THE DISTRIBUTED MEETINGS SYSTEM

Ross Cutler, rcutler@microsoft.com

Microsoft Research, Redmond, Washington, USA

ABSTRACT

Meetings are an integral part of everyday life for most work-groups. However, due to travel, time, or other constraints, people are often not able to attend all the meetings they need to. Teleconferencing and recording of meetings can address this problem. We describe a system that provides these features, as well as a user study evaluation of the system. The system uses a variety of capture devices (a novel 360º camera, a whiteboard camera, an overview camera, PC graphics capture, and a microphone array) to provide a rich experience for people who want to participate in a meeting from a distance. The system is also combined with speaker clustering, spatial indexing, and time compression to provide a rich experience for people who miss a meeting and want to watch it afterward.

1. INTRODUCTION

Meetings are an important part of everyday life for many work-groups. Often, due to scheduling conflicts or travel constraints, people cannot attend all of their scheduled meetings. In addition, people are often only peripherally interested in a meeting such that they want to know what happened during it without actually attending; being able to browse and skim these types of meetings could be quite valuable.

This paper describes a system called Distributed Meetings (DM) that enables high quality broadcasting and recording of meetings, as well as rich browsing of archived meetings. DM has a modular design and can use combinations of a variety of input devices (360º camera, overview camera, whiteboard capture camera, PC graphics capture, and microphone array) to capture meetings. For live meetings, the system broadcasts the multimedia meeting streams to remote participants, who use the public telephone system for low-latency duplex voice communication. The meetings can also be recorded to disk and viewed on-demand. Post-processing of recorded meetings provides on-demand viewers with indexes of the whiteboard content (e.g., jump to when this was written), PC graphics (e.g., jump to when this slide was displayed) and speakers (e.g., only show me the parts when this person speaks). On-demand viewers can also use time compression to remove pauses in the meeting and speed up playback without changing the audio pitch of the speakers.

A typical usage of the DM system is as follows: A meeting organizer schedules a meeting using MS Outlook. In the meeting room, the DM kiosk is used to start the meeting broadcast and recording. Remote participants run a DM client to see video of the meeting, and use the telephone for audio. After the meeting is done, the organizer stops the meeting using the kiosk. After the meeting is postprocessed, an email with a link to the recorded meeting is sent to the meeting participants scheduled with Outlook. People who missed the meeting can use the DM client to view the meeting, which has additional indexing and time compression.

This paper will first describe the system, a user study for the meeting recording scenario, and then discuss conclusions and future work. A more complete description of the system is given in [2].

2. RELATED WORK

There are several commercially available video conferencing systems (e.g., Polycom, Tandberg), and several academic meeting recording systems (e.g., [1][4]). A significant difference in the DM system is the use of omnidirectional video, which allows multiple remote users to view any local participant simultaneously, not just the remote speaker [5].

To better index recording meetings, we use both audio (sound source localization) and video (person detection/tracking) to robustly determine the current speaker [2], which can be used to index and filter the meeting playback. The whiteboard image stokes and keyframes are timestamed [6], which are used for spatial indexing of the meeting. The PC graphics keyframes are also computed and used as an index. Finally, we use time compression [3] to speed up meeting playback.

3. SYSTEM DESCRIPTION

A typical room configured with the DM system is shown in Figure 1. It consists of the following hardware devices:

- RingCam 360º camera (Figure 2)
- Whiteboard capture camera (Canon G2)
- Overview camera (640x480 1394 camera)
- PC graphics capture card (VisionRGB)
- Meeting room server (a dual CPU workstation) with a touch screen display (kiosk).

The DM system architecture is shown in Figure 3. Remote participants are able to view live video via the Intranet, but use the telephone network for audio communication. Archived clients use the Intranet for both audio and video.

3.1. Meeting Room Server

The dataflow for the meeting room server is shown in Figure 6. The input devices are the RingCam, overview camera, whiteboard camera, PC graphics capture, and microphone array (Motu828). The server runs Windows XP and is implemented using C++ and DirectShow. Each node in the dataflow is a DirectShow filter and is described in [2].

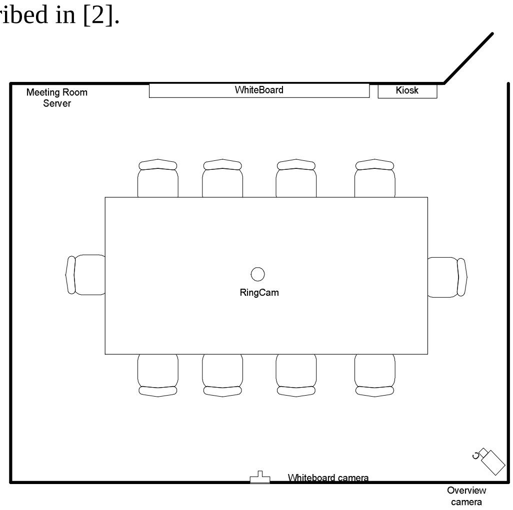

Figure 1: DM room diagram, which contains a RingCam, whiteboard and overview camera, PC graphics capture, meeting room server and kiosk.

The network bandwidth for the live client currently implemented is given in Table 1. This bandwidth is prohibitively high on typical networks, but it can be significantly reduced by using a better codec (WMV 9 instead of MJPEG). Another significant decrease in bandwidth is achieved by sending a low resolution version of the panorama and a high resolution speaker view, rather than the entire high resolution panoramic image. These two changes reduce the bandwidth from 4682Kbps to 513Kbps, a much more practical bandwidth.

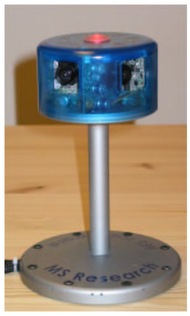

Figure 2: RingCam: an inexpensive omnidirectional camera and microphone array designed for capturing meetings.

3.2. DM Client

The DM Remote Client supports both live and asynchronous viewing of meetings. The user interface for the archived client is shown in Figure 5. The live client is similar, but does not include the timeline or whiteboard/PC graphics key frame table of contents.

A low resolution version of the RingCam panorama image is shown in the lower part of the client. A high resolution image of the speaker is shown in the upper left, which can either be automatically selected by the virtual director or manually selected by the user (by clicking within the panoramic image).

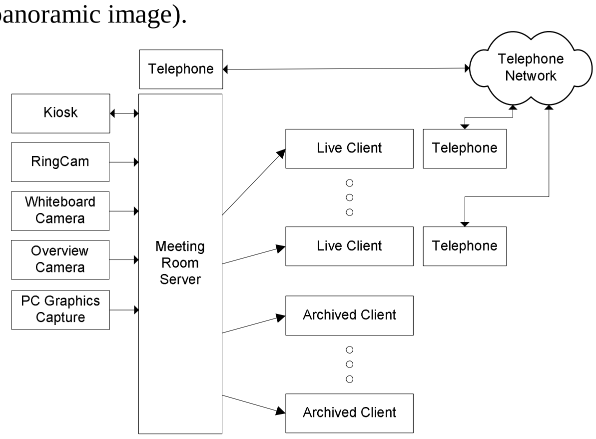

Figure 3: The DM architecture. Meetings are captured and broadcasted by the meeting room server, and stored for offline access by the archived meeting server.

The whiteboard image is shown in the upper right window. Each pen stroke is timestamped, and clicking on any stroke in the whiteboard synchronizes the meeting to the time when that stroke was created. Pen strokes that will be made in the future are displayed in light gray, while pen strokes in the past are shown in their full color. Key frames for the whiteboard are displayed to the right of the full whiteboard image and provide another index into the meeting. The transparency of the current key frame and the current image can be adjusted so that remote viewers can even view pen strokes occluded by a person.

The PC graphics capture is displayed by toggling the WB/PCG tab. Keyframes for the PC graphics capture will be the slides shown during a PowerPoint presentation.

The timeline is shown in the bottom of the window, which shows the results of speaker segmentation. The speakers are automatically segmented and assigned a unique color. The person IDs have been manually assigned, though this process could be automated by voice identification. The remote viewer can select which person to view by clicking on that person's color. The speakers can also be filtered, so that playback will skip past all speakers not selected.

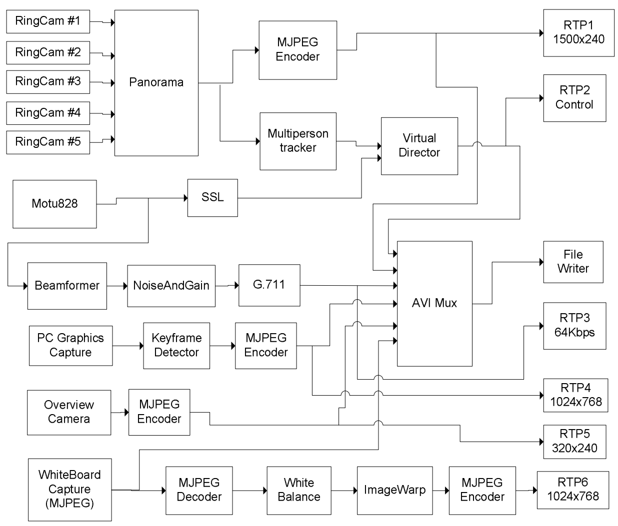

Figure 4: DM Server Dataflow

The playback control section to the left of the panorama allows the remote view to seek to the next or previous speaker during playback. In addition, time compression [3] can be used to remove pauses to and increase the playback speed without changing the speaker's voice pitch.

Just above the playback control is the tab control, which allows the user to display meeting information (time, location, duration, title, and participants), meeting statistics (who led the meeting, number of active participants), the overview window, and whiteboard statistics.

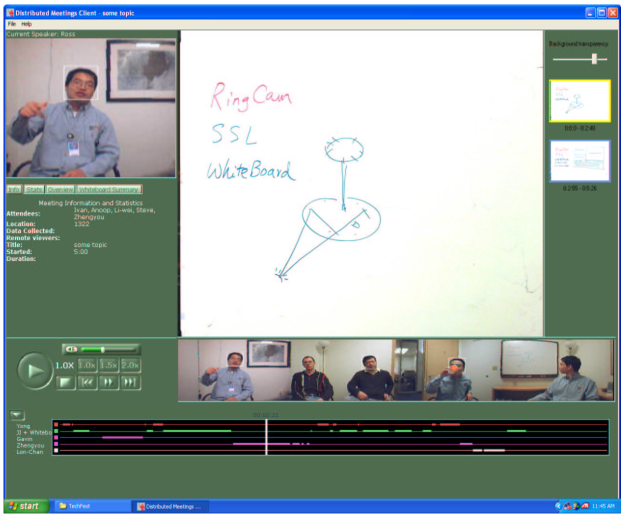

Figure 5: DM archived meeting client: Panorama window (bottom), speaker window (upper left), whiteboard (upper right), timeline (bottom).

Figure 6: Meeting Room Server dataflow

Stream	Width	Height	FPS	Mbits/s
RingCam	1500	240	15	4.120
Overview	320	240	7.5	0.439
Whiteboard	1024	768	0.1	0.060
Audio				0.063
				4.682
1 Hour storage (GB)	2.06			

Table 1: DM Bandwidth using MJPEG

Stream	Width	Height	FPS	Mbits/s
RingCam	750	120	7.5	0.051
Speaker	320	240	15	0.300
Overview	320	240	7.5	0.044
Whiteboard	1024	768	0.1	0.060
Audio				0.063
				0.518
1 Hour storage (GB)	0.23			

Table 2: DM Bandwidth using WMV9 (future implementation)

4. USER STUDY

To evaluate the usefulness of the DM meeting capture system, a user study was conducted. In this study, 10 meetings were recorded from 10 separate MSR groups. At least one group member was absent from the meeting (often due to other reasons). A total of 11 meetings were viewed by absent members in a usability lab. In-meeting and offline participants were interviewed afterwards. The interview questions and results are shown in Table 3 and Table 4. Future details of the study are given in [2].

Question N = 10 groups	Avg	Std dev
I was comfortable having this meeting recorded.	3.9	0.7
The system got in the way of us having a productive meeting.	1.7	0.4
I felt like I acted differently because the meeting was being recorded.	3.1	1.1
It was awkward having the camera sitting in the center of the table.	3.0	0.8

Table 3: Survey responses from people who participated in meetings that were recorded. All questions were answered using the following scale: 5 = strongly agree, 4 = agree, 3 = neither agree nor disagree, 2 = disagree, 1 = strongly disagree

Question N = 11	Avg	Std dev
It was important for me to view this meeting.	3.7	0.5
I was able to get the information I needed from the recorded session.	4.6	0.5
I would use this system again if I had to miss a meeting.	4.4	0.8
I would recommend the use of this system to my peers.	4.0	0.9
Being able to browse the meeting using the whiteboard was useful	3.2	1.2
Being able to browse the meeting using the timeline was useful	4.0	0.9
Being able to speed up the meeting using time compression was useful	4.1	1.3
Being able to see the panoramic (360º) view of the meeting room was useful	4.4	0.9
Being able to see the current speaker in the top-left corner was useful	4.1	1.2

Table 4: Survey responses from people who missed a meeting and watched it afterward. All questions were answered using the following scale: 5 = strongly agree, 4 = agree, 3 = neither, agree nor disagree, 2 = disagree, 1 = strongly disagree

5. CONCLUSIONS AND FUTURE WORK

We have described a system to broadcast, record, and remotely view meetings. The system uses a variety of capture devices (360º RingCam, whiteboard camera, overview camera, PC graphics capture, microphone array), to give a rich experience to the remote participant. Archived meetings can be quickly viewed using speaker filtering, spatial indexing, and time compression. The user study of the recorded meeting scenario shows that users found the system captured the meeting effectively, and liked the panoramic video, timeline, speaker window, and time compression parts of the system.

We plan to greatly extend the capabilities of DM in many ways. Most importantly, we want to add duplex audio and video real-time communication over the intranet, so that the telephone network is not required. This is a challenging task, as it involves significantly lowering the audio/video latency, lowering the network bandwidth requirements, and adding echo cancellation suitable for microphone arrays.

6. ACKNOWLEDGEMENTS

We gratefully acknowledge the following people who contributed to the DM system (in alphabetical order): Alex Colburn, Anoop Gupta, Ben An, Dawson Yee, Gavin Jancke, Ivan Tashev, JJ Cadiz, Li-wei He, Lon-Chan Chu, Mehrdad Basseri, Michael Cohen, Mike Holm, Rico Malvar. Steve Silverberg, Vince Jesus, Yong Rui, Zhengyou Zhang, Zicheng Liu.

7. REFERENCES

[1] P. Chiu, A. Kapuskar, and L. Wilcox, Meeting capture in a media enriched conference room, *Multimedia Magazine*, vol. 7, no. 4, Oct-Dec 2000, pp. 48-54.

[2] R. Cutler, et. al. Distributed Meetings: A Meeting Capture and Broadcasting System, *ACM Multimedia*, 2002.

[3] L. He and A. Gupta. Exploring Benefits of Non-Linear Time Compression. *ACM Multimedia*, 2001.

[4] A. Waibel, et. al. Advances in Automatic Meeting Record Creation and Access, ICASSP 2001.

[5] Y. Rui, A. Gupta and J.J. Cadiz, Viewing meetings cap-tured by an omni-directional camera, Proc. ACM CHI'2001, Seattle, WA, April, 2001.

[6] L. He, Z. Liu and Z. Zhang. Why Take Notes, Use the Whiteboard Capture System. Technical Report MSR-TR-2002-89, September 2002.