

[2] A. Al-Hajri, G. Miller, M. Fong, and S. Fels. Visualization of personal history for video navigation. In *Proceedings of the ACM CHI Conference on Human Factors on Computing Systems*, CHI'14, New York City, New York, U.S.A., April 2014. ACM.

[3] J. Alexander, A. Cockburn, S. Fitchett, C. Gutwin, and S. Greenberg. Revisiting read wear: Analysis, design, and evaluation of a footprints scrollbar. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 1665–1674, New York, NY, USA, 2009. ACM.

[4] L. Bergman, V. Castelli, T. Lau, and D. Oblinger. Docwizards: a system for authoring follow-me documentation wizards. In *Proceedings of the Symposium on User Interface Software and Technology*, pages 191–200. ACM, 2005.

[5] M. G. Christel, M. A. Smith, C. R. Taylor, and D. B. Winkler. Evolving video skims into useful multimedia abstractions. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '98, pages 171–178, New York, NY, USA, 1998. ACM Press/Addison-Wesley Publishing Co.

[6] A. Cockburn and B. McKenzie. What do web users do? an empirical analysis of web use. *International Journal of Human-Computer Studies*, 54:903–922, 2002.

[7] A. Girgensohn, J. Boreczky, and L. Wilcox. Keyframe-based user interfaces for digital video. *Computer*, 34(9):61–67, Sept. 2001.

[8] C. Gkonela and K. Chorianopoulos. Videoskip: event detection in social web videos with an implicit user heuristic. *Multimedia Tools and Applications*, pages 1–14, 2012. 10.1007/s11042-012-1016-1.

[9] S. Greenberg and I. H. Witten. Supporting command reuse: Empirical foundations and principles. *International Journal of Man-Machine Studies*, 39:353–390, 1993.

[10] T. Grossman, J. Matejka, and G. Fitzmaurice. Chronicle: capture, exploration, and playback of document workflow histories. In *Proceedings of the Symposium on User Interface Software and Technology*, pages 143–152. ACM, 2010.

[11] K.-W. Hwang, D. Applegate, A. Archer, V. Gopalakrishnan, S. Lee, V. Misra, K. Ramakrishnan, and D. Swayne. Leveraging video viewing patterns for optimal content placement. In R. Bestak, L. Kencl, L. Li, J. Widmer, and H. Yin, editors, *NETWORKING 2012*, volume 7290 of *Lecture Notes in Computer Science*, pages 44–58. Springer Berlin Heidelberg, 2012.

[12] J. Juran and A. Godfrey. *Juran's Quality Handbook*. Juran's quality handbook, 5e. McGraw Hill, 1999.

[13] D. Kimber, T. Dunnigan, A. Girgensohn, F. Shipman, T. Turner, and T. Yang. Trailblazing: Video playback control by direct object manipulation. In *International Conference on Multimedia and Expo*, pages 1015–1018. IEEE, July 2007.

[14] D. Kirk, A. Sellen, R. Harper, and K. Wood. Understanding videowork. In *Proceedings of the Conference on Human Factors in Computing Systems*, CHI, pages 61–70. ACM, 2007.

[15] I. Leftheriotis, C. Gkonela, and K. Chorianopoulos. Efficient video indexing on the web: A system that leverages user interactions with a video player. In *Proceedings of the 2nd International Conference on User-Centric Media (UCMEDIA)*, 2012.

[16] I. Li, J. Nichols, T. Lau, C. Drews, and A. Cypher. Here's what i did: sharing and reusing web activity with actionshot. In *Proceedings of the Conference on Human Factors in Computing Systems*, pages 723–732. ACM, 2010.

[17] R. Mertens, R. Farzan, and P. Brusilovsky. Social navigation in web lectures. In *Proc. of the Conference on Hypertext and Hypermedia*, pages 41–44. ACM, 2006.

[18] D. A. Shamma, R. Shaw, P. L. Shafton, and Y. Liu. Watch what i watch: using community activity to understand content. In *Proceedings of the international workshop on Workshop on multimedia information retrieval*, MIR '07, pages 275–284, New York, NY, USA, 2007. ACM.

[19] T. Syeda-Mahmood and D. Ponceleon. Learning video browsing behavior and its application in the generation of video previews. In *Proceedings of the International Conference on Multimedia*, pages 119–128. ACM, 2001.

[20] S. Vihavainen, S. Mate, L. Seppälä, F. Cricri, and I. D. Curcio. We want more: human-computer collaboration in mobile social video remixing of music concerts. In *Proceedings of the Conference on Human Factors in Computing Systems*, pages 287–296. ACM, 2011.

[21] B. Yu, W.-Y. Ma, K. Nahrstedt, and H.-J. Zhang. Video summarization based on user log enhanced link analysis. In *Proceedings of the 11th International Conference on Multimedia*, pages 382–391. ACM, 2003.

[22] G. Zipf. Human behaviour and the principle of least-effort. Addison-Wesley, Cambridge, MA, 1949.