

Designing Affordances for the Navigation of Detail-on-Demand Hypervideo

Andreas Girgensohn¹, Lynn Wilcox¹, Frank Shipman², and Sara Bly³

¹FX Palo Alto Laboratory
3400 Hillview Avenue, Bldg. 4
Palo Alto, CA 94304, USA

²Texas A&M University
Dept. of Computer Science
College Station, TX 77843, USA

³Sara Bly Consulting
24511 NW Moreland Road
North Plains, OR 97133, USA

{andreasg, wilcox}@fxpal.com

shipman@cs.tamu.edu

sara_bly@acm.org

ABSTRACT

We introduced detail-on-demand video as a simple type of hypervideo that allows users to watch short video segments and to follow hyperlinks to see additional detail. Such video lets users quickly access desired information without having to view the entire contents linearly. A challenge for presenting this type of video is to provide users with the appropriate affordances to understand the hypervideo structure and to navigate it effectively. Another challenge is to give authors tools that allow them to create good detail-on-demand video. Guided by user feedback, we iterated designs for a detail-on-demand video player. We also conducted two user studies to gain insight into people's understanding of hypervideo and to improve the user interface. We found that the interface design was tightly coupled to understanding hypervideo structure and that different designs greatly affected what parts of the video people accessed. The studies also suggested new guidelines for hypervideo authoring.

Categories and Subject Descriptors

H.5.1. Information interfaces and presentation: Multimedia information systems – *video*.

H.5.4. Information interfaces and presentation: Hypertext/hypertext – *navigation, user issues*.

General Terms

Design, Human Factors.

Keywords

Hypervideo, link navigation, video keyframes, iterative design, user studies.

1. INTRODUCTION

Finding the desired information in video takes time. Typically, this is because video must be viewed linearly and tools for searching are limited to fast forward and rewind. Aids such as chapter boundaries and scene indices in DVDs only provide coarse access to video segments. One approach for providing more rapid access is to support skimming via shorter versions of the videos [2, 7, 16]. Another approach is to support access to pieces of video via keyframes [18].

We are investigating a simple form of hypervideo to support rapid information location within an interrelated video. *Detail-on-demand* video only supports one link at any given time rather than allowing for multiple link anchors in different areas of a video frame. It was designed to allow hypervideo navigation without requiring a complex user interface.

Detail-on-demand video provides a natural mechanism for authoring instructional videos that allow viewers to quickly find the information they need. In this format, the main video stream provides an outline summary of the topics. For each topic, a summary of the sub-topics is provided, with links to more detail. At the lowest level is the contents of the entire video. Compared to current linear access instructional videos, this saves the viewer's time and makes it possible for video guides to be appropriate to a wider audience. It also improves on the typical DVD chapter mechanism which has just one level of detail and does not support jumps to related information during playback.

Previously, we reported on a user interface for authoring detail-on-demand video and on using it as a representation for video summaries [13, 14]. In this paper, we focus on the questions of whether viewers can effectively make use of this form of hypervideo and how a player for such video should be designed. Initial use of our prototype system revealed several potential problems. Some users commented that the interface was not intuitive. Others did not totally understand the structure of hypervideo. And we ourselves were concerned with how to author this type of material for training and other applications. These concerns led us to further investigate the nature of hypervideo through user studies and iterative design.

To answer these questions, we built a prototype interface and had two expert users evaluate it. We then conducted a user study with a detail-on-demand version of a home improvement video on plumbing. The study suggested modifications in the user interface, both in how to interact with the video and how to expose the structure of the hypervideo to the user. After making these changes, we conducted a second user study to verify our design. Although the interface was improved, we still found problems with interactions with hypervideo due to deeper issues.

One crucial issue is how viewers make sense of the hypervideo structure. Through user studies, we have learned that exposing the structure is much more important in hypervideo than in hypertext. The idiosyncratic linking seen among text documents found on the Web is more likely to leave users "lost in hyperspace" in a time-based media like video. Although not all our questions have been answered, we have gained much insight into authoring detail-on-demand video and its potential.

In the next section, we present the concept of detail-on-demand video and its relationship to other efforts. We then briefly describe the Hyper-Hitchcock authoring environment and walk through the

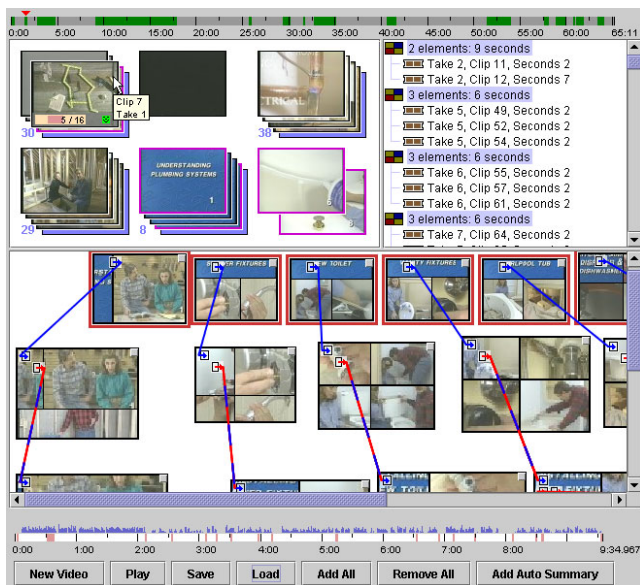


Figure 1: Hyper-Hitchcock Detail-on-demand Video Editor

evolution of the hypervideo player. After a description of the user studies, we conclude with a discussion of lessons learned and future directions.

2. DETAIL-ON-DEMAND VIDEO

Hypervideo allows viewers to navigate between video chunks. General hypervideo allows multiple simultaneous link anchors on the screen, e.g., links from actors on the screen to their biographies [4, 5, 15]. We have concentrated on a simpler form of hypervideo, *detail-on-demand* video, where at most one link is attached to a video sequence. Our notion of detail-on-demand video has been influenced by interactive video that allows people viewing the video to make choices that impact what video they see. Viewing detail-on-demand video combines interaction characteristics from browsing the Web and changing channels on TV.

As the viewer watches a video, the player indicates when links are available and presents labels for them. After following a link, the link destination video plays until completion. Links can be set to automatically return to the jump-off point after the linked video sequence finishes playing. We also allow authors to choose other link return behaviors such as returning to the beginning or the end of the video segment from which a link originated, or stopping. Links can have an offset into the linked video sequence such that the playback starts in the middle of the linked sequence. This supports reuse of video material for different links and allows the viewer to skip back in the linked sequence to get more context.

Many interfaces for accessing video make use of an explicit or inferred hierarchy for selecting a starting point from which to play the video. These vary from the standard scene selection on DVDs to selection from hierarchically structured keyframes or text outlines in a separate window [3, 6, 8, 9, 10]. A difference between interfaces supporting hierarchical access to video and detail-on-demand video is the detail-on-demand viewer may request additional detail while watching the video rather than having to use a separate interface such as keyframes or a tree view. Also, the hierarchical video representation of these tools does not include semantics beyond simple hierarchical composition. Links in detail-on-demand video have labels and a variety of behaviors for when the link's destination anchor finishes playback or when the user interrupts playback. Other hypervideo systems [11, 12, 17] con-



Figure 2: Player with Link Traversal History

centrate on letting the viewer select a story line or on providing context [1] rather than on letting the viewer interactively choose the amount of detail of the presented information.

We previously discussed several applications for detail-on-demand video [13, 14]. In this paper, we focus on the user experience in viewing such video.

3. HYPER-HITCHCOCK

Authoring detail-on-demand video can be supported in a direct manipulation video editor rather than requiring scripting languages or other complicated tools that are unsuitable for a broad user base. Hyper-Hitchcock is a direct manipulation environment for authoring and viewing detail-on-demand video [13]. It presents clips from a video in a selection area (see Figure 1). Authors can drag clips from the selection area into the workspace. There, the length of the segment of clip to be included in the edited video may be changed by resizing the image of the clip. Clips may be grouped into video composites, which in turn may be part of other (higher level) composites. Composites are visualized as collages of up to four keyframes for clips included in the composite. The workspace visualizes links going into and out of a composite.

Because of problems with the audio track in the version of the plumbing video used in the heuristic evaluation, we now determine the audio energy of the video and added a visualization for it above the timeline of the workspace. For a selected video clip, the in/out points can be adjusted easily in the timeline such that they line up with a long enough silence for a sentence boundary.

4. EVOLUTION OF THE HYPERVIDEO PLAYER

We created several different versions of the hypervideo player in response to user feedback, a heuristic evaluation, and two user studies. In a later section, we report in detail on the study results with the different player designs. In this section, we describe the designs and rationales for changes from one design to the next.

4.1 First Version of Hypervideo Player

The first version of the hypervideo player evolved from our player for linear video. Figure 3 includes close-ups of the interaction points for viewers of detail-on-demand video. Below the video playback region is the label for the link followed to get to the currently playing video. In Figure 3, that link is labeled "overview". Beneath that label are control buttons, a label for the currently available link ("connecting copper pipes"), a video timeline slider indicating the current point in the playback and the boundaries of links in the currently playing video, and the time into the currently playing video. The control buttons (from left to right) are for start-



Figure 3: First Design of Hypervideo Player

ing and stopping the video, following links, returning from links, and viewing external information in a Web browser.

When the player follows or returns from a link, a two second video icon is played to indicate the transition. The video icons for following, aborting, and finishing a link have distinct audio tracks¹ but the same video content. After following the link in Figure 3, the main label would be replaced with “connecting copper pipes” and a new label would be displayed if a new link became available. The video timeline slider shows the length of the video sequence currently being played and the current position in that playback. It also shows when links will be active during that playback as blue shaded regions in the slider. The boundaries between links are indicated by darker blue lines in the slider. In the video playing in Figure 3, there are three periods without any link and three periods with links.

We discovered several issues with this player from user feedback and in a heuristic evaluation. First, only information about the currently active link is displayed. To get information about other links, viewers have to skip to the corresponding part of the video. Also, the link label should be clickable. Second, the short text labels were viewed as insufficient to describe the contents of the link destination. Finally, after following a link there is no information about the current place in the link hierarchy or about the place returning from the link would take the viewer. We designed several new player interfaces to address these issues.

4.2 Visualizing Link History and Destination with Keyframes

Early user feedback indicated that users would like to see where a return from a link would take them. We decided to visualize both the link destination and the jump-off point of a followed link with a keyframe, as shown in Figure 4. Both keyframes are mouse-sensitive and can be used to follow or return from a link, respectively. When following a link, the change is animated in the display by shrinking the current frame of the video display and moving it towards the left while at the same time enlarging the keyframe of the link destination and moving it towards the video display area. This animation replaces the video icon transitions used before and was very well received. User observations also indicated the need to allow users to view link labels without jumping to that place in the video, so we added link labels as tooltips to the timeline. Unfortunately, technical problems with the platform prevented us from basing further developments on that version of the player.

¹ Some users later referred to this audio as “noises.”



Figure 4: Player with Animated Keyframes

4.3 Displaying all Links in the Timeline

The results from the heuristic evaluation of the system indicated that it is insufficient to just visualize the existence and location of links in the timeline. Users had to skip the video playback to the points of the links in order to know what material the links would provide. The tooltips introduced for the player shown in Figure 4 were insufficient for this purpose. To address this difficulty, we modified the player shown in Figure 3 and put all link labels into the timeline (see Figure 5). In cases where the text of a link label is too long for the corresponding duration of the link, an ellipsis character indicates that the link label is longer and the complete label is shown on mouse over.

4.4 Modified Link History

Because of technical problems with the platform used for the player shown in Figure 4, we created a version of the link history using our original implementation (see Figure 2). In that version, we display a stack of keyframes representing traversal history in the top-left to make the order from the top-left to the bottom-right more obvious. We also scale down the keyframes for the older jump-off points to enhance the history view. As before, all keyframes are clickable, thus enabling the user to return several link levels at once. Unfortunately, this version cannot animate the keyframes into the video display area and uses the video icons to indicate transitions.

4.5 Displaying Keyframes for all Links

The results of the first user study with the interfaces shown in Figures 5 and 2 made us realize that the placement of the keyframe for the active link to the right of the video display was inconvenient for many users. The single keyframe was also insufficient to give viewers an overview of the link structure. We also noticed that some participants in the first study clicked on the link labels in the timeline and expected the player to follow the link and to play the link destination. Instead, clicking with the mouse on a link label in the timeline would skip to that point in the video and make the corresponding link active but not follow the link.

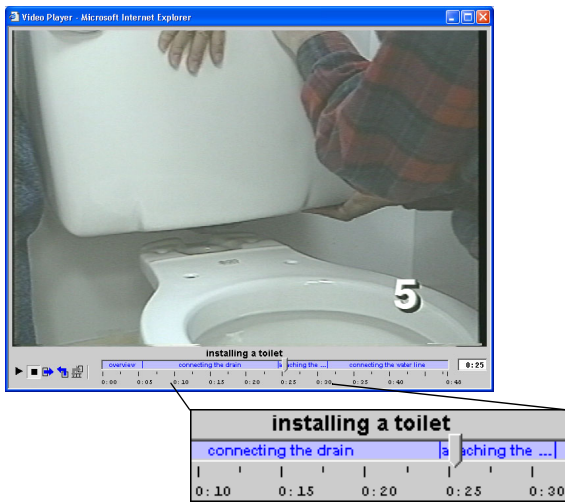


Figure 5: Player with Link Labels in Timeline

To address these issues, we decided to place keyframes for all links along the timeline and to make all links available via keyframes and not just the link attached to the currently playing video segment (see Figure 6). Users can follow the links by clicking on link keyframes or labels without having to first skip the video playback to the appropriate place.

The keyframes for the currently not active links are reduced in size with faded link labels. If the keyframe or the link label are too large to fit in the space for a link, they are cropped. The area of the active link is emphasized in the timeline and separators between links in the timeline are made stronger. Both the keyframe and the label of the active link are expanded if they were cropped because of the lack of space. The keyframe for the link under the mouse is also enlarged and the link is emphasized even more than the active link to indicate that it can be used by clicking.

5. USER STUDIES

The user studies were intended to address how successful people are in navigating detail-on-demand video. We designed the study with two goals in mind: 1) to improve our interface to hypervideo and 2) to gain insight into people's understanding of hypervideo. Our approach was to perform a series of qualitative studies. The heuristic evaluation was used to uncover obvious interface problems and to pilot the user study task. We then conducted a task-oriented lab study designed to test the user interface. A second study was similar to the first, but with emphasis on basic understanding of hypervideo.

5.1 The Task Video

The task-oriented studies were performed using a detail-on-demand hypervideo authored from a linear 65-minute video on home plumbing. Users were asked to navigate the video to find answers to several questions. The hypervideo was structured so that the lower levels contained increasing specific information on a topic. Sometimes, the answer to the question could be found at the higher level, and sometimes the user had to dig down.

The hypervideo was authored using the Hyper-Hitchcock editing tool. Figure 7 shows the structure of the hypervideo as displayed in Hyper-Hitchcock. In this structure, the first level serves mostly as a table of contents. The second level provides brief information about the different topics and allows the viewers to request more detail. The linear nature of the source video meant that once a topic had been discussed in detail earlier in the video, it was only men-

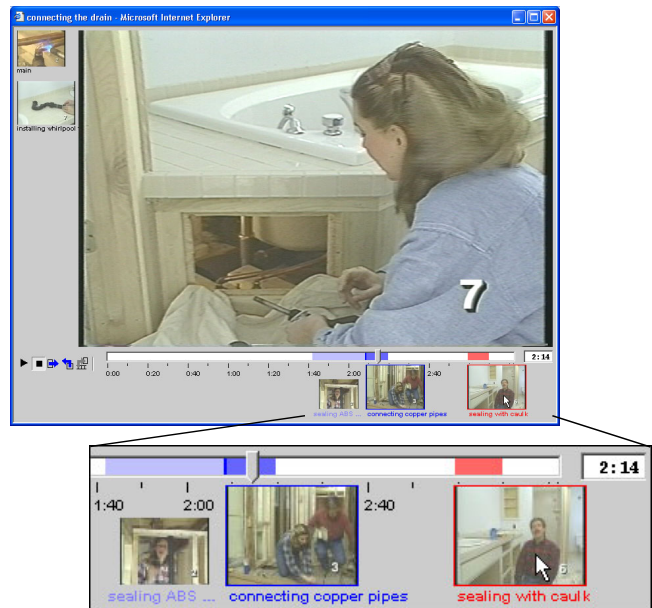


Figure 6: Player with Keyframes Attached to Timeline

tioned in later sections. At the third level, the hypervideo includes additional links from topics in later chapters to earlier parts of the video that provide explanations assumed to already be known. For example, the "installing shower fixtures" section mentions connecting copper pipes without explanation as this was discussed earlier in the video.

The main level (the top row in Figure 7) consisted of six sections totalling 42 seconds of video on topics such as "installing a toilet." Each section contained between 6 and 9 seconds of video. These summary clips were authored from the original video by concatenating several short segments into a video montage. The images in Figure 7 represent these video montages. Each of the top level sections had one link to more detail on the topic.

The second level of the hypervideo contains sub-topics for each of these main topics. For example, under "installing shower fixtures" there are three sub-topics: "overview of installing shower fixtures," "installing shower handles," and "installing shower head." Links from the clips in a sub-topic all go to the same third level video sequence with different offsets into the sequence. The total amount of video at the second level is 5 minutes. Each of the sub-sections is a nested composite with 3 to 18 video clips. The length of the link anchors ranged from 3 to 24 seconds. The third level of the hypervideo contains 27 minutes of video. Unlike the video sequences at the first and second levels that have links everywhere, most of the video at the third level does not have attached links. The two clips in the fourth level contain 80 seconds of video plus 20 seconds that are included in an earlier part of the third level.

5.2 Heuristic Evaluation

We asked two user interface design and evaluation experts to examine two hypervideo player interfaces (see Figures 3 and 4). The first interface visualizes the existence of links in the timeline and displays a link label in a separate location. The second interface provides keyframes for links to follow and for the link history, and puts the link label as a tool tip in the timeline. After the evaluators received basic instructions on using the interface, they were left alone to explore the interfaces and to attempt to answer the study questions with an earlier version of the plumbing hyper-

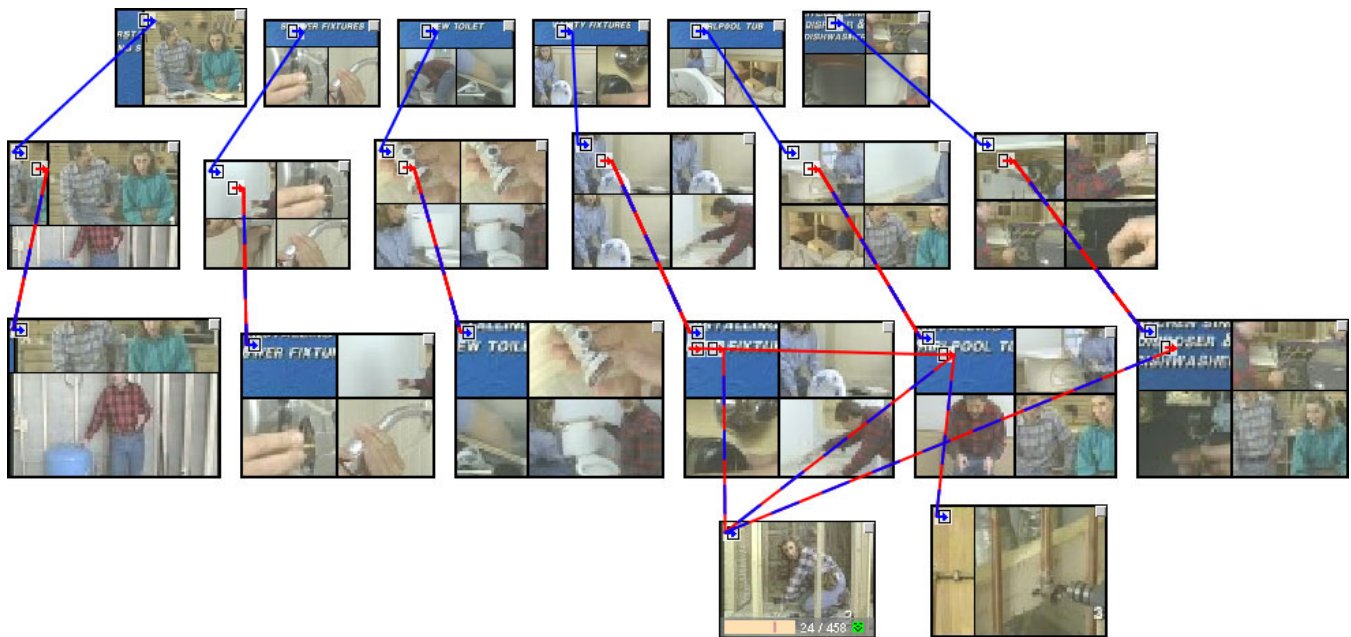


Figure 7: Structure of the Plumbing Hypervideo

video. The experts were asked to comment on the interface in detail.

Both evaluators felt that navigation was possible with the interfaces and that detail-on-demand hypervideo has good potential. They were concerned about three main issues: the difficulty in understanding the video summaries, the difficulty of knowing what material was in the links, and the stability of the system.

Both evaluators commented on the audio choppiness especially in the top-level video caused by the selection of video segments mainly for their visual significance. They felt that complete sentences were necessary. Both said that they found authoring to be critical, perhaps more critical for access to material than the interface itself.

Both saw the value of the tool tip link labels; one frequently used them to orient herself, the other suggested they should appear quickly rather than after a delay. One of the evaluators found herself frequently clicking on the link label in the attempt to follow a link rather than on the link button next to it. They also suggested an index of link labels or a table of contents.

Both evaluators thought the keyframes and animation feedback were helpful for following links. Although it was not clear how to use the keyframes in the history for one evaluator, the other felt the keyframe history let her know where she was in the link structure and that it enabled her to jump back to the top.

As a result of this evaluation, we re-authored parts of the video to improve the quality of the summaries. This involved cutting on sentence boundaries and sometimes removing the audio entirely. We put link labels in the timeline so they could all be seen at a glance. And we decided to continue to experiment with keyframes and keyframe history. As we were unable to make the player depicted in Figure 4 more stable, we added some of its features to the other player and used it as the basis for further studies.

5.3 Task Study 1

The primary purpose of the first study was to improve the interface design. To determine the usefulness of keyframes, we presented participants with two interfaces shown in Figures 5 and 2. Both

interfaces provide buttons for following and returning from links and allow the use of cursor keys for the same purpose. One interface provides clickable keyframes for the active link and the link history. It is notable that the other interface has no link history. Users were given both interfaces in random order for the two halves of the session.

The first task study was performed with six participants aged 18 to 45. In a pre-questionnaire, they were asked about experience with the Web, DVDs, and plumbing. All participants had browsed the Web and navigated to specific chapters of a DVD. Only one had any experience in home plumbing. We asked this user to be sure to answer questions from the video and not using his knowledge of plumbing.

The participants received a brief introduction to detail-on-demand hypervideo, instructions in the interface, a sheet that diagrammed the interface usage, and a list of six questions such as “what type of pipe connects to the water supply system?” Users were then asked to read questions in order and to find answers in the video. Once they heard the answer, they were to tell us and go on to the next question. All sessions were videotaped and logged. Participants were encouraged to talk aloud explaining what they were doing and why. They were also asked to voice their concerns. Afterwards, there was an open-ended interview.

All participants liked keyframes, particularly for their use in displaying the link navigation history. Users given the keyframe interface first disliked when the keyframes were taken away in the second interface. Users given the interface without keyframe first did not use the keyframes as much to navigate when they were switched to the keyframe interface, but commented more about being lost in the hypervideo. Users did not use the keyframe to the right to follow links once they learned to use the follow-link button key because the keyframe is placed too far away from the timeline controlling the video.

Some participants had problems with the readability of link labels in the timeline. If a label was long, they had to move the mouse to expand the label. To have an overview of the available links, they wanted to see all labels at a given level. As in the heuristic evaluation, people confused navigating to a label in the timeline with fol-

lowing the link. Although the absence of the transitional video icon when following the link should have been a clue that the link was not taken, they assumed that because the video jumped, and because the subject of the video was that of the link label, that they had taken the link.

“I didn’t know that there was supposed to be a video effect that goes on after you click a link. If I would have known it could have been easier for me to know that I didn’t actually go somewhere when I was supposed to.”

This caused concern when they needed more information on the topic and felt they had already followed the link. To remedy these problems, in the next interface, link labels were removed from the timeline.

The participants sometimes had problems following links in time, particularly when the video clip for the link was short. By the time they decided to follow the link, it was no longer available because the video had continued playing with the next clip. To follow the link, participants would have to skip back in the video to the missed link and then follow the link.

“I felt — like — this sense of urgency. If it gets to the end I won’t be able to select anything — just seeing that go across almost makes me feel as if I have to do this quick.”

Thus for the next interface, links were made available even when the corresponding video was not playing.

5.4 Task Study 2

The second study was performed with six users aged 22 to 50, with similar experience to those in the first study. They were given the same introduction and the same initial task, so that we could verify the improvements to the player. Afterwards, they were asked two additional questions so that we could explore their understanding of hypervideo after some experience with navigation.

In the interface used for this study, link labels are moved below the timeline and keyframes are provided for all links and not just the active one (see Figure 6). It is possible to follow a link by clicking on the keyframe, even when the video associated with the link is not playing.

People could easily use these keyframes and captions for link navigation. Most people tended to rely on the captions but a few users commented that the keyframes were helpful.

“I’m not sure about the pictures. I’m mostly going by the captions under the pictures. Whatever caption seems like it’s closest to the question being asked is most likely to include the information.”

“I could work with just the captions, but I prefer the images... I could recognize how to get where I was just by the visual.”

People relied heavily on being able to follow a link by clicking on its keyframe even if the video was not positioned for that link, in fact they used these inactive links to navigate over twice as often as the active ones. People rarely used the follow-link button to follow links: 98% of the time they used the keyframes. One problem with uncoupling keyframe links and the video playback is the video position after returning from a link. The video returned to the jump-off point but many participants expected to return to the link anchor.

People liked the keyframe history as it helped them retrace their steps. They used the history keyframes to go back almost twice as

often as they used the link-return button. When they were two and more levels down and used keyframes to go back, they jumped to the main level twice as often as to the previous level.

“I think having this [keyframe history] you always feel like you can retrace your steps. I didn’t feel like I was totally lost. I felt like at a minimum I could go back to main and I kind of remember stuff that was up there.”

People expected the keyframe for a previous level to be fixed instead of showing the point where they took a link. It seems they were using keyframes to recall the overall structure of the video rather than to remember their exact position.

Many people commented that it was distracting to have the video playing when they were looking at keyframes and captions to find a link.

“So it launches straight in again. I want to stop and figure out how to answer the question by looking at the sub-sections.”

Many people stopped playback at this point. Several users offered unsolicited suggestions:

“Why not stop video after you follow a link when there are further link labels. But let it play if you are at the bottom level.”

“You know what would be nice — if you move the mouse from the video then the audio turns off. Cause it’s hard for me to read with the noise going on.”

Link return behavior confused almost everyone. The video was authored such that when the video at a level finished played, it popped back up to the previous level and continued playing there. Even after people had observed this, when we asked what the video would do, they did not know. When asked what they wanted, they said it should stop or go back to the beginning.

Users familiar with video editing used the timeline but others did not. Some felt the timeline was not really necessary for this interface.

“All you really need is to skip between the video segments.”

On the other hand, users familiar with timelines could see how much time there was in a segment and decide whether to listen or to follow a link.

“Okay, that summary is only 4 seconds so I’m assuming it’s not going to tell me anything about that topic so I will dive down.”

6. DISCUSSION

The user studies were intended to explore how successfully people navigate detail-on-demand video. In general it appears that people can recognize and use links. They have no trouble knowing that a link is available and knowing how to follow it. Users appreciate the ability find information quickly.

“There’s more organization to this type of video — if you want to find a specific answer then you could go right to that part and listen to a short summary and dive down for more detail.”

In both studies, on average users completed the questions in 9 minutes, compared to having to watch the entire 27 minutes of video.

“Relative to the prospect of playing a video all the way through, this is definitely an improvement”

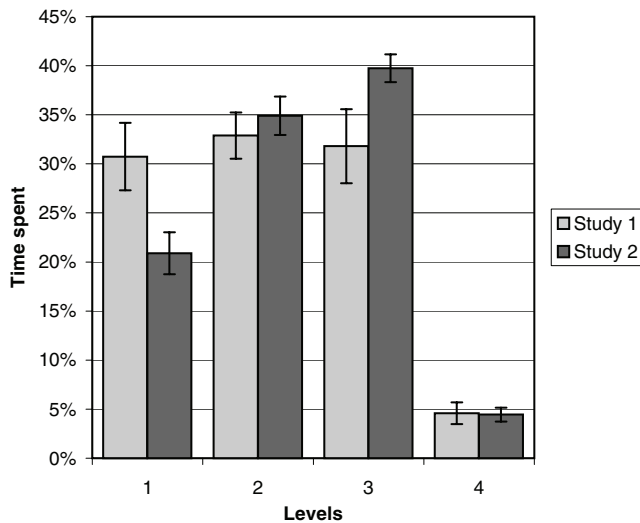


Figure 8: Average Time Spent in Link Hierarchy Levels

Based on the first study, we made several improvements to the interface. However, the second study brought up additional interface problems that can be easily solved. Such changes include having the video return to the link anchor and not to the video that was playing, fixed keyframes for the link anchors in the history, stopping the video after following a link and at the end of a sequence, and the ability to jump forward and backward between links in the timeline.

However, there seems to be a deeper problem related to the interaction between the user interface and people’s understanding of the hypervideo structure. In general, people followed links more quickly in study 2 than in study 1. Figure 8 shows that in the second study (right bars) the participants spent more time at the more detailed levels (Level 2 and Level 3) and less time at the highest level (Level 1).

This can be explained by the additional effort required to follow links in the interfaces used in study 1. Only active links could be followed and the user had to click on a keyframe or link button separated from the timeline. Thus, users tended to spend more time at higher levels and often could find the information there to answer the questions. In study 2, the user could follow a link at any time just by clicking a keyframe. Thus, they tended to navigate to the lowest level before attempting to find the answer in the video. This behavior offset the interface improvements performed between the studies, so that the average time to get answers remained the same.

People in study 2 could have been more efficient if they would have listened at higher levels and found answers. However, even after being reminded of the utility of the higher levels, they continued to quickly follow links to the lowest level. As one person said:

“It isn’t intuitive for me to click on the timeline and think about the summary. I mean, a timeline to me is a fixed space of time and clicking in it for a shortened summary version doesn’t really make sense.”

The disassociation of the video and the links changed the users’ notion from an interactive video interface to a traditional keyframe interface. As one user puts it:

“So I first sort of looked down to see if there’s any keywords that might get me there even before I really started watching the video. It seems as if I’m paying

more attention to the actual interface and looking at keywords for each segment versus actually watching the video. It’s almost as if the video is secondary and I’m actually referring to the descriptions underneath the thumbnails to get where I want.”

While the user interface changes made the interaction simpler, they also caused the users to change their style of interaction such that they navigated the hypervideo structure less efficiently. Further designs need to look more deeply into this issue.

7. CONCLUSIONS

In this paper, we described how we refined the design of a detail-on-demand hypervideo player through several user studies. The results from the user studies are promising with respect to the overall use and benefit of detail-on-demand hypervideo. While the studies uncovered many user interface issues, they also brought up questions about the most suitable form of hypervideo and how to present it to users such that they can understand and use its structure.

Hypervideo, like hypertext, derives its value by allowing users to express their own information needs via navigation. But navigation is only valuable when the viewer can anticipate the outcome of navigation or assess it very rapidly. On the Web, people navigate to and back from an unsatisfying page in a few seconds. With hypervideo, this assessment takes more time. For example, deciding whether an answer was in a summary required listening to the whole summary. As observed in the studies, people often decided to follow links to the lowest level of the video rather than watch the summary.

The implication for hypervideo authoring is that the structure needs to be less idiosyncratic than hypertext structure. Rigorously following a particular design pattern so that the hypervideo structure is predictable may help. In the case of the plumbing video, if the second level video summaries had a common semantic structure (e.g., quick overview of terminology, tools, and tasks), then subjects would have been better able to anticipate whether an answer would be found at that level or required navigation to the bottom level. Alternatively, somewhat longer text descriptions for links could provide a synopsis of the destination but increase authoring effort. Even so, the issue remains of how much detail to put in the summary, and how much to use it simply to guide the user to a longer section of video.

For the user interface, we need to find a better balance between hierarchical keyframes and true hypervideo. One outcome of the iterative design of the player was the move from it being viewed by subjects as interactive video to being a keyframe-based retrieval tool. Once a set of labeled alternatives are provided, users switched to an information retrieval metaphor of interaction. This change in behavior made users miss valuable information from higher level summaries. The user interface needs to present users with an intuitive view of the hypervideo structure. Such a view should be suitable for different tasks and guide the users towards the most appropriate interaction.

8. ACKNOWLEDGEMENTS

We thank Jonathan Helfman for his helpful suggestions on using keyframes in the player interface and for his implementation of the player with animated keyframes. We also thank Candice Kamm for conducting one of the heuristic evaluations and John Adcock, John Boreczky, and Gene Golovchinsky for suggesting improvements to the player interface. Finally, we thank the participants in our user studies.

9. REFERENCES

- [1] Chambel, T. and Guimarães, N. Context Perception in Video-based Hypermedia Spaces. *Hypertext '02 Proceedings*, ACM, New York, pp. 85-94, 2002
- [2] Christel, M.G., Smith, M.A., Taylor, C.R., and Winkler, D.B. Evolving Video Skims into Useful Multimedia Abstractions. *Proceedings of CHI '98*, ACM Press, pp. 171-178, 1998.
- [3] Ding, W., Marchionini, G., and Soergel, D. Multimodal Surrogates for Video Browsing. In *Proceedings of ACM Conference on Digital Libraries*. pp. 85-93, 1999.
- [4] Grigoras, R., Charvillat, V., and Douze, M. Optimizing Hypervideo Navigation Using a Markov Decision Process Approach, in *Proceedings of ACM Multimedia*, ACM Press, pp. 39-48, 2002.
- [5] Hirata, K., Hara, Y., Takano, H., and Kawasaki, S. Content-oriented Integration in Hypermedia Systems, *Hypertext '96 Proceedings*, ACM, New York, pp. 11-21, 1996.
- [6] Li, F., Gupta, A., Sanocki, E., He, L., and Rui, Y. Browsing Digital Video. *Proceedings of CHI '00*, ACM Press, pp. 169-176, 2000.
- [7] Lienhart, R. Dynamic Video Summarization of Home Video, *SPIE 3972: Storage and Retrieval for Media Databases 2000*, pp. 378-389, 2000.
- [8] Mills, M., Cohen, J., and Wong, Y.Y. A Magnifier Tool for Video Data. *Proceedings of CHI '92*, ACM Press, pp. 93-98, 1992.
- [9] Myers, B., Casares, J., Stevens, S., Dabbish, L., Yocum, D., and Corbett, A. A Multi-View Intelligent Editor for Digital Video Libraries, in *Proceedings of the ACM / IEEE Joint Conference on Digital Libraries*, ACM Press, pp. 106-115, 2001.
- [10] Rui, Y., Huang, T.S., and Mehrotra, S. Exploring Video Structure Beyond the Shots. *International Conference on Multimedia Computing and Systems*, pp. 237-240, 1998.
- [11] Rutledge, L. SMIL 2.0: XML for Web Multimedia. *Internet Computing*, Vol. 5, no. 5, pp. 78-84, 2001.
- [12] Sawhney, N., Balcom, D., and Smith, I. Authoring and Navigating Video in Space and Time, *IEEE Multimedia*, 4(4): pp. 30-39, 1997.
- [13] Shipman, F., Girgensohn, A., and Wilcox, L. Hyper-Hitchcock: Towards the Easy Authoring of Interactive Video. In *Human-Computer Interaction INTERACT '03*, IOS Press, pp. 33-40, 2003.
- [14] Shipman, F., Girgensohn, A., and Wilcox, L. Generation of Interactive Multi-Level Video Summaries. In *Proceedings of ACM Multimedia*, ACM Press, pp. 392-401, 2003.
- [15] Smith, J.M., Stotts, D., and Kum, S.-U. An Orthogonal Taxonomy of Hyperlink Anchor Generation in Video Streams Using OvalTime, *Proceedings of ACM Hypertext 2000*, pp. 11-18, 2000.
- [16] Sundaram, H. and Chang, S.-F. Condensing Computable Scenes Using Visual Complexity and Film Syntax Analysis. *Proceedings of ICME 2001*, pp. 389-392, 2001.
- [17] Vardi, G. Navigation Scheme for Interactive Movies with Linear Narrative. *Hypertext '99 Proceedings*, ACM, New York, pp. 85-94, 1999.
- [18] Yeung, M.M. and Yeo, B.-L. Video Visualization for Compact Presentation and Fast Browsing, *IEEE Transactions on Circuits and Systems for Video Technology*. Vol. 7, no. 5, 1997.