Searching and browsing live, web-based meetings

Scott Carter, Laurent Denoue, Matthew Cooper
FX Palo Alto Laboratory, Inc.
3174 Porter Drive
Palo Alto, CA 94304
carter,denoue,cooper@fxpal.com

ABSTRACT

Establishing common ground is one of the key problems for any form of communication. The problem is particularly pronounced in remote meetings, in which participants can easily lose track of the details of dialogue for any number of reasons. In this demo we present a web-based tool, MixMeet, that allows teleconferencing participants to search the contents of *live* meetings so they can rapidly retrieve previously shared content to get on the same page, correct a misunderstanding, or discuss a new idea.

Categories and Subject Descriptors

H.5.2 [**Information Interfaces and Presentation (e.g., HCI)**]: User Interfaces

General Terms

Design, Human Factors

Keywords

Teleconferencing, web, search

1. INTRODUCTION

Past work in teleconferencing recording and retrieval has focused on systems and methods to capture content in instrumented spaces and to access and search media post hoc [1], as well as methods to help users late to a meeting "catch up" to real-time [2]. On the other hand, less work has focused on searching and recovering material presented during lightweight, web-based meetings. This functionality can be particularly useful in a long meeting during which participants are sharing digital content via screensharing. For example, consider the following scenario: John, an interaction designer, is sharing his designs during a teleconference with Maria, a product manager. He uses screensharing to page through several designs constructed in a variety of different applications. Half an hour into the meeting, Maria
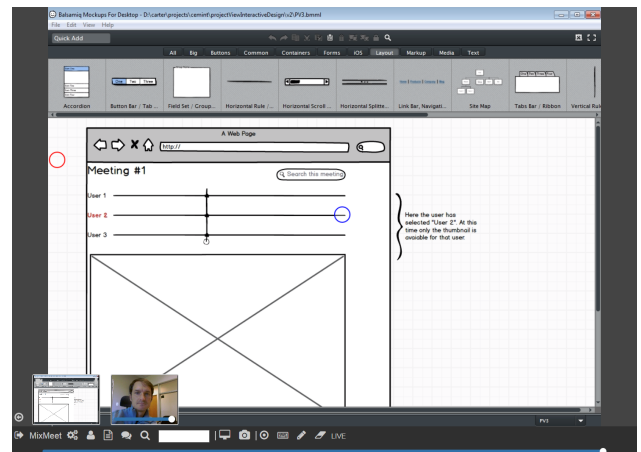
Figure 1: The MixMeet interface. Here, one user shares his webcam while the other shares a region of her screen. Participants can browse captured keyframes for each stream using range sliders. The slider of the currently selected, fullscreen stream is positioned at the bottom of the interface, while the other slider is positioned at the bottom of the remote participant's live, thumbnailed video.

realizes that a design John is currently showing would work best alongside another he showed ten minutes earlier. Using current web-based meeting tools, which provide support for searching through meeting content post hoc, Maria would have to verbally describe the content and hope that John could find it and re-share it.

MixMeet is a web-based teleconferencing tool we built that can simplify this process (Figure 1). With MixMeet Maria can quickly retrieve the design she is looking for either by searching (Figures 2, 3) or browsing (Figure 4). MixMeet creates a browsable and searchable keyframe archive per participant during live meetings. Frames from participants' streams are saved at regular intervals in real time during the meeting. Users can adjust a range slider on each participant's stream to view archived frames (Figure 4 left). At the same time, MixMeet utilizes live content analysis and interface techniques to make it possible to index and search content. Search matches are visualized either as marks on the appropriate range slider (Figure 3) or, when the user searches in a separate search panel, inline.
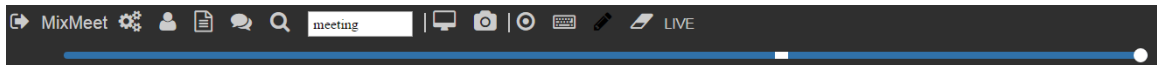
Figure 2: Searching a live meeting. The search term 'meeting' has a single hit, indicated in white on the main timeline.

2. INDEXES

MixMeet analyses video, audio, and metadata streams continuously to generate searchable content:

- *Speech* MixMeet generates a speech-to-text transcript on the fly using Google's Web Speech API running in continuous mode.

- *OCR* MixMeet can stream either webcams or screen regions. A Javascript-based OCR engine analyses screen content. Users connecting via a mobile device can take advantage of a server-side OCR engine.

- *Annotations* MixMeet also supports live, content-based annotations [3]. Annotated screen content is weighted preferentially for matching frames with search terms.

- *Mouse events* MixMeet records and forwards mouse events to remote clients. These events are also weighted preferentially during search.
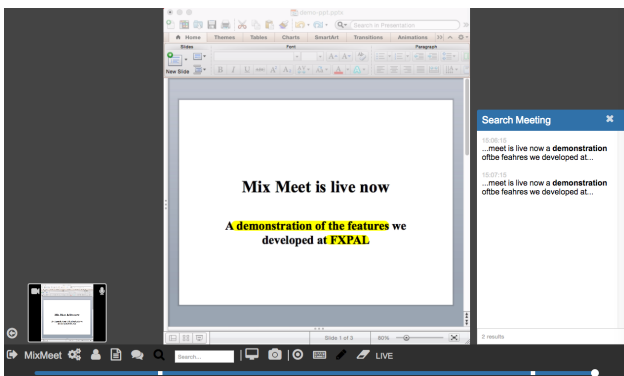

Figure 3: The search box shows OCR or Speech to Text that match the query 'demonstration'; for each result, the system shows the time and a small snippet of text around the match.

3. ORIENTATION-BASED BROWSING

Many remote participants use teleconferencing to meet with a group of collocated people. This scenario often involves a tablet device with a single camera that is rotated during the meeting so that the remote participant can see different co-located members or resources (e.g., whiteboards) in the room. These rotations can be controlled by the remote participant via a table-top device (such as a Kubi[1]) or by a local participant manually moving the tablet to frame other participants when appropriate. MixMeet listens to DeviceOrientationEvents that modern browsers generate to detect these changes and index archived keyframes appropriately. The client further bins these orientation events into
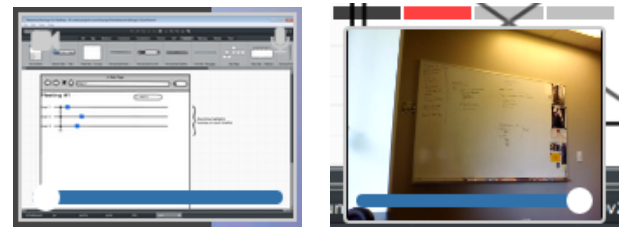
---

[1]https://www.revolverobotics.com/

Figure 4: (Left) Users can browse to frames archived earlier in the meeting with the range slider. (Right) Orientation events are automatically detected and used to help users navigate meetings. In this case, the user moved his tablet device away from his face and toward a whiteboard. The remote interface automatically generated an interactor showing previous directions the camera had pointed during the meeting (black rectangle), the current direction (red), and directions not yet used in this meeting (light gray). Clicking on any black rectangle shows the most recent frame from that direction. Clicking on the red rectangle sets the stream to the live view.

four categories: front, left, right, and back. When a device has keyframes saved from at least two positions, a small interactor appears above the appropriate participant's stream allowing remote users to directly navigate back to the last keyframe stored from each direction (Figure 4 right).

4. CONCLUSIONS AND FUTURE WORK

Using content analysis, indexing, and search and browsing interfaces, MixMeet can make content shared in live meetings more useful. As work practices become more distributed, the context of meetings will continue to evolve, and the types of devices people use to follow meetings remotely will change. As such, we are investigating ways for people to search, browse, and otherwise actively participate in live meetings from other interfaces, including wearables.

5. REFERENCES

[1] Cutler, R., Rui, Y., Gupta, A., Cadiz, J.J., Tashev, I., He, L., Colburn, A., Zhang, Z., Liu, Z., and Silverberg, S. Distributed Meetings: A Meeting Capture and Broadcasting System. *Proceedings of ACM Multimedia*, 503–512. 2002.

[2] Junuzovic, S., Inkpen, K., Hegde, R., Zhang, Z., Tang, J., and Brooks, C. What Did I Miss?: In-meeting Review Using Multimodal Accelerated Instant Replay (Air) Conferencing. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 513–522. 2011.

[3] Denoue, L., Cooper, M., Girgensohn, A., and Carter, S. Building digital project rooms for web meetings. *Proceedings of the ACM Symposium on Document Engineering*, 135–138. 2014.