HyperMeeting: Supporting Asynchronous Meetings with Hypervideo

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

While synchronous meetings are an important part of collaboration, it is not always possible for all stakeholders to meet at the same time. We created the concept of hypermeetings for meetings with asynchronous attendance. Such hypermeetings consist of a chain of video-recorded meetings with hyperlinks for navigating through them. Our HyperMeeting system supports the viewing of prior meetings during a videoconference. Natural viewing behavior such as pausing video generates hyperlinks between previous and current meetings. During playback, automatic link-following guided by playback plans present the relevant content to users. Playback plans take into account the user's meeting attendance and viewing history and match them with features such as topic and speaker segmentation. A user study showed that participants found hyperlinks useful but did not always understand where the links would take them. Experiences from longer-term use and the study results provide a good basis for future system improvements.

Categories and Subject Descriptors

H.4.3 [**Information Systems Applications**]: Communications Applications – *computer conferencing, teleconferencing, and videoconferencing.* H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems – *hypertext navigation and maps, video.*

Keywords

Hypervideo; hypermeeting; meeting capture; asynchronous meetings; teleconferencing.

1. INTRODUCTION

Meetings are an important part of collaboration. Although it is well known that people prefer face-to-face meetings, this is not always possible. For the case where attendees are remote but a synchronous meeting can be arranged, Skype or Google Hangouts can be used. However, when users are in different time zones or are traveling, finding a time for a synchronous meeting among all members is frequently difficult and sometimes impossible. One option is to record the meeting so that missing members can review it asynchronously. While there are third-party tools for recording such videoconferences, experience indicates that people rarely go back and listen to the recordings, resulting in the need for participants to bring missing members up to speed by repeating parts of the past meeting. It can also be burdensome to review the entirety of previous meetings for a few key action items or takeaway points [14]. Formal meetings have minutes but those are not a replacement for access to the meeting discussion as they tend to record decisions and not why the decision was made.

We propose the notion of a hypermeeting as a means to capture and review meetings in an asynchronous manner. Synchronous participants record the meeting, while later participants simultaneously review the meeting and add comments. We hypothesize that the ability to add comments will encourage members who missed it to actually listen to the meeting. In turn meeting attendees will be interested enough to review the comments made about their meeting. Thus rather than isolated recordings, these hypermeetings will allow a single meeting to be held asynchronously. We pose the following research questions:

1. How can technology support the creation, continuation and review of asynchronous, distributed discussions over time?
2. How does allowing people to review content in a non-linear way influence the process of navigating through and understanding a series of past meetings?

The HyperMeeting system combines meeting capture with generating and viewing hypervideo. Video from multiple synchronous videoconferencing attendees is recorded and indexed by speaker and topic. During subsequent meetings, this video is viewed synchronously by potentially geographically separated attendees, and discussions of topics in the previous meeting or new topics are recorded. Pausing prior recordings usually leads to further discussions of the same content so that we chose this interaction as the trigger for automatic hyperlink creation. We use the terms *hypermeeting* and *meeting chain* to describe such a series of interlinked meetings.

These linked video recordings form a hypervideo that can be navigated in different ways. Our novel concept of a playback plan infers information needs of watchers from their roles in a meeting chain to support automatic playback behavior. Plans can take into account features of the current meeting and attendee actions, involvement in prior meetings by current meeting attendees, and topics attached to meeting content to decide which links to follow automatically in a set of interlinked meetings. Plans can include filters to skip undesired content such as silence or previously watched video segments.

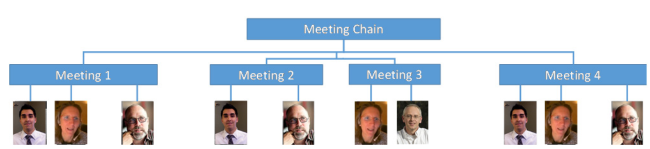

Figure 1. Example of sequence of project meetings.

MM'15, October 26-30, 2015, Brisbane, Australia

DOI: http://dx.doi.org/10.1145/2733373.2806258

The novel contributions of our work are the support for a chain of synchronous meetings forming an asynchronous hypermeeting, the synchronized playback of video streams by attendees of follow-up meetings, the ease of creating hyperlinks from recorded to live video, and automatic playback guided by playback plans.

The next sections present a usage scenario, the conceptual framework, and related work. Section 4 describes the HyperMeeting architecture and its components. Sections 5 and 6 present experiences gained from longer-term use and the results of an initial user evaluation of HyperMeeting. The conclusions in Section 7 indicate directions for future work.

2. SCENARIO AND FRAMEWORK

HyperMeeting can record and support access to meeting content in many contexts but our focus is enabling project teams distributed over multiple time zones. Figure 1 shows an example of such a situation. Consider the case where periodic meetings occur for a project involving a distributed team, John and Ann in San Francisco and Anders in Munich. As such, remote meeting software is required because members of the team have other responsibilities and are not available for every scheduled meeting. Additional team members at both sites are involved in the meetings. This scenario and our first system design focus on egalitarian meetings – the type of meetings most valuable to our initial user community.

In Figure 1, a kickoff videoconference (Meeting 1) occurs with John, Ann and Anders each recorded in their own video (although John and Ann could have been in the same room). The next week, Ann is not available at the normal meeting time but John and Anders meet to discuss progress and look back at the recordings from the prior week's meetings to make sure they remember Ann's perspective on some of the design issues (Meeting 2). As they replay and comment on the prior recordings, links are automatically generated between the original and later meeting and, sometimes, John and Anders attach labels to the links to indicate the subject of the discussion. Later, Ann meets with Bob to further discuss the project (Meeting 3). Their meeting is initially a set of reactions and responses to Meeting 2 but moves on to other topics as well. The regularly scheduled distributed meeting (Meeting 4) reviews the recordings and links from the previous meetings to guide further discussions of the main subjects.

An important aspect of supporting both those attending and missing meetings is to establish relations among meetings as new meeting content is recorded. We define a *meeting chain* or *hypermeeting* to be a recurring meeting or a set of interrelated meetings. Each meeting is composed of synchronized video streams and other associated data. A meeting may consist of just a single video stream in cases when an individual is recording responses to a previously recorded meeting that they could not attend or if an in-person meeting was recorded by a single camera.

Metadata about the meeting series include the list of attendees for each meeting, the topics assigned to segments of the meetings, and the links between meetings. Segmentation of the meetings based on who is speaking and the topics being discussed is used to provide finer access into the recordings. The framework allows for a variety of automatic, semi-automatic, and manual techniques to be used for segmentation, indexing, linking, and playback.

3. RELATED WORK

Previous research has employed 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 aimed at addressing this problem focuses on capturing the content of meetings for later consumption. Meeting capture systems [6][7][10][13] record content for later playback and browsing, taking 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 have been designed to support asynchronous participation in meetings, e.g., Time Travel Proxy [19] and Video Threads [1], by having meeting attendees record and view short videos on topics with responses. In Time Travel Proxy, an absent meeting attendee pre-records a video message that is played in the meeting and the attendees' reaction is recorded. Video Threads puts asynchronous video recordings into a conversation thread similar to an email discussion. These recordings form a discussion thread that supports responses on individual topics but loses the connections between the responses, i.e., the fact that they came from the same meeting.

A third class of systems allow viewers and creators of content to make video more interactive by creating navigable links or annotations within the video itself, e.g., hypervideo. This has been viewed as particularly valuable for education [20]. For example, recent work by [5] in the educational domain allows viewers to annotate and embed discussions at particular locations and times within a video, and to peruse and continue the discussion asynchronously using these "spatiotemporal links." The economic viability for enterprise use of interactive video to support decision making is explored in [11].

General purpose hypervideo systems (e.g., [17][18]) enable rich linking between video content but prior work on hypervideo has focused on authored content. A video-recorded meeting could be taken as the starting point for such authoring 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 nature of the video (e.g., attachment of metadata indicating attendees, topics, and time/date indexing). Hypermedia models like CHM [16] represent relationships between content elements and definition of presentation arrangements. Such models could be used to represent speaker, topic, and link information. Finally, none of these systems support geographically distributed synchronous viewing of hypervideos although there has been work to optimize prefetching of video content in anticipation of user-driven navigation [9].

A last class of related systems support the synchronous viewing of video content on distributed devices [3][15]. These systems can include additional channels of communication (e.g., chat) for synchronous viewers to discuss what they are watching but these records of communication are, at best, attached as metadata to the video content. In our context, it is important that each meeting be treated as primary content whether or not it is the first meeting in a meeting chain.

4. HYPERMEETING SYSTEM

We designed the HyperMeeting system to address the research question of *how technology can support the creation, continuation and review of asynchronous, distributed discussions over time.* The system provides two main functionalities, a video conferencing system with recording capability and an environment for collaboratively viewing a chain of previously recorded meetings. Figure 2 shows the user interface. During a meeting, any attendee

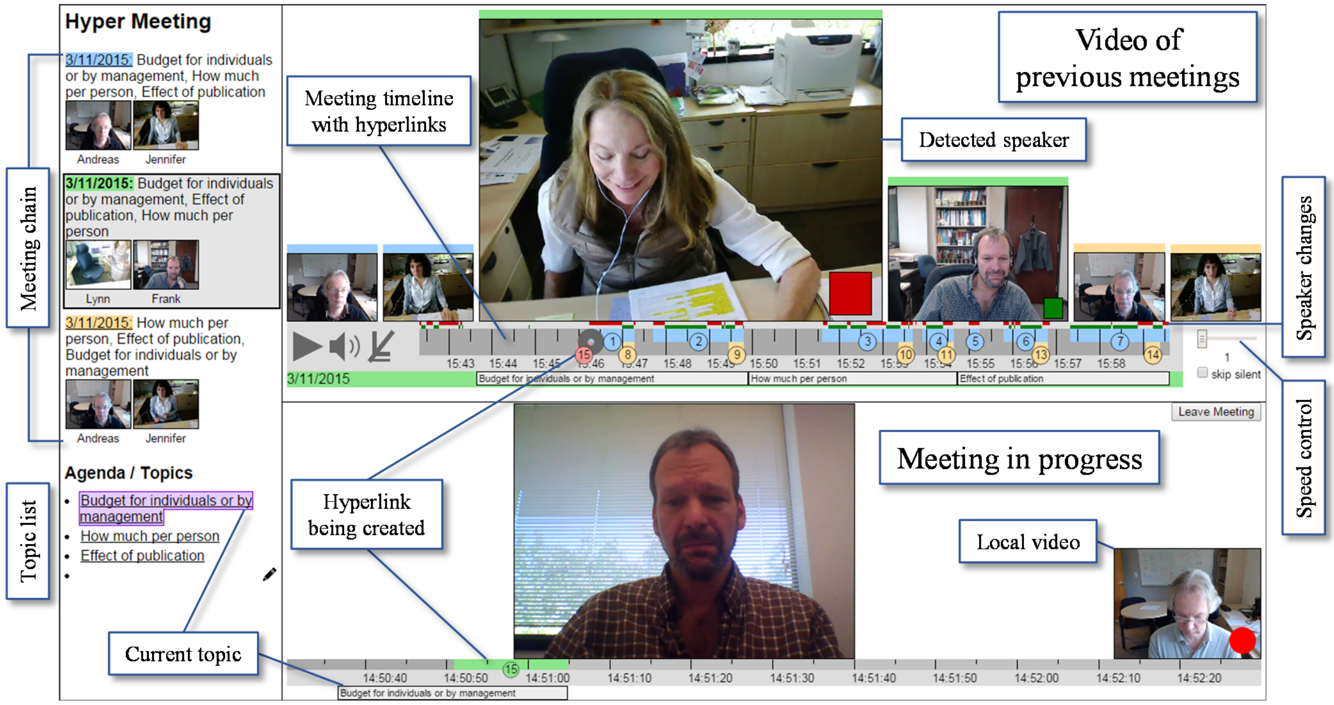

Figure 2. HyperMeeting recording interface continuing a chain of three previous meetings.

may navigate hyperlinks connecting video recordings of previous meetings, while keeping the view of all attendees synchronized.

In the following subsections, we cover how a chain of meetings is recorded and hyperlinks are created to form an asynchronous hypermeeting. We then describe synchronized playback of multiple video streams by multiple attendees of follow-up meetings. We introduce the novel features of automatic link-following and of playback plans that adapt playback for different contexts of use. Finally, we describe the web-based architecture of the HyperMeeting system.

4.1 Meeting Recording

A unique characteristic of HyperMeeting is that video playback and annotation can be a collaborative activity. Previously recorded meetings can be continued in a new synchronous video meeting that is added to the meeting chain. Video players showing oneself and the other attendees of the current meeting are displayed in the bottom right portion of the interface (see "Meeting in progress" in Figure 2). A red circle on top of the local video display in the bottom-right indicates that the video is being recorded. The video of each meeting participant is time-stamped and uploaded to a central server for further processing.

Meeting attendees may collaboratively edit a list of topics (see "Topic list" in Figure 2). Topics can be assigned during the initial recording by attendees, by those watching the videos later, and potentially through automated mechanisms that match the elements of predefined agendas based on content processing. The topics are started and ended during the meeting by clicking in the list. In the recording of the meeting, those topics are attached to the timeline to segment the meeting and to allow for easy navigation to a particular topic. Topics visited during playback of prior meetings are automatically inherited unless a different topic is manually selected. In Figure 2, the topic "Budget for individuals or by management" was started when playback of the recorded meeting entered that topic and would end either when playback moved to a different topic or when a meeting attendee selected a different topic.

4.2 Creating Hyperlinks

Having a record of the relationships among meetings is valuable for following discussions across meetings. The playback activity provides insight into these relations. Consider a later meeting, with different attendees, watching a previously recorded meeting. We observed that attendees in such a context pause the recording of the previous meeting to discuss the topic being presented. We support this behavior by creating a hyperlink between the previous recording and the meeting being recorded whenever playback is paused. At a later time, when both meetings are reviewed together, the hyperlink enables the later watcher to see the continued discussion in the context of the earlier meeting.

Such hyperlinks consist of a video segment in both the older and newer video. We call those segments source and destination link anchors. The destination link anchor in the current video recording starts when video playback is paused and ends when playback resumes (see circle "15" in Figure 2). The source link anchor in the prior meeting ends at the playback position. To provide context for the older meeting, the source anchor starts a fixed duration prior to the playback position. Because playback is sometimes paused without the intention to create a hyperlink, we provide a link-cancel button that removes the automatically created hyperlink with the option to add it later. Users may add a caption to a hyperlink that is displayed as a tooltip during video playback.

There are multiple ways to determine the source link anchor. The obvious solution is to use the point in the older meeting where the link was created. This is likely to be in the middle of a statement and also misses the discussion in the prior meeting that was the impetus for the newer content. Currently, HyperMeeting starts the source link anchor 20 seconds prior to the paused point. Future versions can improve on this by adjusting the start to be at the

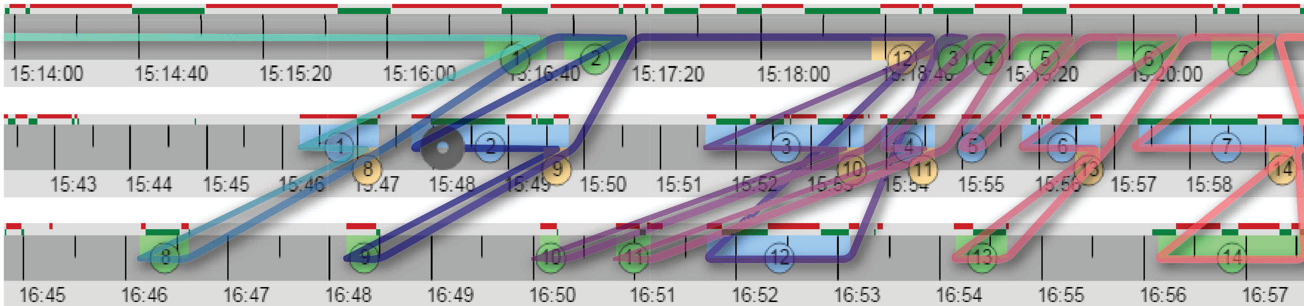

Figure 3. Playback plan for the oldest meeting with automatic link following. Gradient color indicates playback progression.

beginning of a speaker's turn, the beginning of a topic, or some combination of these features.

4.3 HyperMeeting Playback

Any meeting attendee can initiate playback for all current attendees, navigate among the previously recorded meetings in the meeting chain, and pause or skip through the recorded video streams. This synchronization ensures a common context for current attendees continuing discussions based on prior meeting content. All video streams belonging to the same meeting are synchronized as well during playback.

The meeting chain shown at the left of Figure 2 visualizes the context of the synchronous meetings in the chain. Meetings are represented by a combination of the date of the recording, the most-used topics, and the attendees at the meeting. Meetings are color-coded by age in a range from blue to red. Clicking on a meeting in the chain displays the corresponding timeline and starts playback at the start of that meeting. The timeline visualizes the context of the selected meeting with the time-of-day, links to earlier and later meetings, and the selected topics.

Topics are highlighted when the player reaches a segment of the meeting noted as being concerned with a topic. In Figure 2, the current topic is "Budget for individuals or by management". Users viewing the recorded meeting can add or change the topic selected by clicking on the topic in the topic list that is then attached to the timeline. Topics attached to the timeline can be used for navigation. Clicking on one of those topics skips the video playback to the start of the topic. As described earlier, topics encountered during playback are automatically assigned in the live meeting unless the attendees selected a different topic.

The top right portion of the interface is for watching and navigating within the video streams associated with the playback of a recorded meeting (see "Video of previous meetings" in Figure 2). Video players for all videos across all meetings in the meeting chain are available to the user. Each of these video players is color coded above the video frame to indicate the corresponding meeting.

In Figure 2, the middle meeting (with the green color code above the frames) is currently playing. Its video players are larger than the others and show the speaker color in the corner. Speaker transitions are animated with the player associated with previous speaker shrinking as the player of the new speaker grows in size.

The timeline shows the meeting progressing as the time-of-day with a 24-hour clock. Above the timeline, the progression of speakers is color-coded. Topic tags are presented below the timeline next to the date and color-coding of the meeting. Controls to the left of the timeline let the user pause or mute the playback. The link cancel button is only present in live meetings. To the right of the timeline is a slider to speed the playback up to three times the normal speed.

4.4 Link Following

Hyperlinks are shown in the timeline as numbered circles with colors indicating the meeting at the other end of the link. Figure 3 shows timelines of a meeting chain with the newest meeting at the bottom. The middle timeline is also shown in Figure 2, labeled "Meeting timeline with hyperlinks." The vertical position of the link circle indicates whether it is a source link anchor pointing down to a newer meeting or a destination link anchor pointing up.

Links can be followed in both directions by clicking on the numbered circle identifying a link. When following a link, first the link anchor in the older meeting is played to provide context and then the link anchor in the newer meeting is played. In Figure 4, after clicking the green "1", first the green source link anchor around the circle would be played and then the destination link around the blue "1", as indicated by the solid line. The playback would stop at the end of the destination link anchor. Clicking on the blue "1" would cause the same playback sequence. A second click on a link circle advances the playback to the destination link anchor. Our user study described later in this paper uncovered some issues with the sequence of playing the two link anchors.

4.5 Automatic Link-Following

Viewing a hypermeeting can be a purely manual interaction where the user selects which hyperlinks should be followed and which parts of the video should be skipped. However, the system's knowledge of the viewing users' participation and interests can be used to automate the navigation through the hypermeeting.

In addition to user-initiated link traversal, links may also be automatically traversed as they are encountered. Continuous playback of the multipart meeting can move back and forth between the original and additional video content as links are encountered. For example, as links to newer meetings are encountered, the playback automatically follows that link and returns to the original position after playing the destination link anchor.

If the linked video contains a link to a more recent meeting, that link is followed in the same fashion. The dashed line in Figure 4 provides an example of this behavior. While the user played the destination anchor of link "1", link "8" was encountered and automatically followed. Because the ends of the link anchors of links "1" and "8" line up, playback would stop at the end of the green segment depicting the destination anchor of link "8".

4.6 Playback Plans

Depending on the information need of the user, there are different ways links could be followed automatically. We created the abstraction of playback plans to describe such automatic behavior.

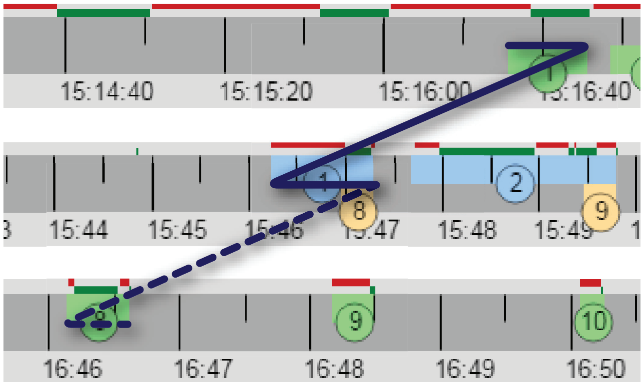

Figure 4. Playback plan for link following. The dashed part indicates automatic link following.

4.6.1 *Types of Playback Plans*

Playback plans are computed when a user clicks on a link, navigates to a different meeting in the chain, or navigates outside the current plan. Playback plans can be used to filter recorded content and are interactive (i.e., users can navigate within a playback plan).

The result of a specific playback plan applied to a specific hypermeeting 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. Unlike edit decision lists, playback plans are computed automatically based on available links, filter conditions, and previous behavior of the user. The playback plans themselves encode logic for deciding when to follow links and when not to. Some plans are a generalization of hypervideo link behaviors [8] – that is they determine what happens when the beginning or ending of a source or destination link anchor is encountered during playback. Other plans may represent filters that skip over undesired portions of a video such as silence. More complex plans can be combinations of such behaviors. Playback plans may also alter the playback speed.

We identified playback plans that are appropriate to particular use scenarios. Table 1 details six expected contexts of use for playback plans, the information goals of the user in these contexts, and the playback plan behavior. This is not meant as a complete list of use contexts. Currently, HyperMeeting only implements playback plans for the more common scenarios that include automatically following links to newer meetings, either for a whole meeting or a single link, in combination with a filter for silence. Additional plans require improved detectors for the context of use.

Selecting a playback plan takes into account participation in previous 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.

Another default plan would be for people who participated in a series of meetings and wished to review comments linked to meeting they attended. In this case the source anchors in the meeting they attended would be played before playing the linked comment, but any links to previous meetings encountered while playing the source anchors would be ignored.

Figure 4 shows a playback plan for playing both halves of link “1” and then stopping, as would happen if a user clicked on the link without automatic link-following enabled. If automatic link-following were enabled, link “8” would be played, too, as indicated by the dashed line.

Context of Use	Likely Viewing Behavior	Plan Behavior
Reviewing a meeting that one participated in	Look for specific segments based on topic	Play segments on a topic
Reviewing multiple meetings that one participated in or has viewed	Look for related segments across meetings based on topic	Play segments on a topic with embedded links to additional related discussions
Viewing a meeting where one was absent	View the whole meeting	Play the meeting
Viewing a meeting where one was absent but at prior meetings	View the whole meeting in the context of prior meetings	Play the meeting with contextual segments from prior meetings
Viewing multiple related meetings where one was absent	View all the meetings with an understanding of relations between meetings	Play the oldest meeting in its entirety along with links to related content in newer meetings, then play additional content in newer meetings
Viewing multiple related meetings where one was absent but at some prior meeting	View all the unseen meetings with an understanding of relations between these meetings and their relations to prior meetings	Play the absent meeting in its entirety along with links to related content in newer meetings, then play additional content in newer meetings together with contextual segments from prior meetings

Table 1. Types of playback plans.

All playback plans can include filter criteria. 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. Attendees of a future meeting would want to skip that period of silence. Other filters could focus on particular topics or speakers.

Plans can also be personalized based on who is speaking in each segment. Thus, a plan could play 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.

4.6.2 *Execution of Playback Plans*

Figure 3 shows the execution order of a plan for playing the oldest meeting and all links to newer meetings. The circles indicate hyperlinks with the numbers matching the two halves of a link. The colored areas surrounding the circles are the link anchors. Superimposed as a gradient line to indicate playback progress is a plan for playing the oldest meeting in the chain while automatically following all links to newer meetings. Note that link "12" to the newest meeting at 16:52 causes out-of-order playback. After returning from link "9", link "12" is followed, then links "3", "10", "4", and "11"

Manual navigation within a playback plan, i.e., skipping to a part of the video that is covered by the plan, just continues the plan from that position in the video. Other user interaction could include clicking on a link circle while playing the source of that link. Such an interaction skips the playback to the destination of that link and continues the plan from there. Navigation to a part of the video not covered by the current playback plan leads to the computation of a new playback plan including that part of the video, by default a plan that plays the video of the current meeting and follow links to all later meetings.

4.7 System Architecture

For the first version of our system, we focus on distributed meetings conducted as a video conference. That approach makes it easy to get a video recording of each attendee and allows for simple speaker segmentation. In such a scenario, multiple video streams are recorded that are played together when the recorded meeting is being watched at a later time. In the future, we plan to support other types of meetings such as local meetings recorded by a single video camera or video conferences between two hubs. Presenting slides to remote attendees is important (e.g., [12]). Such a feature could be integrated into HyperMeeting via screen sharing similar to [4].

Figure 5 shows the primary components of the system. The HyperMeeting client runs in standard web browsers as a single page application. 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 live video conference among the distributed meeting attendees is maintained with WebRTC (Web Real-Time Communication)[1] peer-to-peer connections. Our HyperMeeting server manages the initial signaling to set up those connections. Clients keep a Websocket connection open to the HyperMeeting Manager on the server to maintain meeting state. Initially, each user is in a separate meeting and may join another existing meeting. Video playback is synchronized among users in the same meeting by clients sending playback change messages to the HyperMeeting Manager that are forwarded to other clients in the same meeting. Between changes, each client maintains the requested playback speed.

The HyperMeeting server stores the recorded video and images in the file system to support random access during playback. Other metadata such as the list of participants of a meeting is stored in a database. WebRTC is still an evolving standard that is currently supported in the Chrome and Firefox web browsers. Only Firefox currently provides an implementation of the Media Recorder API.[2] We created a workaround for Chrome to record audio with the Web Audio API[3] and video as a sequence of images. Chunks of the recorded media streams are uploaded to the HyperMeeting server and appended to the video files. After recording is completed, video is transcoded to a more compressed format. In the future, we may use a video conference multi-point control unit with recording capability to avoid the need for sending video content twice. However, it would be important in such a setup to get accurate timestamps to synchronize later playback.

There are several challenges when using a web-based architecture. While the HTML5 video tag as implemented in modern web browsers offers good video playback capabilities, it does not offer as much control as a standalone video player. In particular, it is difficult to control the caching behavior. Preloaded videos may tie up all socket connections to the server until they are completely downloaded. Instead, the system repeatedly attempts to play each video and pauses it in response to the "playing" event. We switched to using a Websocket for communication with the server in part because it works even if all socket connections are tied up.

We also encountered audio clipping at higher playback rates. It is unclear whether the problems were caused by the video player or by our speaker segmentation that muted all but one player. The latter was necessary because there is no echo cancelation for other sources, i.e., the playback of recorded video during a video conference. Also, too frequent synchronization negatively impacts playback. Only correcting drift of more than two seconds is a reasonable compromise. The user evaluation described later in the paper shows that improvements in these areas would be desirable.

4.7.1 Speaker Segmentation

Speaker segmentation serves a utility role in several parts of the system. We currently assume that each attendee has a microphone set up such that the audio from the meeting playback or co-located speaker is not picked up by the microphone. Thus the task of segmenting the audio based on speaker is reduced to determining when each speaker is not silent. We use a low-pass filter and a sampling window for audio energy to eliminate short noise artifacts. We compare the audio energy windows for all attendees and require a minimum speech duration to allow a speaker change.

In the user interface, it is useful to display a timeline showing when each of the attendees is speaking. This allows the viewer, for example, to skip to comments made by a particular person. The speaker segmentation is visualized as red and green lines above the timeline in Figure 2. The "skip silent" check box lets users control whether to automatically skip segments where nobody is speaking (e.g., 15:43:20 to 15:46:15 in Figure 2).

Speaker segmentation is also useful when playing several video streams depicting meeting attendees in parallel. Unless all recorded attendees wore headphones, sounds from the playback of even earlier meetings may be included in the recording. Current echo

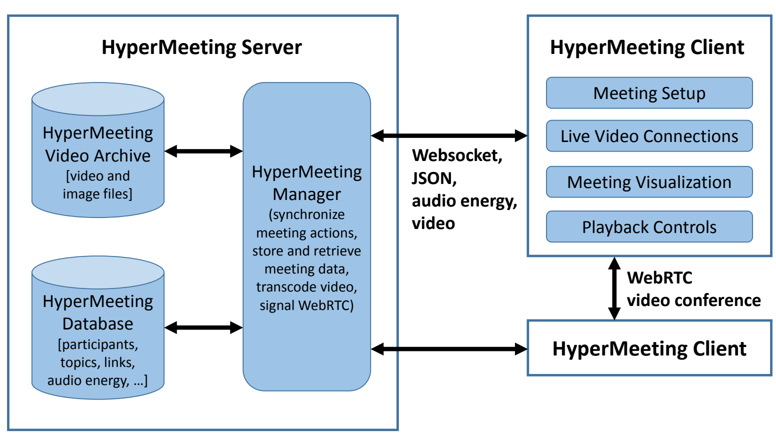

Figure 5. HyperMeeting Architecture

[1] http://www.w3.org/TR/webrtc/
[2] http://www.w3.org/TR/mediastream-recording/
[3] http://www.w3.org/TR/webaudio/

cancelation systems in web browsers remove sounds from the other end of a video conference but not those produced by a video player running at the same time. If several of those "tainted" video streams are played in parallel, slight differences in timing may cause echoes that can be suppressed by unmuting only the video player of the detected speaker.

5. Experience from Longer-Term Use

Our project team has used HyperMeeting over the course of its development. Since one of the members of our team was remote and other members had significant amounts of travel, it was not always possible to hold synchronous meetings. We used the HyperMeeting system to record design meetings so that people who did not attend could later review and add their comments. All team members attended at least some of the meetings, and there was typically a conclusion meeting in which all members were present.

Our experiences showed that there is a significant learning curve in the use of HyperMeeting. One problem is due to differences with common videoconferencing software such as Skype and Google Hangouts. In particular, these other tools provide no insight into the shared synchronized watching of previously recorded content. Issues such as who will drive the playback controls lead to communication among participants about this control. For example, it was common to explicitly say *"I will pause now"* when one of the participants wanted to add a link to a meeting they were reviewing. Similarly, when resuming playback of a recorded meeting after making comments, *"let's take that link to see what they said"* was a common utterance.

However, within a few months of experience, use of HyperMeeting changed distributed group interaction. One example of this is that meeting participants took advantage of knowing that others would review the current meeting. They called out expected future viewers to ask them questions or to respond to some opinion. For example, *"[Name], can you remind me why the system does this when I do that?"* or *"[Name], how can we examine the effect of the hyperlinks on information seeking?"* The request for feedback from someone not in the current meeting became natural in this context. This is something that the system could support more directly by either having people tag such questions/responses or have the system automatically detect mention of non-participating team members.

Due to the learning curve we experienced in recording and linking meetings with HyperMeeting, we decided to focus our first user study on the review of previously recorded meetings. We were aware of issues with link traversal. Through the study, we hoped to better understand their utility and usability. Once these issues are solved we plan to do a longitudinal study of the system that will cover both the creation and review of HyperMeeting.

6. User Evaluation

One research question focused on *how does allowing people to review content in a non-linear way influence the process of navigating through and understanding a series of past meetings*. To gain insights into user behavior with the system, we conducted an evaluation of HyperMeeting that focused just on the playback of a meeting chain. The participants were placed in the role of a decision maker who did not attend any meetings in the chain. We chose to look only at that part of the system because it covers features novel to most users and is amenable to a controlled study.

We particularly wanted to examine the effects of two novel design decisions described in Section 4:

1. The inclusion of hyperlinks between meetings rather than just facilitating synchronized distributed viewing of prior meetings and recording of new content.
2. The inclusion of playback plans (notably automatic link-following) rather than requiring users to control all navigation.

6.1 Evaluation Method

We recorded hypermeetings for training and for two study tasks. Each hypermeeting consisted of a chain of three meetings with an attendee pattern of AB, CD, AB. The training hypermeeting and one of the task hypermeetings included hyperlinks. We edited a copy of the training hypermeeting to remove links and periods of silence so that it could be used for training the no-link condition.

Two pilot participants helped us refine the study setup. In response to their interaction with the study materials, we modified some hyperlinks to avoid overlapping links. Because they had issues with automatic link-following, we decided to turn that feature off for half the participants. We also divided tasks into browsing and question-answering parts, improved the instructions, and added questions about the meeting content to test the participants' retention of the presented material. There was no significant effect of the conditions on retention so we do not further discuss that aspect of the study.

Eight participants (three female; ages 30 to 60) were selected from members of the research lab with no connection to and no experience with HyperMeeting. For their participation they received a chocolate bar. These participants were asked to view two series of three interrelated meetings with a goal of understanding the content of the meetings and answering questions about who said what. The study was a within-subjects design where each participant used both versions of HyperMeeting (with and without links). The order of the versions was counterbalanced. Only half the participants were given a system with automatic link-following enabled.

Each session lasted about 1½ hours consisting of two halves. In each half, participants were given 10 minutes to read written training materials and to play with the user interface to explore training meeting content unrelated to later tasks. They then performed a task with one version of HyperMeeting. In each task, the participants spent up to 15 minutes to review a series of three meetings related to topics that they would have interest in (policy decisions for the research lab related to employee bonuses and travel funding). They then had 15 minutes to answer questions about the meeting contents and could use HyperMeeting to help locate particular content. No note-taking was allowed during the first 15 minutes to better observe navigation behavior during the second 15 minutes. Finally, they answered a questionnaire with the ten-item System Usability Scale [2] to assess overall usability of the system. It also included two Likert-scale questions about the ease and enjoyability of performing the task using the system (i.e., "It was easy to perform the task using this tool to review meetings" and "I enjoyed performing the task using this tool."). Finally, to assess the utility of the hyperlinks in particular, in the condition with links, participants rated the question "Were hyperlinks useful?" on a Likert scale. The questionnaire also contained open-ended questions about the likes and dislikes of the system, and of hyperlinks in particular.

The software was configured to generate a log of user activity in the interface and associated video playback. Finally, one or two team members observed each activity to note behaviors and difficulties that might not come up in user comments.

6.2 Evaluation Results

We compared the participants' experience using the linked version of the system (either with or without automatic link following) with their experience viewing an unlinked series of meeting videos. Here we report results from the questionnaire responses given, open-ended feedback on the two system versions, and behavioral patterns observed in the logs and during the evaluation.

6.2.1 Reactions to hyperlinks and playback plans

A primary evaluation goal was to examine the utility of the hyperlinks and playback plans as a mechanism for navigating content in a new way. Participants' responses to the question "What did you like and dislike about the hyperlinks for reviewing this meeting?" yielded insights into positive and negative aspects of the design of hyperlinks.

Study participants clearly found the hyperlinks useful — 7 of the 8 participants agreed or strongly agreed with the statement about the usefulness of the hyperlinks. Two participants who were given the meeting with hyperlinks first complained about not having them in the second series of meetings: "*Felt like I watched three unrelated meetings. Threads weren't there and I couldn't feel consensus*" (P8) and "*It gets boring to listen through each topic without getting a direct response on each point. Harder to follow each discussion point when chunks of discussion is larger*" (P2).

Other positive comments showed that links allowed users to follow the discussion on a topic. For example, "*It supported topic to speaker navigation very fluidly and helped identify who said what much more easier and in context*" (P4). Links also highlighted important parts of the discussion: "*could easily move to related comments"* (P7) and "*I felt if I listened to linked material, I covered what was important*" (P8).

Although users tended to like hyperlinks, there were clear problems with usability across the different playback plans. A common complaint was feeling lost, meaning not knowing where in the meeting chain you were after following hyperlinks. This was commonly mentioned among users with the automatic link following, and well-summarized by the comment that "*Automatic link following is good for getting related thoughts together, but also leads to a feeling of being lost*" (P1).

Users with manual link-following also felt uncertain about where they were taken when following a link. "*I couldn't tell when the system would jump back out of a hyperlink or when it would pause playback*" (P4). Another explained that "*initially what happened when you click on a future link was counter-intuitive. Not obvious which part of time-line is sensitive to what (what the action will be when you click)"* (P7). P4 suggested*. "Maybe in sequential play / hyperlink do not start from start of segment"* and *"Indicator when a jump back is about to go.*"

Users made other interesting suggestions for improvements to link behavior. One user mentioned visualizing parts of the video that had been seen and optionally skipping already seen material: "*First time viewing, not showing what has been seen or skipping those parts*" (P5). They also suggested a back button for the automatic case because "*Sometimes want to go back to a previous point in the thread, not easy*" (P5).

Two users suggested another visualization of links besides the current numbering system. As one said, *"The links could be visualized differently. They keep on changing numbers which was kind of confusing. Maybe just a different numbering system"* (P2). Another user suggested alternative ways of visualizing the navigation such as using text *"What does the number represent? Perhaps a textual link might be more descriptive"* or a graph *"Need a graph for the links. Could not make sense of the structure"* (P6).

6.2.2 Usability of linked and unlinked versions

Participants rated the overall usability of their experience using the system with and without links. On the 100-point System Usability Scale (where a higher score means greater usability), the mean rating of the non-linked version was 74.38 (SD=14.19) and the mean rating of the linked version was 66.56 (SD=9.06). However, this difference was not statistically significant ($t(1,15)=1.80$, $p=.20$). There were no statistically significant differences between the linked and unlinked versions on individual scale sub-items (scale 0 to 4), with the exception of the rating of complexity. Here, the linked version of the system was deemed to be significantly more complex (M=3.375) than the unlinked version (M=2.625) ($t(1,15)=5.47$, $p=.03$). The difference in the rating of complexity comes from the condition with automatic link-following (M=3.50 vs. M=2.25). There is no significant difference in the other condition. There were no significant differences between linked and unlinked systems in terms of rating of ease of use and enjoyability. There were no significant order effects in the usability ratings with respect to performing the task with hyperlinks first.

Given that participants only used the systems for a short amount of time, it is possible that initial attitudes about usability and complexity could evolve and change over time. However, responses to the open-ended questions about what participants liked and disliked about each version of the system provided more information about how the links (or lack thereof) influenced the experience of reviewing meetings.

6.2.3 Other usability observations

Participants liked several features of the system particularly those that allowed for speedup and skip silences (P1, P3, P5, P8) and those that allowed for color-coding to identify speakers (P1, P6).

There was disagreement on the animation that zoomed in on the video of the speaker: this was mentioned as both a "like" *"Resizing speaker video to highlight who is talking"* (P6) and a "dislike" *"changing video size seemed to busy"* (P1), *"Video Zoom for active speaker was distracting"* (P4).

Several participants requested more fine-grained topics in the non-link condition: *"Topics not fine enough in granularity"* (P5), and in particular they *"would like to search for keywords"* (P7).

Finally, several participants had issues with audio clipping: *"Sometimes on change of speaker it seemed like I lost a bit of audio"* (P1), *"In fast forward mode, some of the audio is skipped"* (P6), *"Audio clipped too much so parts missing"* (P7). There was also some lag: *"lag in playing video, unable to scrub"* (P7)

6.3 Log Analysis

The analysis of the usage log revealed how and when participants used the different features of the system. All but one participant used the "skip silent" feature, and all participants increased the playback speed. There were differences in using the content navigation-based system features (links, topics, and directly clicking on the timeline) depending on what was available. Because each task was separated into two halves, the exploration of the meeting chain and goal-oriented navigation to answer questions, the strategies and features used differ in those halves.

Navigation with links: Links were used more frequently for exploration than for answering specific questions. In the exploration phase, two participants with automatic link-following (P1 and P6) started playback at the beginning of the oldest meeting and watched the complete link chain without manual navigation while automatically following 23 links. The other six participants clicked on links about the same number of times in that half (M=9.0; SD=2.4). In the goal-oriented (question answering) phase, links were used sparingly (M=1.1; SD=0.9). In that half, automatic link-following happened unevenly among the participants that had it (M=3.3; SD=2.2).

Navigation by topics: In the no-link condition, navigation by clicking on a topic was used more. It occurred more in the exploration phase (no link: M=5.3; link: M=1.3) than in the question-answering phase (no link: M=2.1; link: M=1.3).

Navigation by clicking on the timeline: Directly clicking on the timeline to jump to various points in the discussion was about twice as frequent in the question-answering phase (no link: M=21.7; link: M=27.3) compared to the exploration phase (no link: M=10.9; link: M=14.3).[4]

In summary, links were used frequently in the exploration phase. In their absence, navigation by topics was used instead. Links and topics were only used sparingly in the question-answering phase. As with links, topics seem to be better suited for exploration. In the question-answering phase, clicking on the timeline was twice as frequent as in the earlier phase. This seemed to be preferred for finding specific answers. That may explain the comments of some participants that topics were not fine-grained enough.

6.4 Discussion

It is encouraging that the study participants found hyperlinks useful for navigating the video recording of a meeting chain. Several participants commented that they missed them in the condition without hyperlinks. Log analysis shows that links were used for exploration if present and that otherwise topics were used as proxy. At the same time, the study uncovered usability issues with hyperlinks and automatic link-following in particular. To address this issue, several participants suggested visualizations of the link network or of content that they already watched. One participant requested a back button similar to that in a web browser. Such a feature is similar to the navigation stack described in [18].

For visualizing the link network, our first idea can be seen in Figure 3. However, the many crossed paths would make such a visualization non-intuitive. An alternative would be to make the timelines nonlinear such that most corresponding link anchors line up and connecting lines would be mostly vertical. Such a visualization may make the link behavior more intuitive because the users could predict where the link would go and whether playback would stop at the end of the link anchor. Indicating previously watched content or even skipping it by default are good suggestions from our participants. We will integrate this into a future version of the system.

The labeling of links with numbers turned out to be another issue. Our system supports link labels that appear as tooltips on mouse-over but we did not create link labels when we recorded the meeting chains. We believe that many other users will not create link labels, either, so that a less laborious solution would be desirable, maybe using automatic speech recognition. One participant expected the links covering the same content to have the same number. Currently, numbers are assigned sequentially so that a link from the second to the third meeting would not have the same number as a link from the first to the second meeting even if the links line up at the second meeting. Clustering links by time may be a solution for this issue.

Some participants made extensive use of topics for navigation, especially in the condition without links. However, several participants commented that the topics were too coarse. Displaying keywords determined by automatic speech recognition might be an approach for more fine-grained navigation support.

Several participants noticed problems with audio clipping. We decided to use our speaker segmentation to mute all but one video player. That deals with an issue where background audio from playing previous meetings may be present in the recording, leading to echo if more than one instance is played. This can be avoided if all participants wear headphones but it would be challenging to remove it with echo cancelation because of variable latency. It appears that our speaker segmentation is not precise enough to avoid audio clipping. One possible solution would be to unmute the next speaker slightly earlier than muting the current speaker.

The lag observed by a pilot participant and one of the other participants is likely due to caching and network latency. Worse network connectivity would impact it negatively. Unfortunately, our web-based solution does not offer much control over video caching. Keeping a second video player in the background that skips ahead to likely playback positions may be a possible solution. This lag is also a reason why we did not implement video scrubbing requested by some participants.

Finally, as mentioned in Section 5, the current HyperMeeting system seemed to best support a series of three or four meetings. Issues of scalability could arise as the historical record grows beyond this number of meetings, making navigation and information finding more complex and difficult. Designing for longer-term archiving and use of the hypervideo system will require adaptation of the interface and is out of the scope of this initial study, but remains a fruitful area for future work.

7. Conclusions

Meetings often overlap in terms of content. This causes difficulties as meetings that build on what was discussed in prior meetings result in barriers for those not at prior meetings and meetings that re-discuss topics are often viewed as unproductive by the participants that were at prior meetings. HyperMeeting was designed to support geographically-distributed teams that need meetings to make decisions and coordinate without requiring them to be available and on-line at the same time.

HyperMeeting combines in a single interface several novel features that integrate meeting capture, synchronous playback of meeting content, and the recording of new meetings and extensions to prior meetings. As a result, it is possible to use the playback activities (e.g., pausing a recorded meeting) to automatically generate hyperlinks between recordings. A critical component of this is the ability to synchronize playback and annotation by distributed participants.

During playback, automatic link-following guided by playback plans present the relevant content to the users. An analysis of usage scenarios identified a variety of viewing behaviors appropriate as responses to navigational actions based on the user's meeting attendance and viewing history. Playback plans take into account

[4] Excludes outlier P8 who clicked the timeline in one half seven times as often as the mean of the others.

features of the user along with features of the content to present just enough previously viewed content to contextualize new content.

Multiple months of use of HyperMeeting within our own geographically-distributed team showed that users developed new meeting practices based on the ability to review prior meetings and call out to attendees of future meetings. These practices improved understanding of meetings that group members were not at and increased shared understanding within the group.

A user study examined the viewing of hypermeetings. The results show that participants generally liked having links. Log analysis supported user feedback of the value of links. Without them, users relied on topics but found them too coarse-grained. Users found navigating back-and-forth between multiple videos disorienting as documented in prior reports on the evolution of hypervideo playback interfaces [8].

HyperMeeting is a promising option for people who cannot attend synchronous meetings. Future work will address issues uncovered in the user study such as the need for refinements to the interface cues surrounding navigation. A future study will more formally assess the process of synchronously viewing prior meetings and recording extensions to that content. Longer term future work will explore the inclusion of roles for meeting participants (e.g., meeting chairs) and formal or informal results (e.g., official minutes and individual's notes).

8. REFERENCES

[1] J. Barksdale, K. Inkpen, M. Czerwinski, A. Hoff, P. Johns, A. Roseway, and G. Venolia. Video threads: asynchronous video sharing for temporally distributed teams. *Proc. CSCW '12,* 1101-1104.

[2] J. Brooke. SUS: A Quick and Dirty Usability Scale. In: P. Jordan, B. Thomas, B. Weerdmeester, and I. McClelland (Eds.), *Usability Evaluation in Industry*. London: Taylor & Francis, 1996.

[3] J.J. Cadiz, Anand Balachandran, Elizabeth Sanocki, Anoop Gupta, Jonathan Grudin, and Gavin Jancke. Distance learning through distributed collaborative video viewing. *Proc. CSCW '00,* 135-144.

[4] L. Denoue, S. Carter, A. Girgensohn, and M. Cooper. Building digital project rooms for web meetings. *Proc. DocEng '14,* 135-138.

[5] B. Dorn, L.B. Schroeder, and A. Stankiewicz. Piloting TrACE: Exploring Spatiotemporal Anchored Collaboration in Asynchronous Learning. *Proc. CSCW '15,* 393-403.

[6] L. Gericke, M. Wenzel, and C. Meinel. Asynchronous understanding of creative sessions using archived collaboration artifacts. *Proc. Collaboration Technologies and Systems 2014,* 41–48.

[7] W. Geyer, H. Richter, and G.D. Abowd. Towards a Smarter Meeting Record—Capture and Access of Meetings Revisited. *Multimedia Tools Appl.* 27(3), 2005, 393-410.

[8] A. Girgensohn, L. Wilcox, F. Shipman, and S. Bly. Designing affordances for the navigation of detail-on-demand hypervideo. *Proc. AVI '04,* 290-297.

[9] R. Grigoras, V. Charvillat, and M. Douze. Optimizing Hypervideo Navigation Using a Markov Decision Process Approach. *Proc. MM '02,* 39-48.

[10] S. Hunter, P. Maes, S. Scott, and H. Kaufman. MemTable: an integrated system for capture and recall of shared histories in group workspaces. *Proc. CHI '11,* 3305-3314.

[11] F. Lehner, M. Langbauer, and N. Amende. Measuring success of enterprise social software: the case of hypervideos. *Proc. i-KNOW '14,* Article 3, 9 pages.

[12] G. Mark, J. Grudin, and S.E. Poltrock. Meeting at the desktop: an empirical study of virtually collocated teams. *Proc. ECSCW '99,* 159-178.

[13] T.P. Moran, L. Palen, S. Harrison, P. Chiu, D. Kimber, S. Minneman, W. van Melle, and P. Zellweger. "I'll get that off the audio": a case study of salvaging multimedia meeting records. *Proc. CHI '97,* 202-209.

[14] M. Nathan, M. Topkara, J. Lai, S. Pan, S. Wood, J. Boston, and L. Terveen. In case you missed it: benefits of attendee-shared annotations for non-attendees of remote meetings. *Proc. CSCW '12,* 339–348.

[15] J. Pan, L. Li, and W. Chou. Real-Time Collaborative Video Watching on Mobile Devices with REST Services. *Proc. MUSIC '12,* 29-34.

[16] M. Sadallah, O. Aubert, and Y. Prié. CHM: an annotation- and component-based hypervideo model for the Web. *Multimedia Tools Appl.* 70(2), 2014, 869-903.

[17] N. Sawhney, D. Balcom, I. Smith, HyperCafe: narrative and aesthetic properties of hypervideo. *Proc. HT '96,* 1-10.

[18] F. Shipman, A. Girgensohn, and L. Wilcox. Authoring, viewing, and generating hypervideo: An overview of Hyper-Hitchcock. *ACM Trans. Multimedia Comput. Commun. Appl.* 5(2), 2008, Article 15, 19 pages.

[19] J. Tang, J. Marlow, A. Hoff, A. Roseway, K. Inkpen, C. Zhao, and X. Cao. Time travel proxy: using lightweight video recordings to create asynchronous, interactive meetings. *Proc. CHI '12,* 3111-3120.

[20] C.A.B. Tiellet, A. Grahl Pereira, E.B. Reategui, J.V. Lima, and T. Chambel. Design and evaluation of a hypervideo environment to support veterinary surgery learning. *Proc. HT '10,* 213-222.