An Investigation of Textbook-Style Highlighting for Video

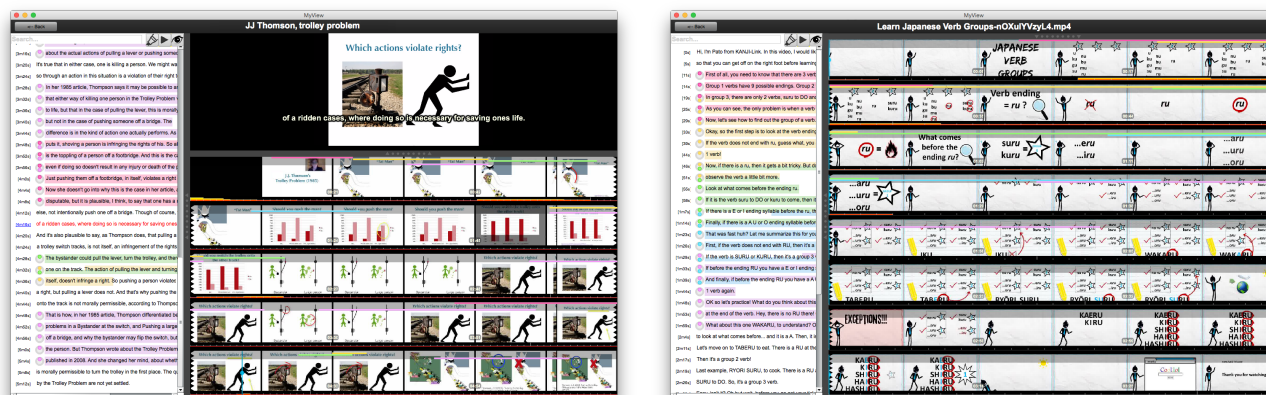**Anonymous for Review**
Submission #798

Figure 1. Our prototype interface allowing users to mark up video in a familiar way using highlights. The player view shows users a transcript of the entire video, as well as visual previews of the entire video (*left*). Text and visual preview can be selected and highlighted to assist users with reviewing the material later on. With the Filmstrip maximized to take up most of the screen and the transcript to the left, the user can perform both a textual search and a visual search.

ABSTRACT

Video is used extensively as an instructional aid within educational contexts such as blended (flipped) courses, self-learning with MOOCs and informal learning through online tutorials. The challenge is providing mechanisms for students to manage their video collection and quickly review or search for content. We provided students with a number of video interface features to establish which they would find most useful for video courses. From this, we designed an interface which uses textbook-style highlighting on a video filmstrip and transcript, both presented adjacent to a video player. This interface was qualitatively evaluated to determine if highlighting works well for saving intervals, and what strategies students use when given both direct video highlighting and the text-based transcript interface. Our participants reported that highlighting is a useful addition to instructional video. The familiar interaction of highlighting text was preferred, with the filmstrip used for intervals with more visual stimuli.

Author Keywords

Video, search, highlighting

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION

There is a long history of video being used as an educational tool, such as documentaries and visual tutorials. Streaming video sites, such as YouTube, have led to an enormous number of instructional videos across many topics, contributed from individuals on hobbies and interests up to formal education contributed from well-known universities. With the rise of online education frameworks (such as edX[1]), video has become a core aspect of courses designed for self-learning. Massive Open Online Courses (MOOCs) use video substantially, often as the central teaching aid in place of lectures, to offer complete courses that are delivered over the web. In 2014, nearly 20,000,000 learners registered to at least one MOOC[2]. Other sources, such as Khan academy[3] offer more targeted videos for learners who seek information on specific topics.

The challenge with using large quantities of video to teach courses is providing students with effective management tools. Traditional playback controls (play/pause, seek, preview) are generally used, which do not support recall, history or interval bookmarking. This requires the user to either note down timecodes for intervals or visually search the video when they view it again. However, video is a complicated medium to navigate, so it is not clear which types

---

[1]https://www.edx.org
[2]https://www.edsurge.com/news/2014-12-26-moocs-in-2014-breaking-down-the-numbers
[3]https://www.khanacademy.org

of interactions would be helpful within a learning context. The first contribution of this work is an investigation with university students to determine which type of video interfaces would help them with their learning experiences. We created a prototype interface with various features including filmstrip preview, view count visualization, personal viewing history, interval bookmarking and playlists. Students were individually interviewed while trying each feature and asked which they found to be useful or if they would modify or add aspects. Based on their feedback, we identified the most promising features and ran a focus group to obtain ideas on how to present them for instructional content.

Our second contribution is an investigation into highlighting of video using a textbook metaphor. Based on the outcomes of the interviews and focus group, we designed a prototype interface supporting video playback, view count visualization, transcript navigation and a 2D filmstrip. The transcript navigator and the filmstrip both allow "highlighting" of video intervals with various colours, exactly as one may do with a textbook. The goal of this prototype interface was to discover if the textbook metaphor of highlighting is effective and the use of highlights within instructional video is helpful for students.

There are various aspects of the video highlighting problem which we investigated with a qualitative evaluation:

- The main research question: Does the textbook metaphor of highlighting work for instructional video?

- If so, then what type of content would they highlight?

- When do students highlight the content? (I.e. while viewing, in advance, or afterwards)

- Given a number of highlight colours, how do students assign meaning to these, and would they prefer pre-assigned meaning?

- How many colours should be available?

- Finally, what are the strategies used to highlight, or later find highlights? For example, did they use the transcript navigator more often for video with similar-looking visual content?

After a review of the literature we present the preliminary investigation, followed by a description of the new interface developed to test video highlighting and the evaluation performed to determine its utility.

RELATED WORK
The development of video interfaces within educational contexts is an important topic for contemporary learning. We first cover prior work from video interfaces for navigation and annotation, and then discuss pedagogical strategies and the use of video for learning.

Video Navigation
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, and allowed users to directly manipulate the preview frame by moving a cursor along a timeline to find the desired clip [11]. Further work by Matjeka et al. [20] utilize multiple thumbnails to help users explore video. Kimber et al. [18], Dragivic et al. [8] and Karrer et al. [15] encouraged users to directly manipulate within-video content along its movement path to seek around video.

Moving towards traditional timelines, we see inherent problems as the length of video becomes larger. Representation of each time-step becomes too small to see and use. Work by Hürst et al. [13, 14] introduced the ZoomSlider, which hid parts of the timeline, and allowed users to shift the slider across the screen to seek across the video. The seeker bar is zoomable, and the granularity of seeking is dependent on the cursor's vertical position, allowing for higher accuracy in seeking for longer videos. This has been adopted commercially and can be seen in Apple's iOS video player.

There has been a recent trend towards allowing users to utilize their navigation histories in video, much like those found in web browsers. Al Hajri et al. visualize detailed video navigation history for general video in [3, 4], allowing users to view and explore previously seen video.

Video Annotation
There are several annotations tools being developed, some commercial, and some open source [4], allowing users to mark up video, involving different methods of annotation, including text, ink, audio, or video. In research there has been work in video annotation tools designed for various devices, such as mobile [7], or stylus based [25, 27], collaborative annotation tools [26, 1, 9], as well as automatic annotation via computer vision techniques [12, 23]. The focus of the work in video annotation has been for general video and allowing In more specific contexts, such as education, annotation systems have been developed more for instructor use. To focus more on student roles in usage of video for education would be invaluable.

Video Education Strategies
One common use of video within education falls under the category of flipped classrooms, in which students watch instructional videos at home in order to prepare themselves for class [5]. The videos are often created by the instructors themselves, and include a standard "lecture", often using slides. The slides or lecture backdrop can lead to highly similar-looking content which can be difficult to navigate visually, so the use of a transcript for navigation was included in our interface.

While the use of video to support learning has become more popular, the style has remained largely intact as a linear lecture [28]. This can be improved in some simple ways; for example, people learn better from instruction that includes text and video next to each other, compared to either in isolation
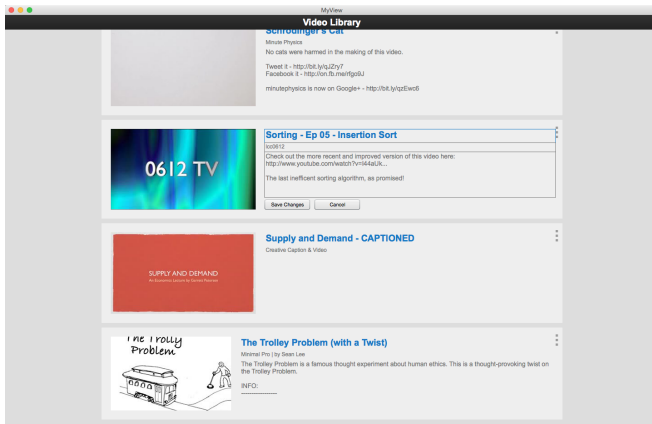
---

[4]http://www.annotations.harvard.edu/icb/icb.do?pageid=icb.page466612

Figure 2. The Video Library screen. Each video has a title, an author, and a description. The second video "Sorting - Ep 05 - Insertion Sort" currently being edited.
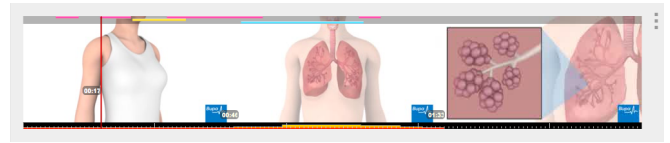

Figure 3. The Filmstrip as seen in the Video Library when the cursor is placed over a video. This at-a-glance view shows the user what parts of the video has been highlighted (top, multicolour) and what parts they have seen (bottom, red/orange). The user can also move the mouse across the Filmstrip to preview the video.

[6]. Additional evidence suggests that learning is improved when students have control over their learning and can watch videos in segments, at their own pace [6]. Our viewer is designed to accommodate this, and provide students with the ability to segment the video in their own way for later review.

To our knowledge, highlighting, a major aspect of textbook learning, has not yet been tested on video. Learners learn better when they can highlight information that they find valuable [19], which we attempt to provide them in our interface.

Video Interfaces for Education
Mertens et al. developed the idea of video footprints, a detailed view count for the length of the video based on personal viewing, visualized as a heatmap timeline [21]. This was presented together with temporal bookmarks (at a single time instant) and navigation using either slides or time. These ideas were further explored by Kim et al. [17], through the use of interaction data (including video footprints) to help students navigate instructional video. Frequently-viewed regions are isolated within a histogram, and video summaries can be generated based on personal or crowd-sourced interactions. The interface also supports transcript visualization and temporal bookmarking. Our work extends these interfaces with interval bookmarking (highlighting) coupled to a highlighted transcript, together with a heat-map based view count to aid navigation. Pavel et al. presented an interface to create video digests, smaller chunks of video with previews and text summaries, to help people skim and browse long videos [24]. The interface is based on a transcript of the video, with automatic and crowd-sourced first passes and a manual refinement step if needed. The creation of chapters using transcripts is similar in principle to our approach, using intervals of the video as the basis for subsequent navigation.

RIMES explored the use of interactive multimedia exercises embedded within lecture videos, with student feedback recorded using video, audio and sketching for the instructor to review later [16]. Each video was self-contained with its exercises, with annotation available to the student to answer questions, but no mechanisms were present for later refer-

ral. Monserrat et al. presented an education environment called L.IVE, which provided students with the ability to discuss parts of the video with temporal comments (marked at a single time in the video), along with assessments included with the video [22]. This provided students with an integrated viewing and assessment environment, however the navigation was based on the traditional controls and did not organize the comments to aid navigation.

PRELIMINARY INVESTIGATION
As a start to determining the usefulness of current video interfaces we decided to perform interviews and focus group studies on students who would be able to provide proper feedback. We chose interfaces that we believed would support users in studying: the navigation history visualizations provided by [4], the revisitation heatmaps seen in [17], a filmstrip type visualization seen in [20], and a form of authoring functionality seen in [10]. We presented these components to students and collected feedback to inform the design of a video highlighting solution.

Interviews
We conducted interviews with twelve participants to identify how students watched videos for their learning online, and in regular classes. The participants were aged 19 to 33, three male, eight female, and all the participants had taken at least one course with video. We introduced to the participants the idea of the flipped classroom, where video lectures are watched at home, and "homework" is completed in class, and used that as the context in which the participants should think about answering the questions we posed to them. We demonstrated implementations of interfaces mentioned above to the participants and asked them about their thoughts on how they would use them while studying. General consensus was that being able to see statistics of how the video was watched, and how other students watched the video would be useful. Furthermore, the participants liked the ability to author their own videos to create their own review material.

Focus Group
In order to better understand the results of the interviews, we conducted a focus group study in which we brought in students who had just completed an offline philosophy class that was taught partially with video using a standard video player (YouTube). We had seven participants in this focus group aged 18 to 26, two male, and five female. This time, we allowed the participants to download a copy of the interface to their laptops and they were given a tutorial for each
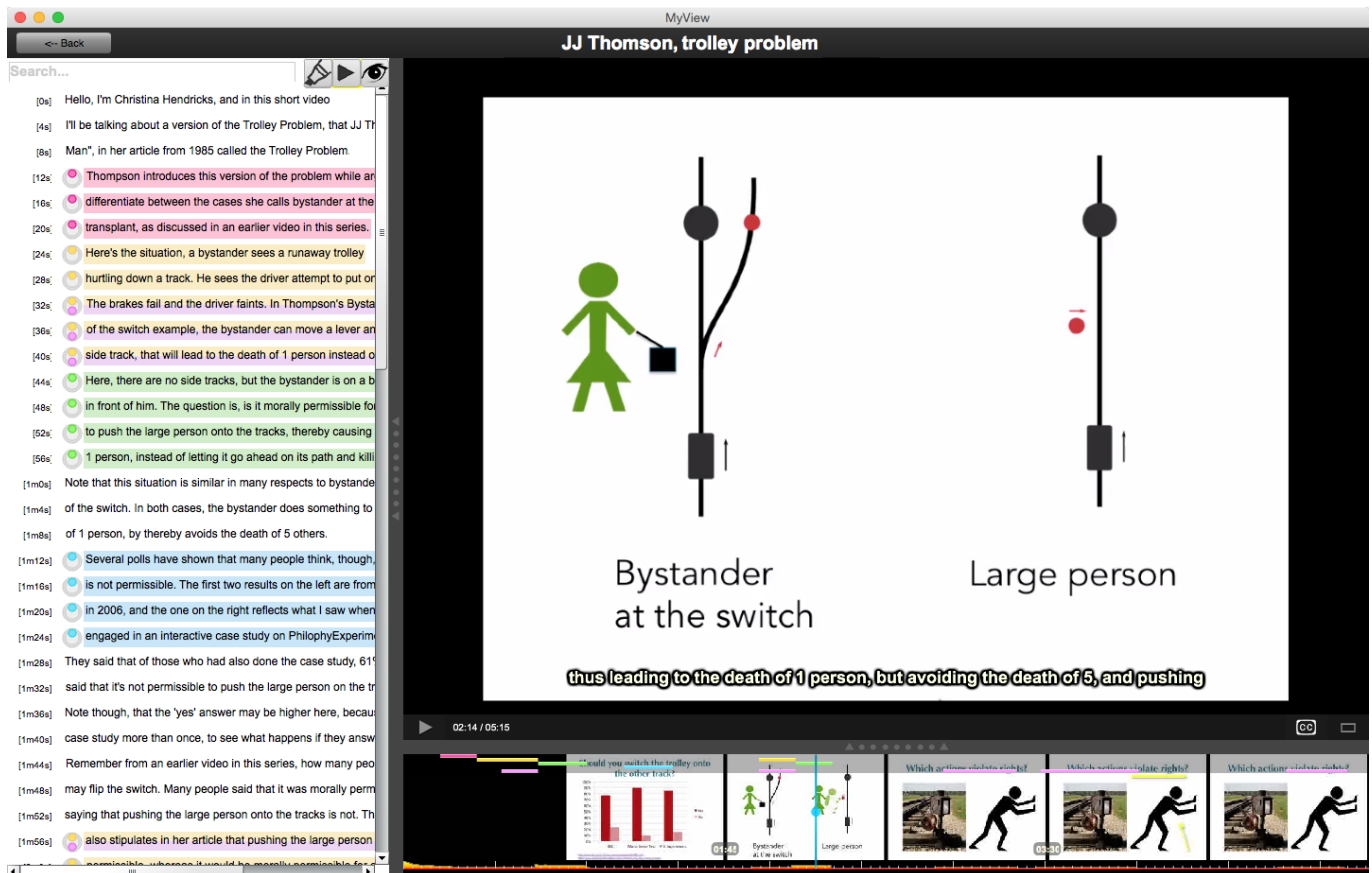
Figure 4. The main video player view. The filmstrip (red), the player (green), and the subtitle viewer (blue), reside here to help the user watch, review, and highlight the video.

interface and given time to use it and acclimatize to the interface. While the participants used the interface, we asked them to speak aloud any frustrations or ideas they had, to better inform us of a design of an interface that can support their studying needs.

Outcomes

The interviews and focus group showed that being able to view a history of a user's behaviour ([4]), however, the interface was difficult to use and the students were frequently confused. Being able to select intervals of video for manipulation ([10]) was useful in terms of organization. The idea of being able to keep track parts of the video that have been seen more often, or parts of the video that have not been seen at all ([17]) was intriguing to the participants of the focus group. They wanted more control over it, so be able to emphasize or de-emphasize certain parts of video without having to watch it over and over. Furthermore, they liked the idea of the playlist and being able to use it as an organizational structure mechanism.

Looking at the implications of user-adjustable emphasis and video organization tools, we were led to the development of the proposed textbook metaphor for highlighting video and the prototype interface described in the following section.

INTERFACE

The main idea behind the interface is to allow users to manipulate emphasis on certain parts of video and be able to use that to organize their video in their own way. Thus, the interface was designed to allow users to highlight video, and then play back those highlights. In this way, they are emphasizing video (by highlighting it), and creating playlists (by being able to play back highlighted parts). We designed a video player around this concept and took cues from physical highlighting tools as well as digital ones found in document readers. The interface consists of two screens and three major elements. The first screen, shown in Figure 2 allows the user to explore the videos offered by the interface. Clicking on one of the videos brings the user to the video player screen, shown in Figure 4. This screen consists of the three elements: the filmstrip (red), the player (green), and the subtitle viewer (blue). This section will describe the design guidelines and each of the interface elements in detail.

Video Library

The video library shows a list of videos offered by the interface. The user can see at a glance, the title, the author, the description, and a preview thumbnail. Upon putting the cursor over an item, a filmstrip, described in the next section, appears, allowing the user to quickly explore the visual contents of the video, see the highlighted parts of the video, as well as see which parts of the video have been seen. This can
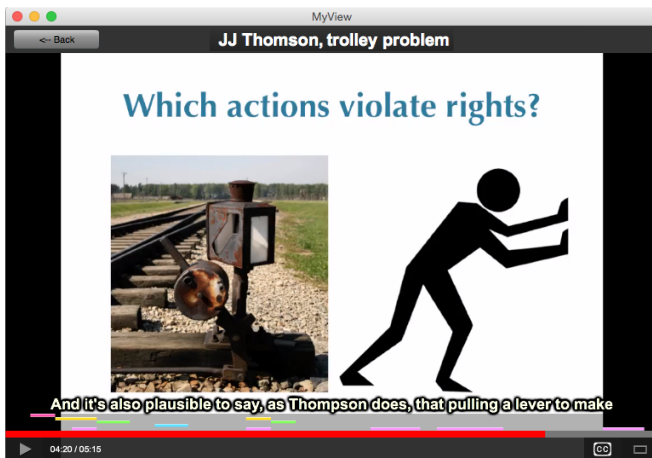
Figure 5. The main player in fullscreen view. In this mode, all the other interface elements are hidden, reducing distractions from the video. The highlights made in the video are visible at the bottom.

Figure 6. Subtitle Viewer Toolbox expanded to show the options of colours available to highlight.

been seen in Figure 3. Each video's details can also be edited to the user's liking.

Filmstrip

The filmstrip is a set of thumbnails from the video arranged side by side, each representing a portion of video. As the width of the filmstrip represents the entire length of the video, each *n* thumbnail represents 1/*n* of the video. In the case of the video library, there are three thumbnails, and each thumbnail represents one-third of the video. Moving the cursor over top causes a floating timestamp to appear, and changes the corresponding thumbnail to show the frame represented by the horizontal location of the cursor. The initial preview in the thumbnail is the first frame the thumbnail represents in the interval.

On the top side of the filmstrip is a bar indicating the highlights made in the video across the filmstrip, with respect to the ratio between the length of the video and the width of the filmstrip. In pilot tests, users found that in certain videos, slide-based ones in particular, the filmstrip resembled a set of slides instead of an actual timeline. In order to rectify this, we chose to insert time-ticks underneath at the bottom, distributing them evenly to a minimum of either one tick per second, or one tick every 10 pixels across the width of the filmstrip. Along the time ticks, there are two timestamps placed at 1/3 and 2/3 of the width of the filmstrip to indicate time. Underneath the time ticks, the filmstrip shows a heatmap of the user's visitation of the video. The height of the graph is logarithmic, but never overlaps with the bottom of the thumbnail, and the colour ranges from red to orange to yellow, with red being least seen and yellow being most seen. This visualization is similar to [17], and deals with scalability better, while less confusing than [2].

Video Player

The video player, Figure 4, allows the user to view and markup the video. It consists of three major elements, the player, the filmstrip, and the subtitle viewer.
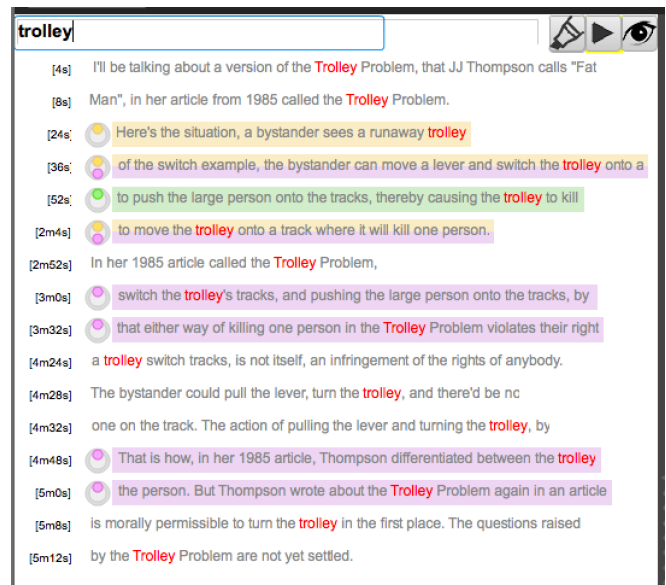
Player

The player, shown green in Figure 4 is the main focus of the interface, and is typical of those usually found in a video player. On the bottom is a toolbar that houses video controls, allowing the user to play (▶) or pause (❚❚) the video, view the currently playing time, turn on and off closed captioning (CC), or make the video fullscreen (▭). When the video is made fullscreen, a seekbar appears on top of the toolbar, allowing the user to seek without view of the Filmstrip, which will be described later on. Furthermore, when the player is made fullscreen, a bar of highlights appears on the top above the seek bar. This can be seen in Figure 5.

Filmstrip

The filmstrip in the video player is very similar to the filmstrip found in the video library, with some extended functionality. As we are in the player, the filmstrip provides the user with the ability to highlight intervals.By clicking and dragging across the thumbnails, the user can select intervals, and then a button will pop up. When the cursor is over the button, more buttons will pop up, allowing the user to select the colour they want to highlight in.

The filmstrip in the player can also be expanded. By clicking dragging on the bar above it, the filmstrip can expand into multiple rows, and the scale of time is expanded across all the filmstrips. For example, in Figure 7, there are three filmstrips and the width of each filmstrip represents one-third of the video. Some of the filmstrips ends and beginnings are jagged, indicating that they represent only parts of the video and not all of it.

Subtitle Viewer

The subtitle viewer, shown on the left in Figure 4 in blue, is a transcript of everything said in the video. This provides the user with an overview of video in textual form, allowing them to quickly scan through the spoken content of the video. As the video plays through, the text being currently spoken turns

Figure 7. The Filmstrip, when resized, will split into multiple rows, each representing a portion of the video. Here, each row represents one-third of the video.

red and acts like a playhead. On the left of each caption is a timestamp, which the user can click, causing the player to seek the the time indicated.

Like the filmstrip, the user can click and drag across text to select it, and using the ✎ button, the user can highlight the selected text in the chosen colour. Like in the filmstrip, text can be highlighted in more than colour, and each colour will stacked on top of each other underneath the text. Once a line of text has been highlighted, a miniature widget appears next to the text indicating to the user the colour the text has been highlighted. This is useful when the text has been highlighted in multiple colours, as seeing which colour underneath the text can be difficult.

The ▶ button allows the user to playback highlights and the 👁 hides the highlights within the text. Included is also search functionality, allowing the user to search for text within the subtitles. When the user enters search terms, captions that do not contain any matches disappear and captions with a match remain on the screen. The text is greyed out with the exception of the matching term within the text, which is marked in red. The result of a search can be seen in Figure 6.

Design Strategies

In an effort to maintain consistency across the entire interface, there were some rules that we adhered to when designing the interface. For each In each of the interface elements in the video player, selections and highlights performed in one element is reflected in the others instantly. Colours always appear in the same rainbow order.

EVALUATION

We ran a qualitative user study to investigate the usability and utility of the various methods for highlighting video presented in our system. The qualitative study provided insight into the patterns of highlighting material.

Participants

11 volunteers participated in the experiment: 6 male and 5 female, ranging in age from 18 to 33. They were monetarily compensated for their time. All participants have taken at

least one course online, and all but one have taken at least one course that used video as a teaching medium.

Design and Procedure

Each participant was exposed to the three methods for highlighting described above. The participants were given eight different videos to choose from, where each video was 2 to 3 minutes long and of the educational category. The videos were found on YouTube:

1. How chronic obstructive pulmonary disease (COPD) develops[5]

2. How to Balance Chemical Equations[6]

3. How To Read Music - Two Minute Music Theory[7]

4. Learn Japanese Verb Groups[8]

5. Schrödinger's Cat[9]

6. Sorting - Ep 05 - Insertion Sort[10]

7. Supply and Demand[11]

8. The Trolley Problem (with a Twist)[12]

By being able to select a video of their choosing, we hoped they would choose a video that is of interest to them and encourage them to pay attention to the video and use the interface in a way that would increase their enjoyment of the interface. They were asked to watch the video and pretend they were studying the video for a course and to highlight portions of video as needed. The experiment took approximately 30 minutes.

---

[5]http://www.youtube.com/watch?v=iYjKZHmzWEA
[6]http://www.youtube.com/watch?v=8KH3laR2iR4
[7]http://www.youtube.com/watch?v=GiM2OKtACAQ
[8]http://www.youtube.com/watch?v=nOXuIYVzyL4
[9]http://www.youtube.com/watch?v=IOYyCHGWJq4
[10]http://www.youtube.com/watch?v=3orUYqcaEEQ
[11]http://www.youtube.com/watch?v=WdPI3hKUJYo
[12]http://www.youtube.com/watch?v=kKHOpw6tpd4

Question	Score
Easy to use	4.18
Powerful	4.27
Flexible	4.09
Aesthetically pleasing	3.82

Table 1. General reactions to the system on a Likert scale from 1 to 5.

Question	Subtitle Viewer	Filmstrip
Learning was difficult	1.18	1.36
Distracting	1.27	2.18
Utility of highlighting	4.27	3.45

Table 2. Reactions to specific interface elemnents on a Likert Scale from 1 to 5.

The evaluation began by introducing the participants to the idea of being able to markup video as they would in a textbook, and a short description of the experiment. We first started by loading up a demonstration video, and introduced the *Subtitle Viewer* and its highlighting and search functionality. The filmstrip, and the player widget were disabled during this section of the experiment. The participant was then asked to pick a video of their choice, watch it as if they were studying it, and highlight it as necessary with the Subtitle Viewer. Once finished the experimenter asked the participants three questions about the video, that they were told to find answers to in the video. The questions were simple content-based questions and the answers were located around one-third, one-half, and two-thirds of each video. The process was repeated for the filmstrip and the player widget, with the opposing interface elements disabled. After introducing and using three elements, the participants were asked to watch one final video with the entire interface enabled. They were then asked to complete a questionnaire detailing their experience with the interface. The questionnaire contained 21 questions that were answered on a 5-point Likert scale. The experimenter then conducted a post-experiment interview asking about their experience with highlighting video, as well as their highlighting experiences in ebooks and physical textbooks.

RESULTS
Results derived from the questionnaire showed that the system was easy to use ($M = 4.18$), and learning to use each of the interface elements was easy (Subtitle Viewer $M = 1.18$, Filmstrip $M = 1.36$). General reactions to the video can be seen in Table 1. Questions related to each particular interface element are shown in Table 2. T-tests were conducted on each question. For "Learning was difficult", there was no significant difference ($t = 0.452$); for "Distracting", there was a significant difference ($t = 0.045$); for "Utility of highlighting", there was no significant difference ($t = 0.126$). We also asked users to rank their preference of interface element to perform the highlighting and found a statistically significant difference ($\chi^2(1) = 11.000$, $p = 0.001$).

DISCUSSION
The Subtitle Viewer was found to be preferred over the Filmstrip for highlighting, and also found to be less distracting while watching a video. Furthermore, while participants found both easy to use and both had useful highlighting functions, the Filmstrip was significantly more distracting than the Subtitle Viewer. The participants who thought the Filmstrip was distracting cited the large playhead moving across as the video played as well as the thumbnails being distracting. The participants who thought the Subtitle Viewer was distracting cited the large amount of text on the screen. When shown that it was possible to hide the panel, they were satisfied. In most cases however, participants would use the Subtitle Viewer when asked to search for something in the video.

Highlighting
Participants utilized many different strategies of highlighting among the different interface elements. When given only the Subtitle Viewer, the participants utilized the following techniques:

1. Pause the video and highlight the past video.

2. Watch entire video first, then go back and highlight.

3. While the video is still playing, highlight parts that have already been played.

4. Read ahead and highlight ahead of the still playing video.

The first three techniques are all retroactive highlighting, but having the Subtitle Viewer makes it easy for users to look ahead of what is currently playing and evaluate the importance of it. For users, this means that going through a video can be done much faster than real-time, making their studying time more efficient. Some participants in our preliminary interviews have stated that when they watch online videos for studying, they oftentimes speed up the video 1.5 to 2 times faster. This current system does not have this, but allowing users to read the lecture at their own pace allows them to control the speed of their learning. In fact, a couple participants stated that they read ahead of the video, and one participant even skipped forward in the video, having evaluated for himself that a portion of video was not worth watching.

When given only the Filmstrip, participants restricted themselves to retroactive highlighting, such as:

1. Pause the video and highlight the past video.

2. Watch the entire video first, expand the Filmstrip, the go back and highlight.

3. While the video is still playing, highlight parts that have already been played.

4. Rewind and highlight the appropriate part while the video is playing.

While the participants remembered to highlight using the Subtitle Viewer, they had to be prompted to highlight when they were given the Filmstrip. When given both interfaces participants, only two participants decided to use the Filmstrip at all to highlight, after having watched the entire video and going through again to highlight important points. Evidently, the highlighting task is not only encouraged by the Subtitle Viewer, but easier and more convenient to do as well. When asked about highlighting in the Filmstrip, participants

said that it was difficult to see what was being highlighted due to the lack of knowledge about what is being said during the highlighted interval.

Search

The search task required the participants to find portions of video using either interface element. Participants had more trouble with the search task using the Filmstrip than the Subtitle Viewer. This is due to the amount of distinguishable information that can be displayed by each interface element. In the Subtitle Viewer, the text is easily distinguishable, and it is easy to pick out certain words, especially when aided by the search function. The Filmstrip, while proved in other works to be useful for general video which contains many scene changes, is difficult to use due to the nature of educational videos. However, in the case that the video contained screens of varying scenes, the searching using the Filmstrip became far easier. In the experiment, sometimes, depending on how the Filmstrip was expanded, the exact frame containing the answer to the search task appeared in the Filmstrip and participants were able to quickly hone in on the correct location.

As far as the search task itself, when the participants were restricted to the Subtitle Viewer, they:

1. Watched through the video again.

2. Skimmed through text.

3. Used the text search.

4. Looked for the highlight when they thought it was highlighted.

Looking at the methods used, one might assume that using the text search is not only the easiest option, but also the most used option. However, participants were more likely to scroll to the portion of the text they though the answer was in, and start skimming the text, searching for keywords manually. In these cases, participants stated that because they knew approximately where the answer was, it was easier to skim the text than perform a text search. This is because the text search required the participant to think of an appropriate keyword to search, and because the non-matching results were removed, breaking the linear flow of the text, and it was not clear where the resulting search results resided in relation to the length of the entire video.

When participants were restricted to only using the Filmstrip, they:

1. Expanded the Filmstrip and searched for visual information in the thumbnails.

2. Looked for highlights if they knew they highlighted the part.

Finally, when given both the Filmstrip and the Subtitle Viewer, participants did the following:

1. Seeked to approximate point in the Filmstrip and played the video.

2. If the portion was highlighted, used the highlight playback function.

3. Used the text search.

4. Looked for the highlights in the Subtitle Viewer.

In terms of highlighting strategy, participants stated that they would highlight what they didn't understand, what they thought were key points in the video, as well as review sections in the video. Oftentimes, when the participants were asked questions about something that they highlighted, the participant recognized it and were able to find the location of the video they were looking for, whether it was in the Filmstrip or the Subtitle Viewer. This indicates that the act of highlighting video acts as a memory aid for recall at a later time.

When asked about highlight colours, participants had varied answers, the most popular of which, was that choosing colours was random. In this case, there was no pattern in which they used highlight colours. For some, they would highlight things in the order that they were given. For example, the first thing they would highlight would be red, the second highlight would be in yellow, the third highlight would be in green etc. It is worth noting that of the participants who utilized this method of highlighting were the quickest to find the answers to the experimenter's questions, immediately know which colour to look for and could very quickly find the appropriate time. One participant stated that they would use red for important points in the video and the rest of the colours would be random. Two participants stated that they would colour code their highlights based on the topic within the video, for example, definitions, or instructions sets. All participants except for one stated that the five colours included within the system was adequate for their needs. The participant who wanted more stated that the number of colours should be an even number, for example, four, six, or eight colours. The reasoning behind this was that she wanted contrasting colours so that she could highlight things positively or negatively. When asked if they highlight digital ebooks or PDFs, nine participants said they did; the other two said that they did not. Following the experiment, some of the participants emphasized how much they liked the system with the following quotes: *"Is the app available now?" "I feel like the idea is really good, because you have highlight for ebook, for physical textbook, then why not for videos?", "Are you guys planning on release this thing at all?"*

CONCLUSIONS AND FUTURE WORK

We presented a novel interface for highlighting video in both text and video representations. We performed two preliminary studies to inform us of what kind of interfaces work in a learning and education context, and developed an interface based on those findings. We then evaluated the interface and compared highlighting and search functions across both the Filmstrip and the Subtitle Viewer and found that the Subtitle Viewer was more preferred. The interface as a whole was positively perceived and we conclude that highlighting is a useful tool for educational video interfaces.

We intend to investigate collaborative video interfaces that can be used in and out of classrooms. This would allow students to share information and help each other, as well as allow the instructor more engagement with students. Many of the participants in the preliminary interviews stated that the face-to-face contact was invaluable, and we believe that adding more interaction between instructors and students in a video interface will aid in that respect. We also intend to test the interface on a larger scale, and deploy it a real world setting at a local college or university. That will allow for even more feedback on the emerging video learning interfaces.

REFERENCES

1. Gregory D. Abowd, Matthias Gauger, and Andreas Lachenmann. 2003. The Family Video Archive: An Annotation and Browsing Environment for Home Movies. In *Proceedings of the 5th ACM SIGMM International Workshop on Multimedia Information Retrieval (MIR '03)*. ACM, New York, NY, USA, 1–8. DOI:`http://dx.doi.org/10.1145/973264.973266`

2. Abir Al-Hajri, Matthew Fong, Gregor Miller, and Sidney Fels. 2014. Fast Forward with Your VCR: Visualizing Single-video Viewing Statistics for Navigation and Sharing. In *Proceedings of Graphics Interface 2014 (GI '14)*. Canadian Information Processing Society, Toronto, Ont., Canada, Canada, 123–128. `http://dl.acm.org/citation.cfm?id=2619648.2619669`

3. Abir Al-Hajri, Gregor Miller, Sidney Fels, and Matthew Fong. 2013. Video Navigation with a Personal Viewing History. In *Human-Computer Interaction – INTERACT 2013*. Lecture Notes in Computer Science, Vol. 8119. Springer Berlin Heidelberg, 352–369. DOI:`http://dx.doi.org/10.1007/978-3-642-40477-1_22`

4. Abir Al-Hajri, Gregor Miller, Matthew Fong, and Sidney S. Fels. 2014. Visualization of Personal History for Video Navigation. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 1187–1196. DOI:`http://dx.doi.org/10.1145/2556288.2557106`

5. Jacob Lowell Bishop and Matthew A Verleger. 2013. The flipped classroom: A survey of the research. In *ASEE National Conference Proceedings*.

6. Ruth C. Clark and Richard E. Mayer. 2011. *E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning*. John Wiley and Sons.

7. Bruna C.R. Cunha, Olibário J. Machado Neto, and Maria da Graça Pimentel. 2013. MoViA: A Mobile Video Annotation Tool. In *Proceedings of the 2013 ACM Symposium on Document Engineering (DocEng '13)*. ACM, New York, NY, USA, 219–222. DOI:`http://dx.doi.org/10.1145/2494266.2494267`

8. Pierre Dragicevic, Gonzalo Ramos, Jacobo Bibliowitcz, Derek Nowrouzezahrai, Ravin Balakrishnan, and Karan Singh. 2008. 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, 237–246. DOI:`http://dx.doi.org/10.1145/1357054.1357096`

9. Roberto Fagá, Jr., Vivian Genaro Motti, Renan Gonçalves Cattelan, Cesar Augusto Camillo Teixeira, and Maria da Graça Campos Pimentel. 2010. A Social Approach to Authoring Media Annotations. In *Proceedings of the 10th ACM Symposium on Document Engineering (DocEng '10)*. ACM, New York, NY, USA, 17–26. DOI:`http://dx.doi.org/10.1145/1860559.1860566`

10. Matthew Fong, Abir Al Hajri, Gregor Miller, and Sidney Fels. 2014. Casual Authoring Using a Video Navigation History. In *Proceedings of Graphics Interface 2014 (GI '14)*. Canadian Information Processing Society, Toronto, Ont., Canada, Canada, 109–114. `http://dl.acm.org.ezproxy.library.ubc.ca/citation.cfm?id=2619648.2619667`

11. Andreas Girgensohn, John Boreczky, and Lynn Wilcox. 2001. Keyframe-Based User Interfaces for Digital Video. *Computer* 34, 9 (Sept. 2001), 61–67. DOI:`http://dx.doi.org/10.1109/2.947093`

12. Dan B. Goldman, Chris Gonterman, Brian Curless, David Salesin, and Steven M. Seitz. 2008. 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, 3–12. DOI:`http://dx.doi.org/10.1145/1449715.1449719`

13. W. Hurst, G. Gotz, and T. Lauer. 2004. New methods for visual information seeking through video browsing. In *Information Visualisation, 2004. IV 2004. Proceedings. Eighth International Conference on*. 450–455. DOI:`http://dx.doi.org/10.1109/IV.2004.1320183`

14. W. Hurst and P. Jarvers. 2005. Interactive, dynamic video browsing with the zoomslider interface. In *Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on*. 4 pp.–. DOI:`http://dx.doi.org/10.1109/ICME.2005.1521484`

15. Thorsten Karrer, Malte Weiss, Eric Lee, and Jan Borchers. 2008. 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, 247–250. DOI:`http://dx.doi.org/10.1145/1357054.1357097`

16. Juho Kim, Elena L. Glassman, Andrés Monroy-Hernández, and Meredith Ringel Morris. 2015. RIMES: Embedding Interactive Multimedia Exercises in Lecture Videos. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 1535–1544. DOI:`http://dx.doi.org/10.1145/2702123.2702186`

17. Juho Kim, Philip J. Guo, Carrie J. Cai, Shang-Wen (Daniel) Li, Krzysztof Z. Gajos, and Robert C. Miller. 2014. Data-driven Interaction Techniques for Improving Navigation of Educational Videos. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14)*. ACM, New York, NY, USA, 563–572. DOI:http://dx.doi.org/10.1145/2642918.2647389

18. Don Kimber, Tony Dunnigan, Andreas Girgensohn, Frank Shipman, Thea Turner, and Tao Yang. 2007. Trailblazing: Video Playback Control by Direct Object Manipulation. In *International Conference on Multimedia and Expo*. IEEE, 1015–1018. DOI:http://dx.doi.org/10.1109/ICME.2007.4284825

19. Detlev Leutner, Claudia Leopold, and Viola den Elzen-Rump. 2007. Self-regulated learning with a text-highlighting strategy. 215(3) (2007), 174–182.

20. Justin Matejka, Tovi Grossman, and George Fitzmaurice. 2013. Swifter: Improved Online Video Scrubbing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 1159–1168. DOI:http://dx.doi.org/10.1145/2470654.2466149

21. Robert Mertens, Holger Schneider, Olaf Müller, and Oliver Vornberger. 2004. Hypermedia navigation concepts for lecture recordings. In *World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*, Vol. 2004. 2840–2847.

22. Toni-Jan Keith Palma Monserrat, Shengdong Zhao, Yawen Li, and Xiang Cao. 2014. L.IVE: An Integrated Interactive Video-based Learning Environment. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems (CHI EA '14)*. ACM, New York, NY, USA, 185–186. DOI:http://dx.doi.org/10.1145/2559206.2579525

23. Klimis Ntalianis, Anastasios Doulamis, Nicolas Tsapatsoulis, and Nikolaos Doulamis. 2008. Human Action Analysis, Annotation and Modeling in Video Streams Based on Implicit User Interaction. In *Proceedings of the 1st ACM Workshop on Analysis and Retrieval of Events/Actions and Workflows in Video Streams (AREA '08)*. ACM, New York, NY, USA, 65–72. DOI:http://dx.doi.org/10.1145/1463542.1463554

24. Amy Pavel, Colorado Reed, Björn Hartmann, and Maneesh Agrawala. 2014. Video Digests: A Browsable, Skimmable Format for Informational Lecture Videos. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14)*. ACM, New York, NY, USA, 573–582. DOI:http://dx.doi.org/10.1145/2642918.2647400

25. Gonzalo Ramos and Ravin Balakrishnan. 2003. Fluid Interaction Techniques for the Control and Annotation of Digital Video. In *Proceedings of the 16th Annual ACM Symposium on User Interface Software and Technology (UIST '03)*. ACM, New York, NY, USA, 105–114. DOI:http://dx.doi.org/10.1145/964696.964708

26. Michael Riegler, Mathias Lux, Vincent Charvillat, Axel Carlier, Raynor Vliegendhart, and Martha Larson. 2014. VideoJot: A Multifunctional Video Annotation Tool. In *Proceedings of International Conference on Multimedia Retrieval (ICMR '14)*. ACM, New York, NY, USA, Article 534, 4 pages. DOI:http://dx.doi.org/10.1145/2578726.2582621

27. João Silva, Diogo Cabral, Carla Fernandes, and Nuno Correia. 2012. Real-time Annotation of Video Objects on Tablet Computers. In *Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia (MUM '12)*. ACM, New York, NY, USA, Article 19, 9 pages. DOI:http://dx.doi.org/10.1145/2406367.2406391

28. A. M. F. Yousef, M. A. Chatti, U. Schroeder, M. Wosnitza, and H. Jakobs. 2014. MOOCs - A Review of the State-of-the-Art. In *Proceedings of the Computer Supported Education Conference*, Vol. 3. 9–20.