Video Navigation with a Personal Viewing History

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

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

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

Abstract. We describe a new video interface based on a recorded personal navigation history which provides simple mechanisms to quickly find and watch previously viewed intervals, highlight segments of video the user found interesting and support other video tasks such as crowd-sourced video popularity measures and consumer-level video editing. Our novel history interface lets users find previously viewed intervals more quickly and provides a more enjoyable video navigation experience, as demonstrated by the study we performed. The user study tasked participants with viewing a pre-defined history of a subset of the video and answering questions about the video content: 83.9% of questions (average) were answered correctly using the personal navigation history, while 65.5% were answered using the state-of-art method; they took significantly less time to answer a question using our method. The full video navigation interface received an 82% average QUIS rating. The results show that our history interface can be an effective part of video players and browsers.

Keywords: Video Navigation; Navigation History; Video Summarization;

1 Introduction

The dramatic increase in the quantity of video now available, either in the online form of services such as YouTube[TM], Vimeo[TM] or the more personal form of home video, provides new challenges such as finding specific videos (or intervals within video), authoring new videos from existing content (e.g. video summarization or home movie editing), sharing video playlists (or even intervals) of online video and annotating dynamic content. The increased volume and shrinking average duration of online video has led to new use cases: users are generally free to skip ahead to find intervals which interest them; users may re-watch part of the video they enjoyed; users may employ temporal video links to provide instant access to a specific time within a video, skipping irrelevant sections; playlists and temporal links provide opportunities for customization and sharing of video consumption among users. In short, users can and want to view personally interesting intervals easily and without viewing the entire video. These use cases are supported within current interfaces, however there are no existing mechanisms to support new use cases which arise as a natural consequence, such as navigation of previously viewed intervals. Additionally, users often re-watch segments of video as part of the contemporary browsing and navigation experience

[24, 25, 28], and may wish to share these intervals with others. A historical record of video viewing would allow users to quickly find popular intervals, comment on and share them.
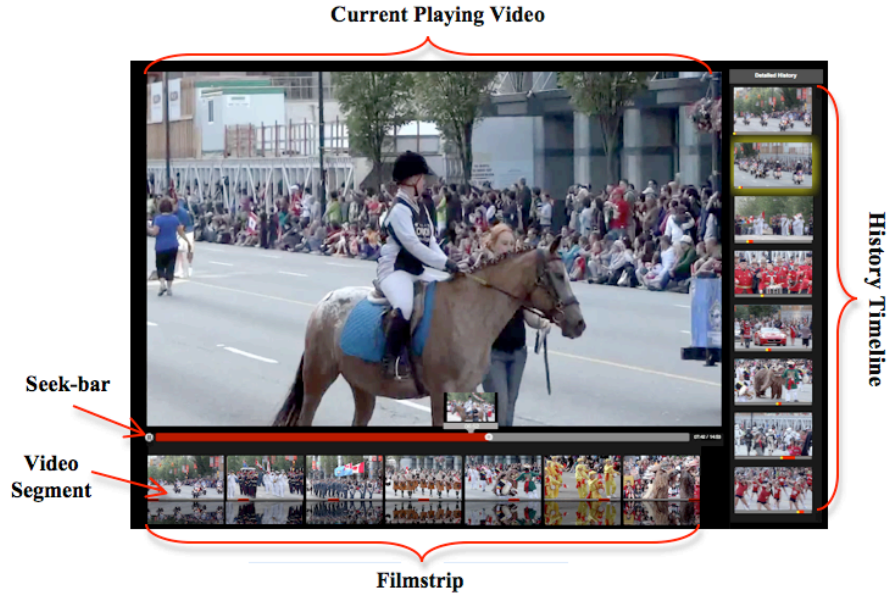

Figure 1. Our video navigation interface: the majority of space is devoted to the currently playing video (top left) with a seek-bar preview; below is a horizontal array of Video Segments arranged by video-time (the Filmstrip), and a vertical array of Video Segments to the right (the History Timeline) ordered top-down by user-time i.e. the order the intervals were viewed. These components described in more details in section 4.

We introduce our novel video interface, which captures the user's viewing history (analogous to a web browser history) and allows them to reuse it as a navigation tool, through a user-time visualization with supporting interactions. The history supports actions such as faster access to previously viewed intervals, visualization of popular intervals within the video based on a user's temporal view count, simpler authoring via selection of viewed intervals and provides the user with the ability to create a personalized map of the video (e.g. deciding on their own separation of the video into chapters, or highlighting interesting intervals). This type of history interface is not currently supported in commercial products, thus, we are studying interfaces for managing personal viewing history to evaluate effectiveness for inclusion in future video interfaces.

Our research focuses on the utility of video history, examining use cases and establishing which are significantly affected given a user-based timeline of video viewing. The interface we designed (shown in Figure 1) includes a filmstrip visualization below the main video player representing the entire video, and a history to the right of the player with a sequence of intervals ordered top-down by user-time. We discovered through pilot studies that the filmstrip and history are equivalent in navigation speed

to find a previously viewed event in the case where the entire video is viewed, although the history interface was preferred since users were able to customize which intervals were visualized. Our study shows that when a subset of the video is viewed, the history significantly outperforms the state-of-art, and 75% of participants preferred the history in this use case. The study found that our history-based authoring interface was useful, easy to use and enjoyable, and participants were keen to see this functionality in YouTube.

We begin with a discussion of the related work, followed by a description of the video navigation interface (including all widgets and interactions used). The evaluation of the use of a history and the interfaces we designed to support it is presented, looking at users' impressions of using a personal video history and the power of using history for fast video navigation.

2   Motivation

Consider the following scenario with Robert and Elisa:

> *Robert received a YouTube link from his friend Tom about his family vacation to Rome last summer. Robert and his wife Elisa watched the video along with other videos about Italy to enjoy a sense of Tom's holiday. Robert enjoyed some segments of the videos and wanted to share them with his kids when they came home. When they arrive, he browses his history view and jumps to the good bits to show them. After seeing how nicely the Italy videos looked, Elisa suggests creating a video of their last vacation to Vancouver using various videos and photos they captured. Robert had watched these videos many times and he was excited about including particular events from the 12 hours of video he had obsessively taken during their 4-day holiday. He goes to his history view and grabs the shots of the parade and fire works of Canada day, the suspension bridge and swimming at Lynn Canyon Park, along with another 10 segments that he had watched at various times with Elisa and the kids before. He then drags them into the authoring window and creates a new video that he shares with his friends. .*

This scenario emphasizes the major contributions of our work: (1) a novel video navigation interface for video viewing that captures and visualizes a user's personal history; (2) an experiment that confirms using personal video navigation history is faster in finding previously seen events within videos in contrast to the state-of-art method, filmstrip; (3) a simple video mash-up tool that can be used to summarize videos by using intervals from history or simply watching the corresponding intervals, in comparison to other video editing software where users need to add a video clip, watch it, and then trim it.

3   Related Work

Retrieving and reusing information from the past is a common activity for users [6]. In the domain of web browsers, researchers have introduced and developed different tools that keep records of user's browsing experiences for later use. However, keeping

and reusing navigation experiences in a video space has not received similar attention. Web browsing history is useful for various domains, thus, we expect our research on effective mechanisms for video watching history management will also benefit various fields where video plays an increasingly important role e.g. education, entertainment, etc. History of user actions has been widely investigated for multiple purposes in different domains including web [2, 19, 15], workflow [12, 21], tutorial generation [1] and information visualizations [14, 13]. These are based on mostly static media and so do not directly apply to video: unlike static media, two timelines exist when capturing user history: video timeline and user (or history) timeline. Recording a user's actions, therefore, requires more sophisticated representations, visualization layouts and interaction techniques. In our interface to eliminate the confusion between different timelines, we used a vertical layout to indicate a user's timeline while a horizontal layout indicates a video timeline consistent with usual video players.

The relatively high interest in the social web, sharing and the use of videos online has motivated researchers to investigate the use of video navigation history. Users unintentionally leave footprints during their video-browsing process, which can add value to the content. Yu et al. [28] and Syeda et al. [25] investigated how users' video browsing footprints can be utilized in video summarization and to facilitate future viewers. They used video browsing footprints to rank different scenes within a video, which can later be used to extract clips when generating video summaries or previews. However, these systems [25, 28] have not given the users any control over their personal history during the viewing process (i.e. users cannot access their history). They recorded the users' actions, which were later used by the researchers to generate previews or evaluate the ranking of interesting parts within videos. In our interface we offer the users the ability to utilize their personal history for navigation and authoring as well as summarization.

Video navigation history can play an important role in user-based information retrieval from videos on the web. Shamma et al. [24] and Yew et al. [27] have proposed a shift from content-based techniques to user-based analysis because it provides a more promising basis for indexing media content in ways that satisfy user needs. Leftheriotis et al. [18] proposed a video indexing tool on the web based on user interactions to rank different videos. They generated a thumbnail to represent each video based on the frame that has been mostly viewed by users. Gkonela et al. [10] also indicated that this simple user heuristic while navigating videos can be used as effectively to detect video-events as when content-based techniques are applied. In our interface, we applied a flexible thumbnail visualization of each history segment: initially, each segment is visualized using a thumbnail that indicate the beginning of that segment which then changes while the user navigates this segment.

Navigating a video space or even long videos can be demanding and time consuming. There have been many interaction techniques proposed in the literature to address this problem in order to quickly navigate and search video content; here, we will only mention a few. Typical navigation tools used in most video systems are the VCR-like controls (play, pause, seek, fast forward, and rewind). DVD systems provided users with a higher level of video access using a random access to video segments based on scene indices and chapter boundaries via a table of contents index. Some researchers

[26, 23, 5, 8] allowed video navigation using their abstraction thumbnails or keyframes (e.g. filmstrip), where clicking on a thumbnail directly positions the video at a particular point in time corresponding to that thumbnail. Others [3, 9] proposed new playback mechanisms to help users rapidly skim through uninteresting content while slowing down to watch interesting parts of the video. Since the video content changes as time passes, the separation between the video content and the controls used to manipulate the video introduces a target acquisition problem. Thus, some researchers have focused in introducing new techniques to manipulate the video timeline rather than using the traditional VCR controls. They allow users to navigate the video by directly manipulating the video content along its natural movement path rather than using the timeline slider [17, 16, 7, 11]. Pongnumkul et al. [22] developed a new navigation tool that uses a map-based storyboard where different locations of video segments are directly visualized on the map.

Video navigation history provides a potential interaction technique that can be used to facilitate fast navigation and event search. According to Shamma et al. [24], the more affective a scene is, the more the corresponding interval of the video is viewed or consumed. Thus, the cumulative seamless user interaction history could be leveraged for the benefit of future viewers (i.e. social navigation). For example in YouTube, when enough user data is available, user behavior will exhibit similar patterns which can then be used by new viewers to quickly navigate to the interesting part of that video. Yu et al. [28] used low-level feature extraction along with user's footprints (i.e. view count) to rank each scene and offer users other scenes that have some correlation with what they are watching. This provides users with a quick navigation method to similar scenes. Mertens et al. [20] visualized users' footprints on the video timeline, which allowed users to quickly navigate to the corresponding scenes in the video. In our interface, we used VCR like controls along with thumbnails of the previously viewed video segments that allow simple navigation to these intervals. Our study revealed that using this method speeds up searching for previously seen events within the video.

4   Video Navigation Interface

Given the volume of available video data, methods to effectively navigate this space are needed to provide users with more enjoyable video viewing experiences. The principle goal of our interface is to provide efficient access to previously viewed video: in particular, intervals within video. We accomplish this by recording a continuous video history for the user as they consume video content. This paper considers only the history of a single video as a starting point, and demonstrates that further research is warranted for multiple-video history based on our success.

Internally, the user's video watching history is represented by a very simple data structure organized by linear user time. Whenever a video begins playback the video ID and current user time are recorded; these are noted again when the video stops playing, to produce interval start and end timestamps. An accumulated view count is

maintained for every instant of time in the video: the structure stores intervals of view count derived from the intersection of all viewed intervals.

We designed our interface in a way that allows users to play, view and navigate videos similar to any video player as we envision our history management to be used to augment it. Since our interface records the user's navigation history, we propose new components that offer the users the flexibility of using this history for fast navigation, summarization (or authoring) and sharing. Thus, our interface is based on five separate components: a Player which provides the user with a familiar video setting, a Filmstrip which provides a navigation tool for the entire video content, a History Timeline which visualizes the personal viewing history, a Video Mash-up and a Preview Window which help users creating and previewing their own edit list for video summaries or previews. The main interface components are: the Player and the History Timeline.

Figure 2. The video player component contains play/pause button, video seek-bar, and a pop-up thumbnail shown when the pointer is over the time bar.

Player. A familiar video player (similar to YouTube, QuickTime, etc.), shown in Figure 2, contains: play/pause, applied via a button (bottom left) or by clicking directly on the video; seeking via playhead (white circle on timeline) or mouse click on the timeline (red/grey bar); a frame preview triggered by the mouse over the timeline (including when seeking interactively i.e. with mouse button held down over time slider); and a current playhead time label (white text, bottom right). This component is used for direct control of the selected video. To watch a specific video from disk, users click on the "Open" button and choose their video file.

Figure 3. The Filmstrip component visualizes the entire video into *n* equal length segments.

Filmstrip. The Filmstrip component, shown in Figure 3, is the state-of-art video navigation tool. It provides a visualization aid to different parts of a video for faster navigation and supports access to the entire video content. The Filmstrip simply consists of a fixed number of Video Segments (*n*) from the playing video. The entire

video is divided into *n* equal length intervals, where each interval is represented by a small seek-able Video Segment described below. These intervals are created systematically based on the length of the video and the number of segments to be visualized.

User Viewing History Timeline (History Timeline). This is the central component for visualizing the personal viewing history of the user. We use a filmstrip metaphor for visualization: the History Timeline consists of thumbnails (similar to interactive storyboards) as shown in Figure 1. These thumbnails represent Video Segments that the user watched. We represent this as a Video Segment component as discussed in the next section. Every time the user seeks to a new temporal location, a new Video Segment is added to the end of the user's history list and represents the interval a user watches in the main video player. The History Timeline is a scrollable box, to provide navigation of the entire viewing history and allows the user to find a specific interval.

The video history is similar conceptually to webpage bookmarks that are used to refer back to visited pages, however, it is more complicated as each Video Segment in the history view is a video interval and also includes a representation of the time the user watched it. Thus representing both of these quantities requires additional mechanisms for visualization and interaction. Moreover, there may be confusion with the currently viewed timeline and with the history views since there exist two timelines: when/what the user watched and the video's timeline. We address this by visualizing the history timeline vertically, while the video timeline remains horizontal, consistent with user's mental model of video. The history is unique in that it records in piecewise linear user-time, not video time (we only display the history of intervals watched and do not represent the complete video from where it came from in the History view.). Each Video Segment is a small interactive video widget described next.

Video Segment. Each interval a user watches is represented as a small video widget that contains the viewed interval only, visualized by the starting frame of that interval. The location of the interval within the entire video is visualized as a red portion underneath the widget to help users spatially contextualize the temporal location of intervals within the complete video where it came from. The video widget, shown in Figure 1, is seek-able and playable to allow users to easily search within the intervals and minimize the time needed to search for a previously viewed scene. On mouse over, a play/pause overlay is displayed over the widget. Clicking on the overlay plays the corresponding video interval within this small widget only. (Note, while technically, this is a rewatching of video, we do not add this activity in the History View.) Moving the cursor over the bottom third portion of the widget pops up a zoom-in visualization of the interval timeline, which can be used to seek within the video interval using mouse motion as a cue. The seeking point is visualized by a yellow line and the thumbnail updates to reflect the current seek position. To play a Video Segment in the main video player, users can drag the entire widget (or the seek location in the widget) to the main video player: dragging the entire widget plays the

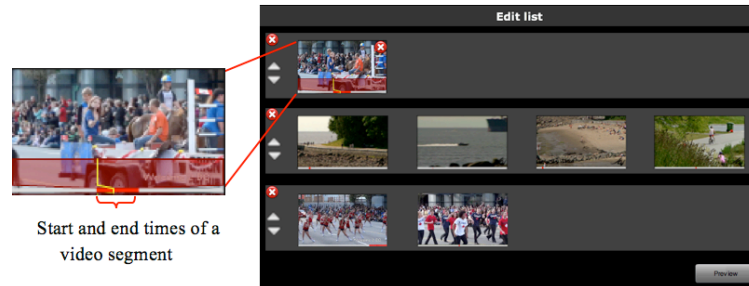corresponding interval from the beginning, while dragging the seek bar location plays from that specific time.

Figure 4. The Video Mash-up component illustrates a user edit list which consists of a user defined Video Segments with the options of delete ✖, reorder ▮, play (on click), navigate (on cursor move shown as a yellow line), and modify segment's start and end times.

Video Mash-up. The video mash-up is used for short video composition such as summaries, previews, or trailers (analogous to video editing software e.g. iMovie, Movie Maker, etc). We included this component in our prototype to evaluate one of the use cases we idenitified that takes advantage of potential strength of having personal viewing history. The video mash-up, shown in Figure 4, represents a video edit list that shows the videos, and intervals within each video, that a user may combine into a summary. In summary creation mode, when the user subsequently plays a video (either through opening a new video, a segment from history or seeking within the current video) a new record is added to the edit list. If the video already exists in the edit list, a new interval is added (unless there is an existing identical interval). Otherwise if the video does not exist, a new video record with the interval is added to the edit list. Thus, adding segments or intervals to the edit list is simply done by watching these intervals, in comparison to other video editing software where users need to add a video clip, watch it, and then trim it. Using this component makes video editing simpler for novice users as we discovered upon evaluation.

The video mash-up component allows the user to preview a specific interval by using the play button for that interval. Users can also modify their edit list by directly updating interval start or end times, delete an interval, delete an entire video record, or reorder video records. Previewing the entire summary is provided by clicking on the "Preview" button.

Summary Preview. After creating a video mash-up, the user can preview it using the summary preview mode, where a new video player window pops up, and the rest of the interface is disabled and darkened. The summary preview takes the user's edit list as input and plays the intervals in order using this content. It allows users to replay, modify (back to the main interface to modify the edit list), or export it if they wish to save it - the edit list is exported as an XML file that can be shared or modified later.

Personal Navigation Footprints. The recorded video history offers additional benefits, such as within-video history and crowd-sourced user histories (neither of these were included in our study). The personal history or footprints offers users a visualization of the view count of every point of time in the video: the blue bar in Figure 5 becomes brighter for the most-viewed intervals, and each one can be selected to play. It provides the user with information on the most-viewed intervals of each video (a temporally tracked frequency). As mentioned previously, a view count is maintained for all points of time for each viewed video. As the user views a video the view count will dynamically change, thus the count is continuously updated and the information propagated backwards through the history. The view counts are then summed and normalized within each video.

Figure 5. A Video Timeline is combined with navigation footprints visualization, where the more an interval within a video is watched the brighter that region becomes. Blue: a single user history (Personal Navigation Footprints), Red: a combined multiple users histories (Crowd-Sourced Navigation Footprints).

Crowd-Sourced Navigation Footprints. The personal history from multiple users are combined to form the most popular intervals for a video (i.e. social navigation similar to these are shown in the red bar in Figure 5. This provides users who are watching a video for the first time with a mechanism to watch what others have found to be most interesting. This is an efficient navigation technique, predicated on trusting the crowd to be "correct" on interesting video content.

5   Evaluation

An extensive set of pilot studies were performed (12 participants total) to investigate a user's preferred scenario for history, and to triangulate the use cases in which a history is beneficial for video viewing. Throughout this triangulation process it became more apparent to the authors how complex and intricate the video-viewing task can be. For tasks such as seeking to a specific time or finding a particular event from a previously watched video, using a history was found to be as good as using a filmstrip methodology. Under more modern viewing patterns (e.g. nonlinear viewing behavior) where users view only parts of the video (e.g. trailers, summaries, playlists or direct temporal links) using a history to find events and seek to them was found to be more efficient.

Using this result from the pilot studies, a full comparative user study was performed to evaluate the design and performance of our interface and to demonstrate the utility of personal video navigation histories. We developed an evaluation protocol which satisfies the use cases previously defined (without biasing our design or the control) and provides the user with sufficient viewing history while keeping the ex-

periment relatively short. For fair comparison we ensured our interface mimics currently adopted approaches with logical extensions.

Using our protocol, we investigated whether visualizing and using a video navigation history would make searching for previously seen events more efficient. We conducted a comparative between-subjects user study, comparing the performance of users employing a personal navigation history against those with the state-of-the-art navigation method (Filmstrip) to find events within a previously seen video (i.e. finding answers to questions similar to [28, 10]).

Both methods have similar layouts and functionality, differing on the process of segment creation described in section 4. In Filmstrip the intervals are created systematically, while in the History Timeline segments are constructed based on a personal navigation history; each participant tried both methods, on different videos. In Filmstrip tasks the participants were presented with the main Video Player and a horizontal Filmstrip components only. They only used these components and their features to find the answer to each question. In History Timeline tasks the participants were presented with the main Video Player and a vertical History Timeline components.

Our experiment was divided into two phases: phase 1 was conducted to compare the two methods (Filmstrip and History), and phase 2 was performed to qualitatively analyze the entire proposed interface, described in section 4. In phase 1, participants experimented 3 components (Video Player, Filmstrip, and History Timeline) of the interface described in section 4, while in the second phase they were exposed to all components and features.

5.1 Apparatus

The experiment application was developed in Flash CS4 and ActionScript 3.0. The experiment ran on an Intel dual-processor dual-core 3 GHz Mac Pro desktop with 8GB RAM and equipped with a 24" Dell LCD monitor with a resolution of 1920 x 1200 pixels at a refresh rate 0f 60Hz. A Microsoft optical wheel mouse was used as the input pointing device with default settings and the Adobe Air environment was set to 1920 x 900 pixels while running the Flash program.

5.2 Participants

Twelve paid university students (different from those in the pilot studies), 6 female and 6 male, participated in the experiment. Participants ranged in age from 20 to 30, all were experienced computer users and have either normal or corrected-to-normal vision. Participants reported watching videos either on a daily basis (10 participants) or 3-5 times a week (2 participants). Each participant worked on the task individually.

This project targets heavy video users, and according to Purcel[1], video is a growing trend especially amongst those aged 18 to 29. Thus, students are a primary target audience as they are representative of the current/near-future heavy users of video interfaces. We used students across the University to capture this demographic.

---

1 http://www.pewinternet.org/ /media//Files/Reports/2010/PIP-The-State-of-Online-Video.

5.3 Design

To evaluate the efficiency of using navigation history as an aid for efficient search within a previously seen video, we gave participants a pre-defined history, which they had to watch and subsequently answer questions based on the content. We followed this procedure rather than allowing participants to create their own history: this decision was made based on the pilot studies where participants found creating a history based on an unknown list of questions was confusing. Some participants said, "What interests me might not be what you are looking for". Thus, they tried to create many history segments at each point they thought there was potential for a question. The result was a long list of short segments which required significant scrolling when searching. In short, the uncertainty of what should be included on the list led to a large number of segments and created some confusion that affected the performance while using the method.

In order to tackle this problem, the participants could be allowed to create their own history based on their interest after which they can be asked questions that exist within these segments. However, the variation between participants interest would make the comparison difficult. Thus, instead of asking participants to create their own history we decided to give them all the same history from which to answer the list of questions (analogous to watching a playlist created by a friend or a video summary). All participants watched the same clips from the video and were asked questions from these clips only. Thus, we ensured they had experienced the same clips and eliminated the personalization factor within the history. Our method was compared with the state-of-the-art method (Filmstrip). The History Timeline contained the segments of the predefined history only while Filmstrip contained seven evenly divided intervals over the entire video.

In this experiment, we used five different short videos (V1-5). V1 was used to explain the interface's components and features and demonstrate how to use them. The other four videos were used for the actual experiment where each participant used a single method for each video. Each participant experienced each method in a different video to eliminate the learning effect. The number of questions per video were: V1: 20; V2: 38; V3: 25; V4: 38; and V5: 32.

The participants were divided into two groups, A and B. Both groups experienced the videos in the same sequence but with different method sequences. They both started the familiarity phase where they tried a hybrid of the two methods to answer example questions and to become familiar with both methods. Group A was given Filmstrip for V2 and V4 and History Timeline for V3 and V5, while group B used the opposite. Thus, the results were compared between the two groups for each video.

To quantitatively compare the results of the two methods, we measured five variables for each video: (1) The percentage of questions answered for each video within a specified time range. The total time given is 15 seconds multiplied by the total number of questions (chosen based on the pilot studies); (2) The time needed to answer each question - this is the time measured between two subsequent correct submission clicks; (3) The number of wrong submissions (errors) for each question, which is measured by counting the number of submission clicks that did not result in

a correct answer; (4) The number of seeks within a video for each question; and (5) The number of previews performed in a video for each question, which is measured by counting the number of seeks and playbacks within the small Video Segments. These counts were tracked to analyze the user's behaviour and to investigate the difference between the two methods.

The participants were also asked to fill out questionnaires and give their comments on the entire interface, and suggestions for improvement. This was conducted to qualitatively compare the two methods and to provide an indication of the importance of correcting each particular aspect of the proposed interface. Participants rated different interface features and tools using a 7-point Likert scale. The experiment lasted approximately two hours per participant.
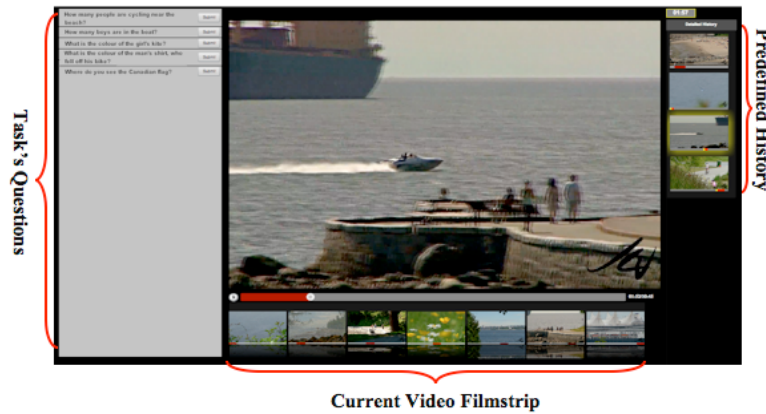

Figure 6. The Experiment Interface[2] illustrates the familiarity phase.

5.4   Procedure

Participants were given a description of the procedures to be employed in the study, informed of the goals and purposes of the study, and informed consent was requested to participate in the study. Two methods were evaluated in the study: Filmstrip and History Timeline.

The researcher started by showing the interface (shown in Figure 6), describing its components, explaining the functionality of the tools within the interface and how to answer a question in order to complete each task. Participants were then asked to try the same interface with the first video and task. They were given the freedom to play with the interface and ask any questions. This familiarity stage was intended to allow the participant to try all the interface components and become familiar with the available functionality. The participants started watching the video provided without knowing the questions they will be asked.

---

2 The video images are for illustration purpose only as different stimulus was used in the experiment (i.e. Toy Story 3, Alma, Geri's Game, and For the Birds.).

For each video, the task started by playing the segments from the pre-defined history in the main video player. Participants were asked to watch the playback and pay attention to the video content in order to answer questions later. Once all segments from the history were viewed, the video paused and a list of questions (compiled by researchers) was displayed to the left of the main video player, as shown in Figure 5. Participants were asked to read these questions before clicking on the "Start" button, when the timer starts and they could begin providing answers. After clicking on the "Start" button, the method's corresponding components were only shown and participants were advised to answer as many questions as they can within the given time by watching the clip which contains the answer.

The questions were randomly ordered (i.e. they were not temporally ordered according to their occurrence in the video). Each participant had the freedom to choose the order they would like to follow to answer these questions (e.g. temporally, linearly, difficulty, memorability, etc.). The participants started answering these questions by clicking on a "Start" button after reading the list of questions. In the familiarity phase, the participant could use either the Filmstrip or History Timeline to navigate the video in an attempt to find an answer. To answer a question the participant needed to re-watch the clip that contains the answer in the main video player. They navigated Video Segments in the provided component (Filmstrip or History Timeline) to find a clip or a video frame that contains the answer. Once they found the frame, they dragged the Video Segment of a specific video frame using the seek thumb of the Video Segment to the main video player to start playing that clip. After watching the interval that contains an answer for a single question they clicked on a submit button beside the corresponding question to proceed to the next question.

Each submitted clip was automatically evaluated: it was considered correct if it contains at least one frame from the ground truth answer. If the answer was correct, the question was faded out from the list and a message was displayed to participants (e.g. "Great. You have 2:14 to answer the rest. Hurry up.") to encourage them and maintain time pressure. However, if the answer was wrong, a message is shown which asks the participant to try again "Sorry. Try again". Participants repeated the same actions till they answered all questions or the time elapsed. A new list of questions was displayed and the participants continued applying the same procedure to answer these questions. The questions for each video were divided into two lists to avoid overloading participants with so many questions. A short break was provided between question sets.

After completing the second list of questions, the familiarity stage ended and the participants were advanced to the phase 1 of the experiment where they watched a new video in similar settings to the watching phase of the familiarity stage. Upon the completion of the watching phase, a new list of questions was given based on the content of this video. In this stage the user was provided with only one method (Filmstrip or History Timeline) to be used in finding the answers. Once all questions were answered or the time elapsed, the participants proceeded to the next video where they repeated the previous stage but with a different video and using a different method than the one used for the previous video. Participants continued tasks until they completed all four videos.

During the experiment, participants were asked to fill out three questionnaires. These questionnaires were designed using the standard Questionnaire for User Interface Satisfaction (QUIS version 6.0 [4]), modified to reflect the functionality and tools applied and the usability of our interface. Some sections were removed (e.g. Terminology and system information, and System capabilities), some questions were modified, and others were added. The first questionnaire was after V4 in which participants evaluated the method they used for that video. The second questionnaire was after V5 where participants rated the usability of the second method they used. After completing the second questionnaire, participants were advanced to the second phase of the experiment. They were given the full interface, described in section 4, where they have the freedom to create their own history and explore use cases which might be applicable to them. Every performed seek within the video or any watched interval in the main video player was recorded and visualized as a Video Segment in the History Timeline. This helped the participants to understand how history is created for later usage. Once they were satisfied playing with the interface, they were asked to fill out the last QUIS questionnaire, which evaluates the entire interface.

Table 1. Methods Descriptives Using t-test for Equality of Means

	Filmstrip		History Timeline		
	M	SD	M	SD	t-test
Mean % of answered questions	65.53	14.69	83.93	14.87	4.53^{**}
Average time per question	24.11	5.85	18.52	3.85	4.13^{**}
Average no. of seeks	1.59	0.45	1.50	0.48	ns
Average no. of previews	14.37	5.14	12.74	4.91	ns
Average no. of errors	0.15	0.09	0.14	0.11	ns

$^{**} p < .01$.

Note. M = Mean. SD = Standard Deviation. ns = not significant. Average time per question is measured in seconds.

5.5 Results and Discussions

The study showed positive and promising results on the features of the interface. Most participants commented that they enjoyed their time using the interface and they can imagine seeing its features applied, especially in social networking websites (e.g. YouTube). They were impressed with the ease of learning and using the features within the interface.

T-test analysis results, illustrated in Table 1, showed that our novel method (History Timeline) was significantly faster than the state-of-the-art (Filmstrip) which allowed participants to answer more questions within the same time. However, both methods demonestrated similar behaviour in terms of average number of seeks, previews and errors.

Each participant answered two identical questionnaires, one for each method, to compare the overall reaction, the time it takes to learn and general impressions. Each questionnaire contains thirty-seven 7-point Likert scale questions. The data collected

for the two methods were analyzed and compared in terms of these 37 questions. The results showed a significant difference taking the overall average rating of each method (Filmstrip ($M = 5.45$, $SD = 1.35$), History Timeline ($M = 5.55$, $SD = 1.06$), $t(886) = 2.08$, $p < 0.04$). The History Timeline showed no significant difference from the Filmstrip in all questions except for the question "Learning to use the method features" where (1: Difficult, 7:Easy) History Timeline ($M = 5.4$, $SD = 1.07$), and Filmstrip ($M = 4.7$, $SD = 1.16$), $t(22) = 2.27$, $p < 0.04$. This also coincided with participants preference where 9 participants out of 12 stated that the History Timeline was faster and easier in finding the answers for the questions. Participants reported that History Timeline was faster because it is based on a personal mental context map created for the video. It was easier for them to refer back to the corresponding segments when needed which is not the case for Filmstrip. Since Filmstrip segments was created systematically, participants needed at least to navigate one segment. If the answer was within that segment they submitted it, otherwise they needed to navigate the preceding or subsequent segment. This was also demonstrated by the participants' quantitative results where they were able to answer more questions using History Timeline. All questions for both methods were rated above 5 out of 7 except for four questions: "Satisfaction with the method", "Learning to use the method features", "Number of steps per task", and "Using the method is effortless".

After acquiring some familiarity with the overall interface in phase 2, participants were asked to complete a third questionnaire to rate the entire interface, described in section 4. Responses were compiled for each of the 12 participants in the study, along with any written comments that the participants had. The overall mean rating of all sections of the QUIS was 5.73, on a 7-point scale, and all questions were rated above 5. For the "Overall Reaction" section, "Satisfaction" overall rating was the only factor that was rated significantly lower than the mean response ($M = 5.2$, $SD = 0.57$), indicating that this area may need additional consideration. The "Ease of use" overall rating ($M = 6$, $SD = 0.53$) was significantly more than the mean response, indicating that users can easily learn and use the interface. From these "Overall Reactions" to the interface, we can conclude that users found our interface easy to use, helpful, useful, and flexible. In the "General Impressions" section, the overall ratings for "Screens are aesthetically pleasing", "Screen designs and layout are attractive", "Interface is impressive", "Interface can do a great deal", and "Interface is fun to use", were rated significantly better than the mean response. This indicates that users found our interface aesthetically pleasing and fun to use. The remaining items of this section were not significantly different from the mean response.

In addition to the QUIS item findings, participants were also asked to list three most negative and positive aspects of the interface. Some participants reported that "I found it weird to have a vertical list of video pieces", "Not being able to delete segments from history", and "Not being able to favorite some segments from history". The comment about the vertical scrolling coincided with their responses to the location of the component where the mean rating for the layout of History Timeline was significantly lower than that of the Filmstrip. We designed the component in this layout to eliminate the confusion between the History Timeline and the video timeline; this may need additional investigation. Participants found the interface helpful and

impressive in being able to dynamically create points or bookmarks, which would allow them to skip to a favorite clip. Most participants stated that the interface is fun to use once you get the hang of it.

In order to explore whether users foresee some potentials for the interface we asked them "Where and how do they think video navigation history can be useful". Participants provided us with valuable responses that gave us some insights of how this interface can be further modified and tested. They foresee that it would be useful in educational environment as well for home usage. There were some overwhelming comments made by the participants about the interface. One participant commented "I definitely see how this would be really helpful for long videos because I will not have to waste my time watching the whole video again to get to the important stuff. I could directly use my previous history to navigate to these intervals". Others said, "I would love to see this implemented within social websites. I could see how I would use it and definitely I will have more fun"; "I need to have this. Could we have it in YouTube?"; and finally "I really do like this interface and I would love to have it to create wonderful clips from my home videos".

Based on the participants' valuable comments and suggestions, we believe using personal navigation history is helpful in navigating a video space. This could work for different applications such as highlights or summary of videos, a movie using home videos, sharing interesting clips, quickly navigating and skimming previously watched videos, and finally watching new videos' interesting parts using crowd-sourced viewing histories. Receiving these positive feedbacks from the participants and how they welcomed the idea of using their personal navigation history motivate us to investigate how our interface would work if integrated with social websites (e.g. YouTube, Vimeo) and explore how this could change the way users consume and navigate video.

6   Conclusion and Future Work

The work we have presented describes a new way to navigate and view video media using a personal video history. We introduced a novel interface for video navigation through the creation and subsequent use of a user's personal viewing history. The interface is based on observations of new use cases involving re-watching of within-video intervals and the increasingly temporal nature of video navigation. We performed a study that found positive results and highly positive affect and comments from participants. A comparative study based on a use case of fast navigation found significant results in favor of our method. To our knowledge, this is the first paper on using history navigation through video browsing, and as such, not enough work has been done in this area to evaluate the effectiveness on a larger scale. So we intend to investigate scalability, filtering mechanisms, interactions required when using a multiple-video history to allow users with large history to accurately navigate through them. Once these mechanisms have been tested, we aim to deploy a field study to check the validity and the scalability of history in action.

References

1. Bergman, L., Castelli, V., Lau, T., and Oblinger, D. Docwizards: a system for authoring follow-me documentation wizards. In Proceedings of the 18th annual ACM symposium on User interface software and technology, UIST '05, ACM (New York, NY, USA, 2005), 191–200.

2. Card, S. K., Pirolli, P., Van Der Wege, M., Morrison, J. B., Reeder, R. W., Schraedley, P. K., and Boshart, J. Information scent as a driver of web behavior graphs: results of a protocol analysis method for web usability. In Proceedings of the SIGCHI conference on Human factors in computing systems, CHI '01, ACM (New York, NY, USA, 2001), 498–505.

3. Cheng, K.-Y., Luo, S.-J., Chen, B.-Y., and Chu, H.-H. Smartplayer: user-centric video fast-forwarding. In Proceedings of the 27th international conference on Human factors in computing systems, CHI '09, ACM (New York, NY, USA, 2009), 789–798.

4. Chin, J. P., Diehl, V. A., and Norman, K. L. Development of an instrument measuring user satisfaction of the human-computer interface. In Proceedings of SIGCHI '88, ACM/SIGCHI (1988), 213–218.

5. Christel, M., and Moraveji, N. Finding the right shots: assessing usability and performance of a digital video library interface. In Proceedings of the 12th annual ACM international conference on Multimedia, MULTIMEDIA '04, ACM (New York, NY, USA, 2004), 732–739.

6. Cockburn, A., and Mckenzie, B. What do web users do? an empirical analysis of web use. International Journal of Human-Computer Studies 54 (June 2001), 903–922.

7. Dragicevic, P., Ramos, G., Bibliowitcz, J., Nowrouzezahrai, D., Balakrishnan, R., and Singh, K. Video browsing by direct manipulation. In Proceedings of the twenty-sixth annual SIGCHI conference on Human factors in computing systems, CHI '08, ACM (New York, NY, USA, 2008), 237–246.

8. Drucker, S. M., Glatzer, A., De Mar, S., and Wong, C. Smartskip: consumer level browsing and skipping of digital video content. In Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves, CHI '02, ACM (New York, NY, USA, 2002), 219–226.

9. Foote, J., Boreczhy, J., Girgensohn, A., and Wilcox, L. An intelligent media browser using automatic multimodal analysis. In Proceedings of the sixth ACM international conference on Multimedia, MULTIMEDIA '98, ACM (New York, NY, USA, 1998), 375–380.

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

11. Goldman, D. B., Gonterman, C., Curless, B., Salesin, D., and Seitz, S. M. Video object annotation, navigation, and composition. In Proceedings of the 21st annual ACM symposium on User interface software and technology, UIST '08, ACM (New York, NY, USA, 2008), 3–12.

12. Grossman, T., Matejka, J., and Fitzmaurice, G. Chronicle: capture, exploration, and playback of document workflow histories. In Proceedings of the 23nd annual ACM symposium on User interface software and technology, UIST '10, ACM (New York, NY, USA, 2010), 143–152.

13. Guimbretiere, ́ F., Dixon, M., and Hinckley, K. Experiscope: an analysis tool for interaction data. In Proceedings of the SIGCHI conference on Human factors in computing systems, CHI '07, ACM (New York, NY, USA, 2007), 1333–1342.

14. Heer, J., Mackinlay, J., Stolte, C., and Agrawala, M. Graphical histories for visualization: Supporting analysis, communication, and evaluation. IEEE Transactions on Visualization and Computer Graphics 14, 6 (Nov. 2008), 1189–1196.

15. Hupp, D., and Miller, R. C. Smart bookmarks: automatic retroactive macro recording on the web. In Proceedings of the 20th annual ACM symposium on User interface software and technology, UIST '07, ACM (New York, NY, USA, 2007), 81–90.

16. Karrer, T., Weiss, M., Lee, E., and Borchers, J. Dragon: a direct manipulation interface for frame-accurate in-scene video navigation. In Proceedings of the twenty-sixth annual SIGCHI conference on Human factors in computing systems, CHI '08, ACM (New York, NY, USA, 2008), 247–250.

17. Kimber, D., Dunnigan, T., Girgensohn, A., Shipman, F., Turner, T., and Yang, T. Trail-blazing: Video playback control by direct object manipulation. In Multimedia and Expo, 2007 IEEE International Conference on (July 2007), 1015 –1018.

18. Leftheriotis, I., Gkonela, C., and Chorianopoulos, K. 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).

19. Li, I., Nichols, J., Lau, T., Drews, C., and Cypher, A. Here's what I did: sharing and reusing web activity with actionshot. In Proceedings of the 28th international conference on Human factors in computing systems, CHI '10, ACM (New York, NY, USA, 2010), 723–732.

20. Mertens, R., Farzan, R., and Brusilovsky, P. Social navigation in web lectures. In Proceedings of the seventeenth conference on Hypertext and hypermedia, HYPERTEXT '06, ACM (New York, NY, USA, 2006), 41–44.

21. Nakamura, T., and Igarashi, T. An application-independent system for visualizing user operation history. In Proceedings of the 21st annual ACM symposium on User interface software and technology, UIST '08, ACM (New York, NY, USA, 2008), 23–32.

22. Pongnumkul, S., Wang, J., and Cohen, M. Creating map-based storyboards for browsing tour videos. In Proceedings of the 21st annual ACM symposium on User interface software and technology, UIST '08, ACM (New York, NY, USA, 2008), 13–22.

23. Sawhney, N., Balcom, D., and Smith, I. Authoring and navigating video in space and time. IEEE Multimedia 4, 4 (Oct. 1997), 30–39.

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

25. Syeda-Mahmood, T., and Ponceleon, D. Learning video browsing behavior and its application in the generation of video previews. In Proceedings of the ninth ACM international conference on Multimedia, MULTIMEDIA '01, ACM (New York, NY, USA, 2001), 119–128.

26. Yeung, M. M., and Yeo, B.-L. Video visualization for compact presentation and fast browsing of pictorial content. Circuits and Systems for Video Technology, IEEE Transactions on 7, 5 (oct 1997), 771 –785.

27. Yew, J., Shamma, D. A., and Churchill, E. F. Knowing funny: genre perception and categorization in social video sharing. In Proceedings of the 2011 annual conference on Human factors in computing systems, CHI '11, ACM (New York, NY, USA, 2011), 297–306.

28. Yu, B., Ma, W.-Y., Nahrstedt, K., and Zhang, H.-J. Video summarization based on user log enhanced link analysis. In Proceedings of the eleventh ACM international conference on Multimedia, MULTIMEDIA '03, ACM (New York, NY, USA, 2003), 382–391.