Guiding Users through Asynchronous Meeting Content with Hypervideo Playback Plans

Andreas Girgensohn[1], Jennifer Marlow[1], Frank Shipman[2], Lynn Wilcox[1]

[1]FX Palo Alto Laboratory
3174 Porter Drive
Palo Alto, CA 94304
{andreasg,marlow,wilcox}@fxpal.com

[2]Department of Computer Science
Texas A&M University
College Station, TX 77843-3112
shipman@cs.tamu.edu

ABSTRACT
We previously created the HyperMeeting system to support a chain of geographically and temporally distributed meetings in the form of a hypervideo. This paper focuses on playback plans that guide users through the recorded meeting content by automatically following available hyperlinks. Our system generates playback plans based on users' interests or prior meeting attendance and presents a dialog that lets users select the most appropriate plan. Prior experience with playback plans revealed users' confusion with automatic link following within a sequence of meetings. To address this issue, we designed three timeline visualizations of playback plans. A user study comparing the timeline designs indicated that different visualizations are preferred for different tasks, making switching among them important. The study also provided insights that will guide research of personalized hypervideo, both inside and outside a meeting context.

Author Keywords
Hypervideo; videoconferencing; interactive video; visualization.

1. INTRODUCTION
As the prevalence of distributed work teams increases, the volume of video meetings also grows. This increases the likelihood that not all participants will be able to synchronously converse. Currently, asynchronous review of video-recorded meetings can be a tedious process of going through large sections of irrelevant or old content in order to find a few key relevant points (which may be different from person to person). We identified a series of different meeting-review scenarios and developed an approach for proactively and intelligently inferring users' information-seeking goals. The concepts and design alternatives presented here provide an initial step towards understanding the potential benefits and challenges of enabling personalized information retrieval of relevant meeting details.

We make use of hypervideos that are a form of interactive media with links between a series of videos. With more videos being linked, it becomes difficult to navigate between the content and to keep a coherent awareness of one's place in the flow. If the hypervideo system knows about the user's motivations or information seeking goals, it can anticipate valuable paths through the network of linked videos. Such personalization is particularly valuable for users acquiring knowledge. There are many hypervideo applications aimed at knowledge acquisition, including education and training [5][27] and corporate memory [16].

Our research applies the concept of hypervideos to a video conferencing system that links conceptually related discussions created in a chain of geographically and temporally distributed meetings. Attendees may not attend all meetings, particularly when based in different locations and time zones. The HyperMeeting system introduces a new paradigm for enabling and supporting a sequence of videoconferences involving attendees with differing availability, locations, and participation. It lets attendees record and review a series of videoconference meetings with asynchronous attendance. Attendees of a later meeting simultaneously review video recordings of earlier meetings and add comments. During the course of watching a prior meeting, viewers can pause playback to have a discussion among those present. This creates a hyperlink from the paused video to the video being recorded in the follow-up meeting. In subsequent meetings, further discussions may be added to the hyperlinked meeting content, creating a meeting chain or hypermeeting (see Figure 1).

Viewing the meeting content may be a purely manual interaction where users select hyperlinks to be followed. To address the interaction challenge with complex, branching video topics and threads, we created the concept of playback plans to represent automatic navigation paths through hypervideo. For example, playback plans can help find key action items without having to review all previous meetings [20]. Also, we incorporate novel approaches aimed at improving user understanding of and control over playback plans. This includes a dialog for selecting the most appropriate playback plan, three alternate timeline visualizations of playback plans, transitions between playback plan visualizations, and an interactive recorded history of user interaction and navigation caused by playback plans. Our results inform the visualization of paths through hypervideo more generally.

A prior publication describes the core HyperMeeting system and an early instantiation of playback plans [8]. This paper presents extensions to the capabilities of playback plans, an interface for

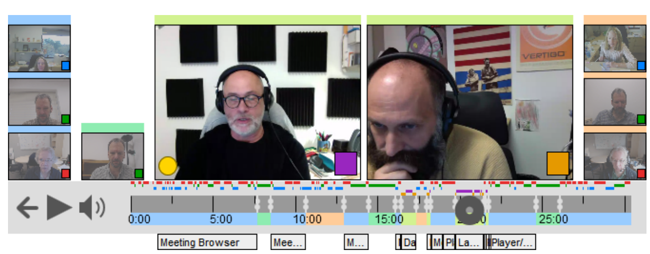

Figure 1. HyperMeeting with Four Meetings. Meeting Colors above the Videos Are Shown in the Timeline. Video of the Selected Meeting Is Larger.

HT '16, July 10-13, 2016, Halifax, NS, Canada

DOI: http://dx.doi.org/10.1145/2914586.2914597

users to select and tailor playback plans, and an evaluation of different UI design choices aimed at improving user understanding of and control over playback plans. The paper examines and discusses the technical and user-experience aspects of these new contributions.

The next section discusses related work. This is followed by a brief description of the HyperMeeting system. Then, playback plans are introduced including a description of how to generate, visualize, and interact with them. A user study of three timeline visualizations of playback plans is then presented. Finally, we discuss lessons learned and next steps in our research.

2. Related Work

Hypervideo consists of video segments that are connected with links. In the course of watching a video, users navigate to related videos through manually authored or automatically generated links. Early applications of hypervideo include tourism or sightseeing (e.g., Aspen Movie Map [15]), moving between recordings of semi-independent conversations (e.g., HyperCafe [24]), news [2], and education [4]. There are many hypervideo applications aimed at knowledge acquisition, including education and training [5][27] and corporate memory [16]. We focus on hypervideo where links connect video segments rather than media objects (e.g., web pages on topics in the video.) Also, our application focus is teleconference meetings.

Prior work on hypervideo (e.g., [24][25]) has focused on authored content where a person intentionally creates links between video segments. A video-recorded meeting could be taken as the starting point but prior systems lack the ability to identify topics and relations among topics as a way of authoring or generating the hypervideo. Also, they do not include features to support the meeting-specific aspects of the video (e.g., attachment of metadata indicating attendees, topics, and time/date indexing). Hypermedia models such as CHM [23] represent relationships between content elements and definition of presentation arrangements. Such models could be used to represent speaker, topic, and link information. However, none of these systems support geographically distributed synchronous viewing of hypervideos. There has been work to optimize prefetching of video content in anticipation of user-driven navigation [11].

The hyperlinks used in HyperMeeting are similar to those in previous work on detail-on-demand video [10]. Those links are attached to video segments and not to objects depicted in the video. Another similarity is that playback returns automatically to the source link anchor after the linked content is played. A central distinction between the work discussed here and prior research on hypervideo is that prior video players require user interaction to follow a link. The playback plans described in this paper identify the links to be followed automatically. User interests determine which links are followed and which links are ignored when the video player reaches their anchors.

While prior hypervideo research has not addressed distributed meetings, there are a number of teleconference systems that have. They employ a variety of approaches to address the issue of allowing people to asynchronously revisit and review content from discussions or meetings. One class of systems focuses on capturing the content of meetings for later consumption [6][7][12][19]. Such systems record content for later playback and browsing. They take advantage of natural activity during meetings (e.g., slide presentations or activities on an electronic whiteboard) to generate navigation points back into the record. These systems allow users to review and find information in a meeting but do not incorporate the ability to extend a previous meeting or create links between meetings.

A second class of systems that have supported asynchronous meetings, e.g., Time Travel Proxy [26] and Video Threads [1], include short videos on topics with responses. The model of threaded messages loses the connections between topics and the larger activity context. In particular, each recorded video response is attached to a single prior video discussion. While the work on Time Travel Proxy and Video Threads supports video annotations to prior content and assumes a hierarchy of video recordings and responses, our approach interlinks full meetings with longer recordings with multiple links among pairs of recordings.

A last class of related systems support the synchronous viewing of video content on distributed devices [3][21]. These systems can include additional channels of communication (e.g., chat) for simultaneous viewers to discuss what they are watching. However, the records of communication are, at best, attached as metadata to the video content. Attached metadata related to the synchronous viewing of prior meetings is less powerful than hyperlinking between meetings.

3. The HyperMeeting System

The HyperMeeting system specializes hypervideo capabilities to support distributed meetings. The web-based HyperMeeting user interface is shown in Figure 2. It includes a live video conference and a player for a previously recorded meeting chain. In the depicted meeting chain, a fourth meeting with two participants is currently in progress.

A hypermeeting starts with a video conference where the video of each meeting attendee is recorded. During the meeting, attendees may enter topics and manually mark when a topic starts or ends.

The hypermeeting is continued in a second video conference where the attendees play the video recorded in the original meeting. Playback is synchronized among the recorded streams and across attendees such that all recorded streams play at the same time and all attendees see the same content. Each attendee can control the video playback for the current participants.

When the attendees of the second meeting continue the discussion started in the original meeting, they pause video playback to discuss their response. The action of pausing the playback creates a hyperlink from the first video to the second video.

When reviewing the previous two meetings in a third meeting, attendees may use the hyperlinks to navigate the recorded content. Links may be followed by clicking on them. Following such links allows the viewer to watch the original discussion together with responses from later meeting. Such a viewing pattern provides more continuity for the conversation compared to viewing the meetings in sequence.

Currently, we focus on distributed meetings conducted as a video conference. Separate video recordings of each attendee allow for simple speaker segmentation. As a result, multiple video streams are played together when the recorded meeting is being watched at a later time.

The HyperMeeting client runs in standard web browsers as a single page application. The live video conference among the distributed meeting attendees is maintained using WebRTC (Web Real-Time

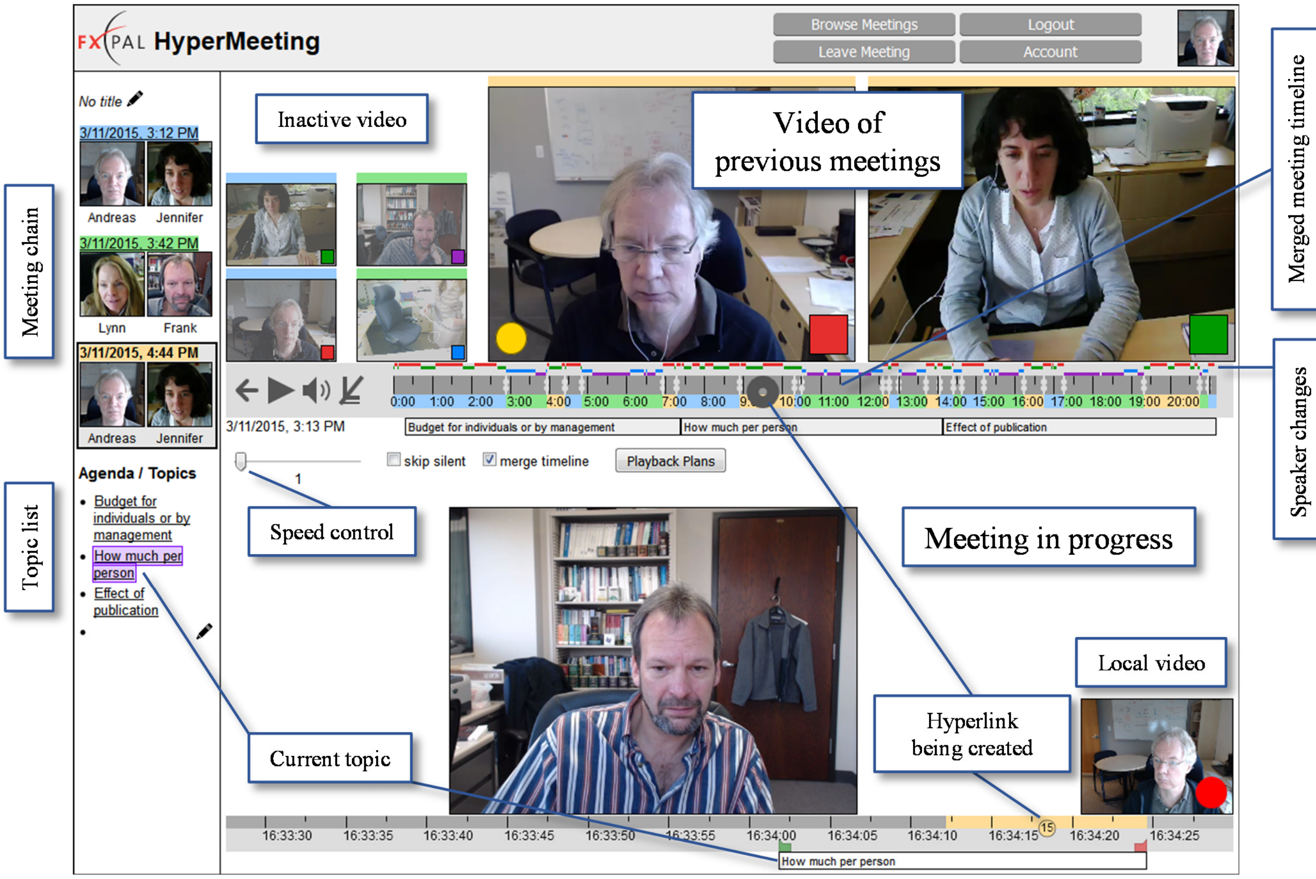

Figure 2. HyperMeeting User Interface with Video Players at the Top and Live Video Conference at the Bottom.

Communication)[1] peer-to-peer connections. Meetings are visualized by manipulating the web page with JavaScript. Recorded video is displayed in HTML5 video tags and timelines are drawn in canvases.

The list of meetings, attendees, and topics are shown on the left of the player (see Figure 2). At the top is the bank of video players for the individual video streams that make up the hypermeeting. Below these is the timeline visualization that presents a variety of information about the hypermeeting. Timelines and video players are color-coded to indicate the meeting they belong to. Colored circles represent links that use the color of the meeting they point to. Links are numbered in creation order with the two ends having the same number. Colored regions around link circles indicate link anchors that would be played when traversing links. Video players of non-active meetings are paused and reduced in size.

4. Automatic Navigation via Playback Plans

The focus of this paper is a study of different designs for timeline visualization and the effect they have on user disorientation and comprehension of the video content. Before we discuss the different timeline designs, we present the concept of playback plans that these visualizations are supporting.

While links can be followed manually, it is often convenient to automatically follow relevant links. The system can determine paths through the hypervideo that cover relevant information for the viewers based on their attendance of prior meetings. We developed the concept of *playback plans* to describe different paths through a hypervideo that are automatically followed during playback. Playback plans for hypermeetings are meant to follow the discussion distributed over multiple recorded meetings by presenting related content together. Additionally, playback plans can be personalized to reduce the quantity of content that is re-watched by individuals.

A playback plan is similar to an edit decision list for video editing. It consists of a sequence of video segments from multiple videos that are played in order. Some plans are a generalization of hypervideo link behaviors [10] – that is they determine what happens at the start or end of a source or destination link anchor during playback. Unlike edit decision lists, playback plans are computed automatically based on previous behavior of the user, available hyperlinks, and specified filter conditions. Filters can identify parts of the videos to include, for example, particular topics, speakers, or non-silent portions.

Different user actions, such as selecting a meeting or a particular link, result in different playback plans. Personalized playback plans can be based on whether the user took part in prior meetings. Consider the case of three meetings shown in Figure 3. An initial meeting is recorded and then additional discussions occur during Meeting 2 and are attached to the initial meeting via two links (Links A and B). A final meeting includes two links connecting the discussions to Meeting 2 (Links C and D). In the case of a user who participated in none of the meetings, it would make sense to play

[1] http://www.w3.org/TR/webrtc/

the video of the first meeting and to automatically follow all links to later meetings such that comments could be viewed in context. Such a plan would perform a depth-first traversal on all of the discussion attached to the originally selected meeting. After Segment 2, playback would visit Segments 7, 8, 15, and 9 before continuing with Segment 3. For participants in earlier meetings, an alternative playback would avoid watching recordings of meetings that they were at.

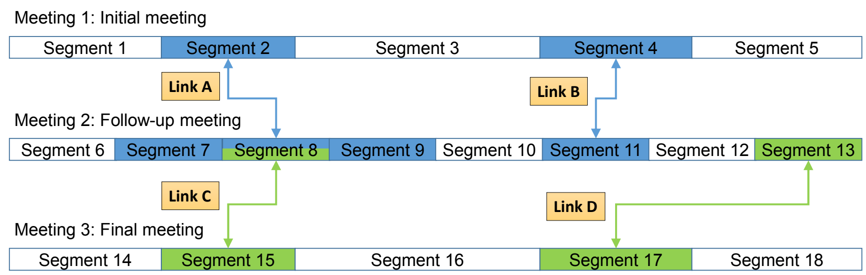

Figure 3. Three video meetings interconnected by hyperlinks.

We identified a set of playback plans that are appropriate to particular use scenarios. The plans are divided into two classes: when selecting a whole meeting and when selecting a hyperlink. Playback plans can be used to filter recorded content and are interactive (i.e., users can navigate within a playback plan). For playing one meeting within a hypermeeting, a plan could (a) just play that meeting without following links, (b) follow only links to later meetings, either the next level or all levels, (c) follow only links to earlier meetings, or (d) follow all links. Plans can also be filtered based on who is speaking in each segment or which topics are covered.

In a meeting chain, each meeting may consist of multiple video streams depicting meeting attendees. Those video streams use time-synchronized playback. When user actions result in the generation of a new playback plan, the new plan is distributed to all meeting attendees so that their video playback can be synchronized. The video player of each attendee can follow the same plan independently, requiring only occasional synchronization messages to avoid drift.

4.1 Intelligently Inferring Playback Plans

We conducted informal user studies where remote meeting attendees used different versions of the HyperMeeting system. Those sessions provided guidance for system improvements and helped identify different use scenarios and playback plans appropriate for those scenarios. The plans contain elements from previous work on taxonomies of meeting-search needs (e.g., [22]), but are adapted specifically to the context of asynchronously reviewing a chain of ongoing meetings. Selecting a playback plan takes into account participation in previous meetings. As the system has a record of meeting participation, it can generate all possible playback plans that fit the participation history and present them to the user in order of the likely viewing behavior.

Review all meetings. For a user who has not participated in the meeting chain, one could play the video of the meeting that started the chain and automatically follow all links to subsequent meetings such that those comments can be viewed in context. Such a traversal through the hypervideo would skip the parts of the subsequent meetings that are not linked, presumably because the content in these portions is not related to the older meetings. This plan was described earlier in the context of Figure 3.

Catch-up. For a user who missed a meeting, the missed meeting could be played in the context of the earlier and later meetings. In this case, source anchors of links pointing to the missed meeting would be played before destination anchors in that meeting. Links to subsequent meetings would be followed as before. For example, a viewer who missed Meeting 2 in Figure 3 would see Segments 6, 2, 7, 8, 15, 9, 10, 4, 11, 12, 13, and 17.

Filter. Users may only be interested in particular topics or speakers. Thus, a plan could playback any responses to segments in which the current user was a speaker, first playing the original segment followed by the responses. Similarly, topic-oriented plans can play through segments tagged with particular topics and any content attached to those segments. A common filter can skip silence in the video. Such a filter is useful in situations where meeting attendees are listening to an older meeting without speaking at the same time. Watching the silent part of the meeting would be of little use to the viewer. Because filters may skip passages as short as a few seconds, many video segments could be produced by such filters. Rather than distributing long video segment lists to other meeting participants and storing them in the database, an abstract description of playback plans is used. Each client can compute the video segment list from the abstract description and the currently active filters.

4.2 Selecting Among Playback Plans

Users may select a plan from a set of candidate playback plans. For example, Figure 4 shows a selection based on the user's attendance of meetings. The plan generation uses the earlier discussed patterns and checks which of them are applicable. If different patterns would be instantiated as the same playback plan, only one of the instances is presented.

As the user selects a plan, it is visualized in the timeline. Once the plan is as desired, the user can play the described video. If none of the plans meet the user's needs, they can cancel the plan.

The plan offered first allows the viewer to catch up on the meeting by starting with the earliest meeting that the user did not attend. Whenever a destination link anchor is encountered in that meeting, the source anchor of the link is played first to provide context. If that source anchor contains destination anchors for other links, those links are handled in the same fashion. When a source link anchor is encountered, the corresponding destination link anchor in the later meeting is played next. If it contains links as well, they are followed. Figure 4 shows an example of that plan in combination with filters for skipping silence and for only including the first topic. To review all meetings, users may choose to start with the oldest meeting regardless of attendance.

The effects of the selection choices are visualized in a stacked timeline depicted at the bottom of Figure 4. A red circle indicates the meeting where the playback plan starts. Translucent red overlays mark the parts of the meetings that are included in the playback plan. The selected *catch-up* plan starts with the middle meeting that the user missed and follows encountered hyperlinks to older and newer meetings. In addition, the first topic is selected as a filter and silent passages are skipped. While the start of the green meeting is not silent, it does not have the topic of interest selected so it is skipped. A brief utterance at 2:30 is included. Otherwise, only link anchors inside the first topic are included, with some of them being slightly trimmed due to silence. In each link-anchor pair, the older source anchor is played before the newer destination anchor.

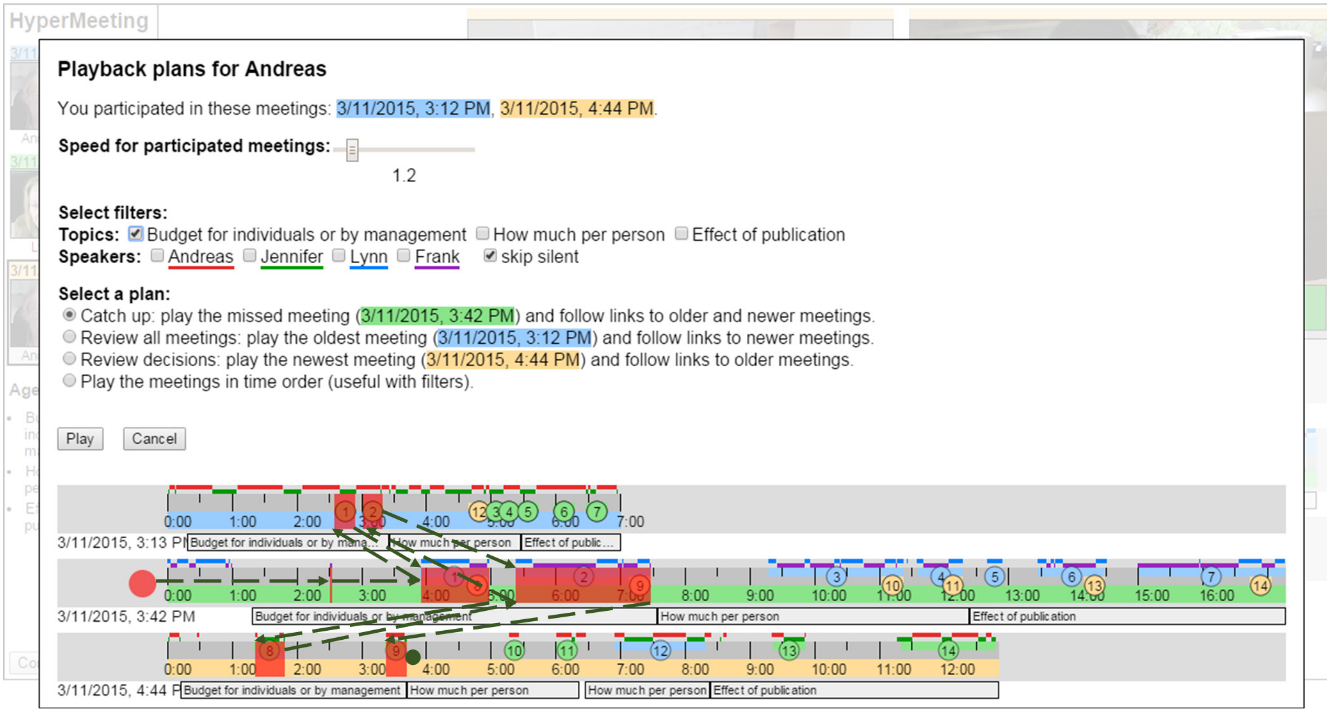

Figure 4. Selected "Catch Up" Playback Plan Combined with Topic and Silence Filters. Red Areas in the Timelines Indicate the Video Included in the Plan Starting at the Red Dot.

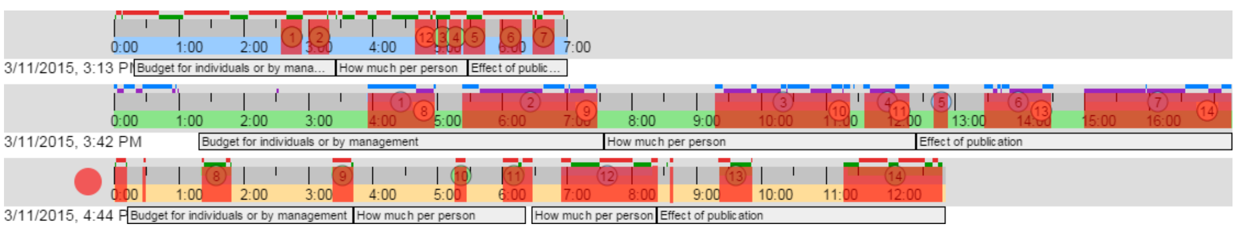

Figure 5. Playback Plan for "Review Decisions" and "Skip Silent" Starting at the Newest Meeting (Red Dot).

Figure 5 shows an example of the *review-decisions* plan for a user who wants to just look at the decisions made in a meeting chain. The plan starts with the newest meeting and plays linked content from older meetings. It is the same as the *catch-up* plan for somebody who missed just the newest meeting. In this example, the plan is combined with a filter for skipping silence.

Link anchors played for context and links to later meetings may refer to meetings that the users attended. Users may play those meetings at a higher speed.

For all meetings in the playback plan, silence may be skipped. Skipping silence is most beneficial when starting with a later meeting in the chain because time in those meetings is spent listening to earlier meetings. Most of the non-silent meeting content may be hyperlinked. However, new topics may be introduced in later meetings that would not have links to earlier meetings.

Users may play only the content that they missed and not follow links leading to meetings they attended. This can be controlled via an age cutoff, e.g., by skipping all known content from the past week.

In addition to selecting the general playback plan strategy, users may indicate that they are only interested in certain topics or people. Check boxes can be used to turn topics or particular speakers on or off. For speakers, one can select to also include adjacent passages from other speakers to provide some context. When selecting only one topic, automatic link following may be less important because the topic provides the context. In such a situation, selecting to play the parts of the meetings tagged with the topic in time order may be more desirable.

Overall, the selection of a playback plan combines automated decisions based on the user's prior participation with options for user adaptation. By associating playback plans with goals and visualizing the plans, HyperMeeting supports user understanding of the effect of their selections.

5. Plan Visualization and Control

In an initial study [8], participants found hyperlinks between meeting discussions to be useful. However, usability issues arose regarding the different playback plans invoked during interaction with a hypermeeting. A common complaint was feeling lost, not knowing where in the meeting chain you were after following hyperlinks. Participants suggested improvements to link behavior such as visualizing the parts of the meeting that had been seen and providing a back button to undo the effect of automatic link following.

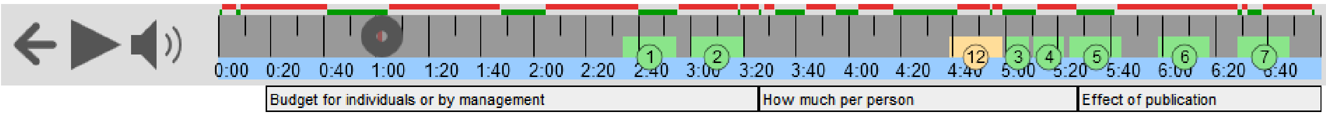

Figure 6. Single timeline showing the oldest meeting of three meetings. Colors above the timeline indicate speaker changes.

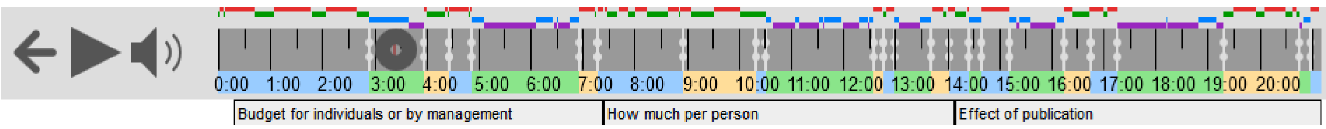

Figure 7. Merged timeline for a playback plan covering oldest meeting and links from it. Colors indicate the meetings.

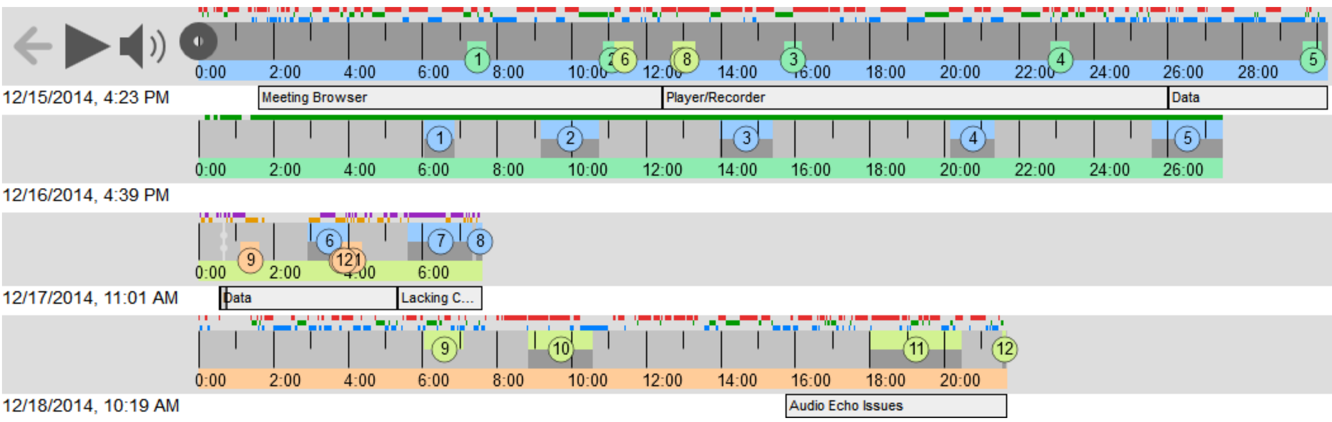

Figure 8. Stacked timelines of four meetings. Dark gray indicates the parts covered by the playback plan.

To improve the use of playback plans, we built on initial findings from [8] to further explore techniques for their visualization, transition, and interaction.

Designing a visualization of a complex set of time-distributed conversations across topics and people can be challenging. There is a tradeoff to be made between providing a great degree of detail (which can be informative but cognitively challenging to interpret) and hiding this detail from the user (which can be easier to visually parse but may obscure useful connections between topics).

No single visualization is likely to be best in all use contexts. Hence, there needs to be a user selectable and automatic way of transitioning between the views of the playback plan. HyperMeeting includes a toggle for users to switch between the views (see checkbox "merge timeline" at the center of Figure 2. HyperMeeting User Interface with Video Players at the Top and Live Video Conference at the Bottom.). Furthermore, double-clicking on the timeline toggles between a merged and non-merged timeline from the other two views.

Prior work in information visualization has explored ways of depicting the flow of events or information over time (e.g., [17][22][29]). However, these visualizations focus on static (non-animated) displays of information. When the information visualized is dynamic and presented in video format, additional cues such as changing speakers, voices, and content being discussed add additional complexity and cognitive load. We focused on three different visualizations of playback plans that revolve around timelines. These timelines show information such as speaker and topic changes and indicate the presence of hyperlinks.

In designing alternate visualizations of the hypermeeting timeline, we incorporated Wang Baldonado et al.'s [28] guidelines for using multiple views for information visualization. These guidelines were initially developed for the display of a complex data set in multiple representations. In the context of HyperMeeting, we apply the design principles to address various tradeoffs in providing multiple views of multiple meetings and helping guide viewers interested in following multiple topics over time. In particular, we incorporated the principles of parsimony, self-evidence, decomposition, consistency, and attention management in different combinations to compare and contrast the relative advantages and disadvantages of each. These principles are elaborated in their application for the respective visualizations that are described in the next subsections.

5.1 Single Timeline

The first visualization, shown in Figure 6, is the standard view of the timeline. This timeline follows the principle of self-evidence [28] and employs perceptual cues (in this case, colored and numbered circles) to make the relationships among the multiple meetings evident. In this view, the hypervideo structure (a set of recorded synchronous meetings and a set of links between segments of these recordings) is presented to the user with one recorded meeting being visualized at a time. After navigating to another meeting, the timeline for that meeting is shown. The playback control to the left of the timeline controls the currently shown meeting. In this view, there is no visualization of the overall playback plan generated by the user's most recent navigation action (e.g., selection of a meeting, link or topic or navigation within the current meeting) as meetings that are not currently being viewed are not visualized.

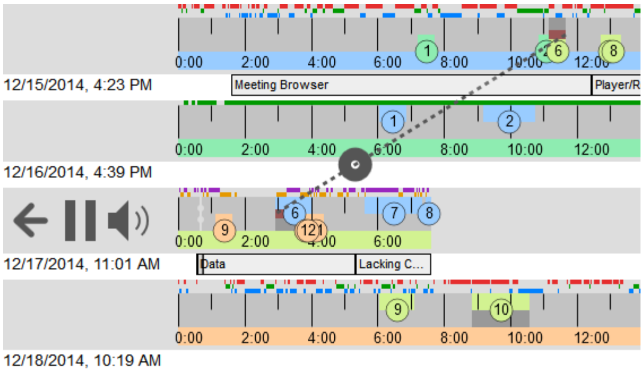

Figure 10. Animation of playback indicator along dashed line during link traversal.

Early users of this timeline design did not have an accurate mental model of the relationships among meetings and topics. Although the single timeline took up less display space, it presented challenges with context switching as the meeting shifted.

5.2 Merged Timeline

To improve users' expectations of what they are going to see, we developed the merged timeline shown in Figure 7. Like the single timeline, the merged timeline is also parsimonious and consistent by using a single view to provide a user with a stable context for analysis [28]. Specifically, the merged timeline does not change as the meeting progresses. This design uses the attention management principle by showing only the relevant parts of the timelines. A benefit of such a design is that the need for display space is reduced. However, a potential drawback is usability issues with memory, comparison, and context switching. This timeline concatenates the portions of the recorded meetings that will be shown based on the current playback plan. Links are not visualized because the automatic link traversal is implied in the playback plan. Links outside the playback plan cannot be selected in this view. Color is used to indicate which meeting will be shown when. Colors are also merged and standardized for the meeting attendees so the same person has the same color across meetings when presenting speaker segmentation information along with the merged timeline.

It has been suggested to us to create a variant of the merged timeline in form of a Gantt chart[2] that would make it look similar to the stacked timeline described in the next section. While such a variant would introduce redundant coding for the meeting a video segment belongs to, it adds no other advantage and takes up more space. We chose not to test this variant in the interest of time.

5.3 Stacked Timeline

A weakness of the merged timeline is that it does not provide information about the material not included in the current playback plan. Hence, users do not know how much of the content they are seeing relative to the amount available and the relative positions of the segments in the meetings. Thus, we designed a third visualization to display this content. It does so at the cost of providing multiple timelines simultaneously rather than a single timeline as in the two previous visualizations. By using the rule of decomposition to visually split the individual meeting timelines from each other, the problems of cognitive overload of a single complex view could be reduced. Multiple views can help the user to "divide and conquer" by reducing the amount of data they need to process at one time [28]. The stacked timeline shows the hypervideo structure (as in the single timeline) but also shows the segments of all meetings that are part of the current playback plan (as in the merged timeline). The coverage of the playback plan is indicated by a darker shade of gray. Figure 8 uses a chain of four meetings instead of the earlier used chain of three meetings to better illustrate the concept of the stacked timeline. The three-part stacked timeline can be seen in Figure 4.

Unlike the other two timeline designs, the user's attention has to shift when playback moves to another meeting. In the stack of timelines shown in Figure 8, the playback control moves to the timeline of the meeting currently being played. While following a link, the playback indicator is animated in its path from one meeting to the other (see Figure 10). This animation was designed to use the rule of attention management to help ensure that the user focuses on the right view at the right time [28].

6. Evaluation Setup and Participants

A study comparing the three timeline visualizations was performed to assess their relative impact on user disorientation, comprehension, and satisfaction when watching an automatic meeting playback. We used the *review-all-meetings* playback plan in the study because that would be the plan used most frequently by viewers who did not attend any of the meetings. Watching previous meetings as an observer is a common playback plan use case, especially for viewers who did not attend those meetings. The plan starts at the oldest meeting and follows all links to newer meetings. The study was a within-subjects design where each participant was exposed to the three different timeline visualizations and provided comments on the advantages and disadvantages of each.

6.1 Method

This study was conducted using an Amazon Mechanical Turk Human Intelligence Task (HIT). Each participant watched a hypermeeting and answered questions about their experience with each of the three designs. The study was developed using best practices from the literature [13][14][18].

The overall structure of the HIT was training followed by three sections for the three designs. There was additional training between sections. Each section consisted of two halves of watching video followed by questionnaires.

Because hypervideo is not a common concept, training included several instructional resources beginning with a textual description of the task. This was followed by a storyboard explaining how the hypervideo was recorded (shown in Figure 9). Finally, a video was presented introducing the interface and its features containing scenes from a similar hypermeeting to the one they would see in the actual task.

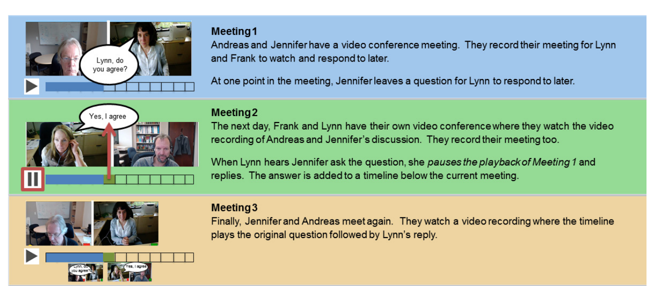

Figure 9. Storyboard for study.

[2] https://en.wikipedia.org/wiki/Gantt_chart

Before switching to a different timeline design, participants were shown a video explaining the features in the particular timeline design. During training, participants answered seven questions about the videos to ensure attention and comprehension.

The hypermeeting in the actual task consisted of three separate meetings between two pairs of individuals and covered three topics. While the order of topics was fixed – participants saw the meetings from beginning to end – the order of designs was counterbalanced using a Latin Square design to account for learning and fatigue effects. There were six possible orders for viewing the timelines: two in which each design (single, merged, or stacked) appeared first, two in which each design appeared second, and two in which each design appeared third.

The overall goal of the presented hypermeeting was to establish a conference travel policy for a research lab. The three topics were (1) whether to have individual budgets or budgets controlled by managers, (2) how much per person, and (3) what effect having publications at the conference should have, if any.

To make sure that all participants experienced the playback plan in the same way, we recorded a screen-capture movie of the playback of the hypermeeting in each timeline design. This gave participants experience with each design while minimizing the required time. The playback speed was set to 1.2 times real-time. Participants could not pause or rewind the movie. The movie paused twice for each topic/design. At each pause, the participant was asked 3-4 questions about the content of the segment they had just watched.

After the second segment on a topic, participants were asked to answer eight 5-point Likert-scale questions (*strongly disagree* to *strongly agree*) about the usability, understandability, and design of the visualization (see Table 1). Finally, after watching the last segment of the hypermeeting, participants were asked which were their favorite and least favorite interfaces and why. They were also asked to provide feedback on the trade-offs between the designs and any other comments they might have on their experience.

6.2 Participants

The 126 participants in the study were part of the US Mechanical Turk population who had at least a 95% acceptance rate for their past work. They were paid $4.50 for the approximately half hour it took to take part in the study regardless of whether we used their data or not. Nine participants who made eight or more errors on the 26 factual questions (seven from training and 19 from actual meeting footage) were removed from consideration to ensure data was from participants who paid attention to the content and activity. To ensure proper balance, data for the first 18 of the remaining participants in each of the six orders of exposure to the designs were included in the final analysis. Thus the remaining discussion concerns data from 108 participants.

The participants included 43 women and 65 men. Most of the participants were in their 20s (47) or 30s (39) although there were also participants in their 40s (13), 50s (5), teens (3), and 60 or over (1). All but 16 participants had at least some college education with 12 having graduate degrees, 47 having Bachelor degrees and 11 having Associates degrees. Nine of the 22 participants reporting some college time but no degree were among the 16 current students.

We also asked about participant experience with videoconferencing as that is the most common related class of system to the one being viewed in the study. 61 reported at most infrequent use. This included 13 reporting never using videoconferencing software, 22 reporting use of a couple times, and 26 reporting a few times a year. There were 47 participants that reported regular use including 25 reporting a few times a month, 16 reporting weekly use, and 6 reporting daily use. While only half the participants regularly use videoconferencing software, the assessments of the other participants are still valuable because the use of videoconferencing in the office increases at a rapid pace.

7. Results

Results are presented regarding usability assessments for the designs, the favorite and least favorite designs as identified by participants, and comments on trade-offs between designs or other suggestions/feedback. Participants were also asked to answer comprehension questions about the content of the meeting but there was no effect of design on the correctness of answers so that data is not reported further.

7.1 Usability Assessments

To analyze the effect of timeline type on user experience, we combined the responses to the eight Likert-scale questions (see Table 1) into an overall "usability" scale with a maximum score of 40. We inverted the scores of the odd-numbered questions such that a score of 5 was always most positive. We then performed a within-subjects analysis, using a 3x6 repeated measures mixed-model ANOVA where usability was the dependent measure and the three timeline types, six possible timeline orders, and the interaction between timeline type and order were independent measures. We also included participant familiarity with video conferencing as a covariate, and Participant ID was modeled as a random effect.

There was a significant main effect of timeline type on usability ($F(2,204)=13.53$, $p<.001$). A Tukey HSD post-hoc test revealed the merged timeline was rated as significantly more usable ($M=29.69$) than either the single timeline ($M=27.52$) or the stacked timeline ($M=26.44$).

We also expected that there would be some learning effects as participants gradually gained familiarity and confidence with hypermeetings. There was a significant main effect of the order in which participants experienced the timelines on perceived timeline usability ($F(5,102)=2.7$, $p=.02$). People who saw the single timeline first rated all timelines (on average) as significantly more usable than people who saw the stacked timeline first.

Prior experience has shown that it can take users some time to understand the structure of a hypermeeting. This challenge likely affected participant experience during the study. As a result, we compared the usability results between the 36 participants that saw each design last. In this case, the people who saw the merged timeline last rated it as significantly better ($M=29.47$) than those

1. The timeline helped me understand the flow of the conversation between people.
2. The timeline distracted me from the content of the meeting.
3. The timeline helped me understand relationships between the meetings.
4. The timeline was confusing.
5. I expect I could use the timeline to accurately find a particular point in the discussion.
6. The timeline hid information I would find valuable.
7. The order of the segments of meetings played made sense to me.
8. The transitions between meetings were confusing.

Table 1. Likert-scale questions about each timeline design

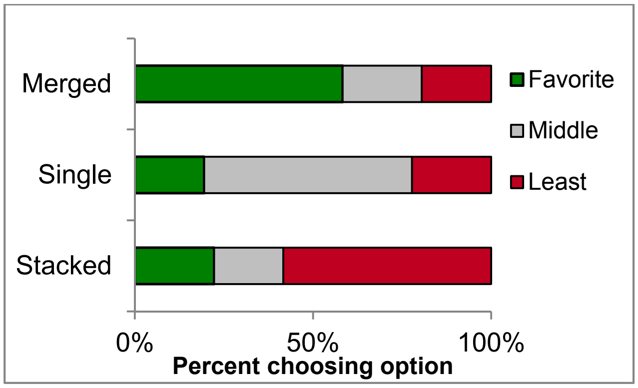

Figure 11. Relative preferences of each timeline design.

who saw the single timeline last (*M*=25.67) ($F(2,105)=3.08$, $p=.04$).

There was a near-significant interaction between timeline type and order on usability ($F(10,204)=1.78$, $p=.06$). In this case, the single timeline was rated as significantly more usable (*M*=30) when it came first in the order of timeline evaluation compared to when it came last in the order (*M*=22.66). This suggests that when participants had more experience with the system and with other alternative designs, the single timeline was viewed as even worse.

Finally, there was a significant effect of familiarity with video conferencing; as participants' prior experience using video conferencing tools increased, it had a positive effect on overall perceptions of timeline usability ($F(1,101)=16.26$, $p<.001$).

7.2 Favorite and Least Favorite Designs

At the end of the study, we had participants compare the three timelines against each other and discuss their favorite and least favorite designs (and why). This was done to mitigate the impact of potential order effects. These comments provided more insight into the design elements that worked and were challenging across the different designs.

The merged timeline was identified by 63 of the 108 participants as their favorite. Stacked was favored by 24 and single was favored by 21. Figure 11 shows the percentages of respondents who chose each timeline type as their favorite, middle, and least favorite design. The open-ended comments gave further insights into the most prevalent positive and negative aspects of the various designs.

7.2.1 Merged Timeline

The most positive aspect of the merged timeline was its general ease of interpretation. 19 of the 63 participants who chose it as their favorite mentioned that it was easy to tell where the meeting would switch and who would be speaking next. We initially expected that the parsimonious nature of the merged timeline design would create a stable context for interpretation that was also not too complex. Multiple comments echoed agreement with these principles. The merged timeline was described as straightforward or easy to follow and understand by 14 individuals. 14 respondents also mentioned it was compact, simple, or consolidated.

For me it was easiest to follow. I always knew when it would transition and could anticipate which group would be speaking. I could also see what the conversation was about if I looked at the bottom. With this timeline I could skip around and still know where to find the information I needed.

Another positive aspect of the merged timeline was that it presented a continuous view of the meetings (5), and unlike the other two views, there was no movement (5). This made it less distracting (4) and less intimidating (1).

It made the most sense for someone watching the entire set of meetings in one shot, as for the viewer, it's just one meeting that jumps around.

There were several negative comments about the merged timeline. Seven people found the view to be too crowded, with too much information in too little space.

One conceptual challenge in the design of timeline visualization was how to avoid creating distractions with a lot of movement and visual cues jumping around in the timeline making it confusing to anticipate when the meeting would change and when. The merged design avoided this issue by signaling when a change would occur without using visual motion over several areas of the screen.

7.2.2 Single Timeline

Overall, the single timeline fell between the merged and stacked timeline in terms of preference. Positive aspects of this design also focused around the relative simplicity of the design, which was seen as being clean and digestible (5) and easy to follow and understand (5).

However, unlike the merged timeline, that showed all the information in a single view, the single timeline transitioned between three individual meeting timelines. This led some people to feel as though information was being hidden from them since they could only see one meeting at a time (5). Also, unlike the merged timeline, more people found it difficult to know when the meeting would jump and where it would jump to (5).

7.2.3 Stacked Timeline

The stacked timeline was the most detailed and also required the most screen real estate to show the individual meeting timelines all in one place. As a result, it was also necessary to use attention management principles to help users understand where they were at any given point and where they were going when a jump occurred. This attention management device was implemented as a moving gray circle that showed the flow of the conversation. Eight people appreciated the use of the circle to show the flow of conversation. The use of decomposition to divide up the different meeting timelines was also a frequent positive aspect mentioned by five people who preferred the stacked timeline. One person appreciated that this timeline showed the "most data."

However, while some people appreciated the gray dot moving between individual timelines, nearly twice as many disliked this feature. In particular, many viewed the moving dot as distracting, which made it hard to follow the conversation with the visual movement. Additionally, eight users found the information density to be too great, describing it as "too spread out" or "cluttered and messy." The following quotes illustrate some of the frequently-mentioned negatives about the stacked timeline:

I found this to contain too much extraneous information and it was difficult to really piece together the timeline of when it was going to jump to another meeting and where in the meeting it would jump to. There is way too much clutter and confusion in this design. More is better in most cases, but not here. It didn't really make intuitive sense and all of that information is not only confusing, but also distracting.

It was confusing to try to look at the timeline and understand what was going on and follow a coherent discussion at the same time.

This presentation can get busy, and it feels too easy to get lost in the flow of the different responses and answers to one another. It feels like something a physics professor would love, though! It's *like some play of linear time presented across three separate possibilities.*

Our initial hypothesis was that the stacked timeline would be preferred because it provides the clearest visualization of the playback plan within the context of the whole hypermeeting. As the study participants could not access the video outside the playback plan, this information may not have been relevant for them. Also, the comments make clear that dividing attention between the video and the timeline was difficult. Further studies will investigate whether the stacked timeline is more useful for information seeking tasks.

7.3 Comments on Trade-Offs

After they had experienced all three timelines, users were asked to comment on the trade-offs between each version. Overall, there is a tradeoff between decomposition and providing too much information (the main complaint for the stacked timeline). Also, there is a tradeoff with using parsimony to visually navigate the user through multiple streams without being distracting, but essentially hiding most of the information from the user (the complaint with the single timeline). Many of the users realized the tradeoffs between a large amount of detailed information (which could induce cognitive load) and a summary of the information, which is less informative but easier to mentally follow.

You can get more information in a design like Design 2 [stacked], but you keep it less confusing if you limit it to one timeline (like in 1 [single] and 3 [merged]). Loss of information isn't great but the simplicity of the system is apparent in both 1 [single] and 3 [merged]. While the advantages of more information is obvious, the information needed in these types of meeting timelines can definitely be condensed.

The merged and single timelines can create clutter that the stacked timeline alleviates, but the stacked timeline introduces a flow problem.

The merged is easy to see everything all at once while it may take some time to figure it out. The stacked is easier to see where things come from but it's very hard to put it together mentally and the single is basic and easy to understand but feels like its missing things.

For many users, the merged timeline seemed to be a happy medium between these two extremes: *I thought that the merged design had the best of both worlds.*

However, at the same time, the relative advantages or disadvantages of the different timelines could vary depending on the user's goals and tasks in using the timeline to review a meeting. For example, if the goal is not to passively watch and follow the meeting but to understand the evolution of a specific discussion topic, the decomposed view of the stacked timeline could be helpful. As one participant explained,

I think it might be easiest to find a specific and particular piece of a conversation or response using the exploded, stacked visual representations of conversation. I think if you were sitting do [sic] to watch this as a unified whole to have it make sense, that the merged would be best. Merged definitely, I think, gave me the best feel for the different meeting flows.

Another also mentioned the potential need to find a specific point in the conversation:

The single is simple, but maybe too simple. If I was looking for a specific conversation that was being had in the conference it might be a little more difficult to spot that exact moment. The merged was a great balance between the stacked and single with minimizing the timeline without cutting out too much of the important details.

In future work, we will evaluate the timeline design in a variety of information-seeking use case scenarios with different playback plans to see how a system might dynamically support the same user performing different types of meeting review and comprehension tasks.

8. Conclusions and Future Directions

Reviewing a chain of meetings that span time and location is not well-supported by existing systems. Using a series of different review scenarios, we built upon prior work [8] by intelligently inferring playback plans that express users' information-seeking goals. We provided an in-depth description of alternative ways of generating, visualizing and playing back a series of linked asynchronous meetings. We created three visualizations of the hypermeeting timeline, incorporating guidelines for using multiple views for information visualization. These visualizations help users understand automatic link following within a sequence of meetings.

We conducted a user evaluation of the *review-all-meetings* playback plan. Watching previous meetings as an observer is a common use case, especially for viewers who did not attend those meetings. The results showed that the merged timeline design was the most beneficial presentation. This design presented all information, including transitions, in a single condensed view. However, this visualization required users to make tradeoffs between obscuring details of the topics and limiting distraction during navigation flow. Participants indicated that switching among visualizations would be beneficial for different tasks.

This result generalizes to additional hypervideo contexts. In particular, the visualizations studied are reasonable choices for adaptive hypervideo when providing paths through educational or how-to content. Playback plans in such contexts would reason about the user's understanding of prerequisite knowledge and the need for detailed description of a how-to process much as they reason about prior knowledge of meetings in our context.

By building on the design insights uncovered in this work, we can continue to explore how to design personalizable hypervideo tools. In particular, such tools should enable users to quickly and easily extract even more fine-grained and useful information from a time-distributed series of videos, both in the context of reviewing meetings and beyond. For example, we will explore the interaction between timeline visualizations and alternate types of information-seeking scenarios, such as finding a particular topic in the meeting.

We will also explore more organic user behavior when navigating through the meeting using various approaches. An example is to use the timeline versus the navigation history stack to have more user-defined control in moving back and forth throughout the content.

Lastly, we plan to add support for other types of meetings, such as local meetings recorded by a single video camera or video conferences between two hubs. In both cases, there would be more than one person per video feed requiring more advanced approaches to participant and speaker identification.

9. REFERENCES

[1] Jeremy Barksdale, Kori Inkpen, Mary Czerwinski, Aaron Hoff, Paul Johns, Asta Roseway, and Gina Venolia. 2012. Video threads: asynchronous video sharing for temporally distributed teams. *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work,* ACM, 1101–1104.

[2] Guillaume Boissière. 1998. Automatic creation of hypervideo news libraries for the World Wide Web. *Proceedings of the ninth ACM conference on Hypertext and hypermedia,* ACM, 279–280.

[3] Jonathan J. Cadiz, Anand Balachandran, Elizabeth Sanocki, Anoop Gupta, Jonathan Grudin, and Gavin Jancke. 2000. Distance learning through distributed collaborative video viewing. *Proceedings of the 2000 ACM conference on Computer supported cooperative work,* ACM, 135–144.

[4] Teresa Chambel, Carmen Zahn, and Matthias Finke. 2004. Hypervideo design and support for contextualized learning. *IEEE International Conference on Advanced Learning Technologies,* IEEE, 345–349.

[5] Brian Dorn, Larissa B. Schroeder, and Adam Stankiewicz. 2015. Piloting TrACE: Exploring Spatiotemporal Anchored Collaboration in Asynchronous Learning. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing,* ACM, 393–403.

[6] Lutz Gericke, Matthias Wenzel, and Christoph Meinel. 2014. Asynchronous understanding of creative sessions using archived collaboration artifacts. *International Conference on Collaboration Technologies and Systems (CTS),* IEEE, 41–48.

[7] Werner Geyer, Heather Richter, and Gregory D. Abowd. 2005. Towards a smarter meeting record—capture and access of meetings revisited. *Multimedia Tools and Applications* 27, 3: 393–410.

[8] Andreas Girgensohn, Jennifer Marlow, Frank Shipman and Lynn Wilcox. 2015. HyperMeeting: supporting asynchronous meetings with hypervideo. *Proceedings of the 23rd ACM international conference on Multimedia,* ACM, 611-620.

[9] Andreas Girgensohn, Frank Shipman, and Lynn Wilcox. 2011. Adaptive clustering and interactive visualizations to support the selection of video clips. *Proceedings of the 1st ACM International Conference on Multimedia Retrieval,* ACM, Article 34.

[10] Andreas Girgensohn, Lynn Wilcox, Frank Shipman, and Sara Bly. 2004. Designing affordances for the navigation of detail-on-demand hypervideo. *Proceedings of the working conference on Advanced visual interfaces,* ACM, 290–297.

[11] Romulus Grigoras, Vincent Charvillat, and Matthijs Douze. 2002. Optimizing hypervideo navigation using a Markov decision process approach. *Proceedings of the tenth ACM international conference on Multimedia,* ACM, 39–48.

[12] Seth Hunter, Pattie Maes, Stacey Scott, and Henry Kaufman. 2011. MemTable: an integrated system for capture and recall of shared histories in group workspaces. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* ACM, 3305–3314.

[13] Markus Jakobsson. 2009. Experimenting on Mechanical Turk: 5 how tos. *ITWorld, September* 3: 2009.

[14] Aniket Kittur, Ed H. Chi, and Bongwon Suh. 2008. Crowdsourcing user studies with Mechanical Turk. *Proceedings of the SIGCHI conference on human factors in computing systems*, ACM, 453–456.

[15] Andrew Lippman. 1980. Movie-maps: An application of the optical videodisc to computer graphics. *ACM SIGGRAPH Computer Graphics,* ACM, 32–42.

[16] Franz Lehner, Michael Langbauer, and Nadine Amende. 2014. Measuring success of enterprise social software: the case of hypervideos. *Proceedings of the 14th International Conference on Knowledge Technologies and Data-driven Business,* ACM, 3.

[17] Shixia Liu, Yingcai Wu, Enxun Wei, Mengchen Liu, and Yang Liu. 2013. Storyflow: Tracking the evolution of stories. *IEEE Transactions on Visualization and Computer Graphics,* 19, 12: 2436–2445.

[18] Catherine C. Marshall and Frank M. Shipman. 2013. Experiences surveying the crowd: Reflections on methods, participation, and reliability. *Proceedings of the 5th Annual ACM Web Science Conference*, ACM, 234–243.

[19] Thomas P. Moran, Leysia Palen, Steve Harrison, Patrick Chiu, Don Kimber, Scott Minneman, William van Melle, and Polle Zellweger. 1997. "I'll get that off the audio": a case study of salvaging multimedia meeting records. *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems,* ACM, 202–209.

[20] Mukesh Nathan, Mercan Topkara, Jennifer Lai, Shimei Pan, Steven Wood, Jeff Boston, and Loren Terveen. 2012. In case you missed it: benefits of attendee-shared annotations for non-attendees of remote meetings. *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work,* ACM, 339–348.

[21] Juan Pan, Li Li, and Wu Chou. 2012. Real-Time Collaborative Video Watching on Mobile Devices with REST Services. *Third FTRA International Conference on Mobile, Ubiquitous, and Intelligent Computing (MUSIC),* IEEE, 29–34.

[22] Adam Perer and Fei Wang. 2014. Frequence: Interactive mining and visualization of temporal frequent event sequences. *Proceedings of the 19th international conference on Intelligent User Interfaces,* ACM, 153–162.

[23] Madjid Sadallah, Olivier Aubert, and Yannick Prié. 2014. CHM: an annotation-and component-based hypervideo model for the Web. *Multimedia tools and applications* 70, 2: 869–903.

[24] Nitin Sawhney, David Balcom, and Ian Smith. 1996. HyperCafe: narrative and aesthetic properties of hypervideo. *Proceedings of the the seventh ACM conference on Hypertext,* ACM, 1–10.

[25] Frank Shipman, Andreas Girgensohn, and Lynn Wilcox. 2008. Authoring, viewing, and generating hypervideo: An overview of Hyper-Hitchcock. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* 5, 2: 15.

[26] John Tang, Jennifer Marlow, Aaron Hoff, Asta Roseway, Kori Inkpen, Chen Zhao, and Xiang Cao. 2012. Time travel proxy: using lightweight video recordings to create asynchronous, interactive meetings. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* ACM, 3111–3120.

[27] Claudio AB Tiellet, André Grahl Pereira, Eliseo Berni Reategui, José Valdeni Lima, and Teresa Chambel. 2010. Design and evaluation of a hypervideo environment to support veterinary surgery learning. *Proceedings of the 21st ACM conference on Hypertext and hypermedia,* ACM, 213–222.

[28] Michelle Q. Wang Baldonado, Allison Woodruff, and Allan Kuchinsky. 2000. Guidelines for using multiple views in information visualization. *Proceedings of the working conference on Advanced visual interfaces,* ACM, 110–119.

[29] Krist Wongsuphasawat, John Alexis Guerra Gómez, Catherine Plaisant, Taowei David Wang, Meirav Taieb-Maimon, and Ben Shneiderman. 2011. LifeFlow: visualizing an overview of event sequences. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* ACM, 1747–1756.