Webcasting Made Interactive: Integrating Real-Time Videoconferencing in Distributed Learning Spaces

Conference Paper · July 2007

DOI: 10.1007/978-3-540-73354-6_30 · Source: DBLP

Alison Weir
University of Toronto

Webcasting Made Interactive: Integrating Real-time Videoconferencing in Distributed Learning Spaces

Ronald Baecker[1,2], Jeremy Birnholtz[2], Rhys Causey[1,2], Simone Laughton[3], Kelly Rankin[2], Clarissa Mak[2], Alison Weir[3], and Peter Wolf[2]

[1] Department of Computer Science and [2] Knowledge Media Design Institute, Univ. of Toronto, 40 St. George Street, Toronto, ON M5S2E4, Canada
[3]Univ. Toronto Mississauga, 3359 Mississauga Rd. N. Mississauga, ON L5L 1C6 Canada
rmb@kmdi.toronto.edu, jbirnhol@gmail.com, r.causey@gmail.com, slaughto@utm.utoronto.ca, kelly@kmdi.toronto.edu, clarissa.mak@utoronto.ca, aweir@utm.utoronto.ca, wolfpet@kmdi.toronto.edu

Abstract. This paper presents an extension to the ePresence Interactive Media webcasting infrastructure to support real-time voice and video conferencing for a few attendees while concurrently streaming an event to many others. We emphasize the tools for enhancing awareness of the remote attendees in the lecture hall or seminar room, and for facilitating communication between remote attendees and the lecturer. We present preliminary results of a field study of the use of this environment for a multi-campus university class.

Keywords: webcasting, streaming media, eLearning, digital media, digital video, videoconferencing, synchronous communications, awareness.

1 Introduction

Synchronous technologies such as video can link learners and instructors at the "same time, anytime and anywhere", allowing them to learn from each other [1]. Video learning systems provide enriched interaction through words, voice tone, environmental and nonverbal cues, and immediate feedback [2]. These technologies are used in academe and business [1,] [3], [4], but questions of technical implementation and pedagogical methods and effectiveness remain. We have been developing and studying an Internet streaming media infrastructure that enables remote attendance at learning events such as lectures, both concurrently and retrospectively, with maximum engagement, interactivity, and support for community learning.

Audio/web conferencing and videoconferencing are often used for real-time Internet communication and collaboration. Audio/web conferencing (e.g., www.webex.com) allows the real-time multipoint transmission of voice and slides, but lacks the media richness, sense of presence, and ability to engage participants afforded by dynamic visual media. Internet desktop video conferencing (e.g., www.microsoft.com/windows/netmeeting/) supports real-time multipoint audio and video communications as well as shared workspaces, but still does not provide reliable video performance at reasonable cost, and is not scalable to large numbers.

Our approach is based on a third technology known as *webcasting* — the Internet broadcasting of streaming media to be viewed via a Web browser on almost any suitably Internet-connected personal computer. Webcasting is scalable to large numbers of participants, but, due to buffering requirements, is typically a one-way broadcast medium with only limited provision for interaction via time-delayed text messages. Yet one could imagine a webcasting system that would allow significantly greater interactivity just in those cases where it is needed.

This paper presents an extension [5], [6][1] of our ePresence Interactive Media infrastructure [8-11] that supports real-time voice and video conferencing for a *few* attendees while streaming an event to *many* others[2]. We also report preliminary results of a field study of using the environment in a multi-campus university class.

2 Background and Previous Work

To allow the scalability and low barriers to access that characterize webcasting, while also supporting real-time interaction between speaker and audience, we have begun to combine webcasting and conferencing in ePresence [5], [6] (Figure 1).

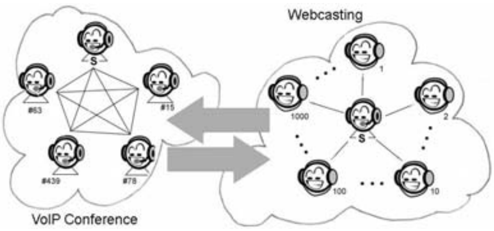

Fig. 1. Bridging webcasting for the many with conferencing for the few

To make such technology effective, at least the speaker and ideally also others in the presentation room must be aware of the remote audience. For example, the speaker may be interested in knowing answers to the following questions: How many are watching? Who is watching? Are they paying attention? Do the remote viewers seem to comprehend what is being said? Do any of them have questions?

The value of a pioneering system for distributed presentations was demonstrated in [13]. The design and evaluation of a large wall display that presents outlines, still pictures, and video streams of up to 50 remote participants, and supports audience interaction via text, is discussed in [14]. The Virtual Auditorium employed video conferencing occupying real estate on a huge wall display to link a speaker with as many as 25 remote participants [15]. Methods for providing the speaker with other forms of remote audience feedback were also investigated [16].

[1] Design of this system was grounded in part on the results of an observational study of classrooms [7].

[2] A companion paper [12] discusses enhancing interactivity via text chat both live over a webcast and asynchronously over archives of a webcast.

3 Conferencing and Awareness Extensions to ePresence

We made two assumptions: 1) The new conferencing/awareness and existing webcasting capabilities should not be interdependent; thus a passive observer can watch the webcast from a PC that does not support the new features. 2) The system should be scalable and customizable based on available bandwidth and user needs.

3. 1 Usage Scenario: Lecture Environment

In a distributed classroom, our approach allows a few individuals or remote viewing groups to engage actively in a videoconference discussion with the speaker at any time, while tens or hundreds of individuals and groups can watch the webcast and participate by submitting text questions. The remote audience may include passive observers who want to simply watch the webcast, and interested students who cannot attend the lecture in person, e.g., because they are in a remote area, ill, or traveling. We therefore have four distinct stakeholders: the speaker, the local audience, the remote audience and, to a lesser extent, a webcast moderator[3] who configures various aspects of the system (e.g., which remote audience members are permitted to speak).

Figure 2 depicts the layout of a small lecture hall we have configured for webcasting. There are two projection screens. The one in front shows the speaker's PowerPoint slides. The second displays a large "in-room awareness display" (explained below) containing visual representations of remote participants, and also their text questions, comments, and contributions to discussion. Thus the speaker and local audience members can be aware of the remote audience, which facilitates interaction.

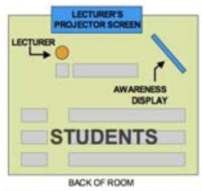

Fig. 2. Lecture hall organized for a webcast

Remote participants have the option of using slow-frame-rate video representations or still photos of themselves. Video conveys more information to the speaker and likely allows for improved engagement, but may be seen as too intrusive by some participants who wish to watch passively or even multitask without the knowledge of the presenter. Remote participants can watch the webcast and actively participate via voice interaction (using VoIP) or text chat.

[3] Our use of a moderator parallels the use of a "technology facilitator" to aid distributed teams working together with desktop conferencing and application sharing [17].

3.2 The In-Room Awareness Display

This display (Figure 3) is intended to give the speaker the ability to quickly assess the composition of the remote audience and their level of engagement. There is also a persistent chat window (implemented with BackTalk [12], [18]) that displays text questions, comments, and contributions to discussion. Text is not intended as the primary means of communication, but instead as a back channel [19] to augment voice conversation and also help the speaker gauge audience comprehension. Participants can color-tag or increase the size of text they contribute on a message-by-message basis; with this, participants could, for instance, emphasize certain contributions, or distinguish with pre-defined tags questions to the speaker, comments, and within-audience chat that the speaker need not necessarily see.

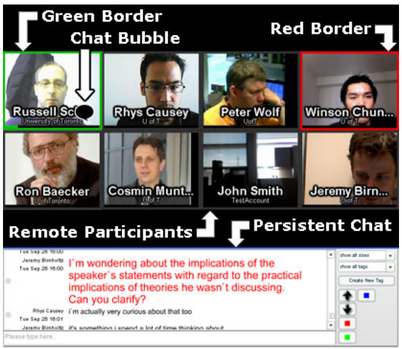

Fig. 3. The in-room awareness display

Video representations give the speaker the ability to visually monitor attentiveness and facial expressions, and see whether remote participants are engaged in the lecture. While our system does not allow for the subtleties of gaze awareness or simulated eye contact, it does provide information to the speaker that can be useful in tailoring the lecture to audience needs. While all remote students may not choose to use webcams, we conjecture that many of those who elect to let the presenter see them will likely be more engaged. Our lecturer interview participants (see section 4) suggested that these are often the ones on whom they tend to focus the most.

The awareness display also simplifies voice interaction between remote participants and the speaker. When remote participants interact by voice, they are transferred from the one-way, slightly time-delayed webcast video and audio feed to a "voice conversation," which allows for two- or multi-way voice interaction with a video feed with a nearly-imperceptible delay. This voice conversation then becomes part of the webcast that is streamed to remote participants not engaged in the voice conversation.

On the awareness display, colors and icons identify the states of remote audience members to make it easy for the speaker and local audience to determine who wants to speak and who is speaking. Remote students who wish to speak indicate this by clicking an icon to "raise their hand." When this icon is clicked, their representation on the awareness display has a red border. Once given permission by the moderator to enter the voice conversation, the participant's representation has a green border.

Because there can be multiple simultaneous participants in the voice conversation, we additionally distinguish between those who simply have permission to speak and those who are actively speaking. When the system detects that a remote participant is speaking, a chat bubble icon appears over his representation. This allows the speaker and local audience to identify which remote student(s) are speaking.

3.3 The Presenter Interface

The presenter can also use the ePresence presentation tool to easily send material to the remote audience. This tool also makes it easy for the presenter to display and send Powerpoint slides, videos, fully-navigable websites, and live desktop demonstrations.

3.4 The Remote Viewing Interface

Our primary goal for the remote viewing interface is to increase the ability of remote participants to engage and interact naturally in real-time. We allow participants to see each other through slow-scan, low-resolution video or digital images, and give them the option of interacting with text or voice. The remote viewing interface is shown in Figure 4. As with traditional webcasts, the participant receives the video feed (Fig. 4a) in sync with presentation material, such as slides or video (Fig 4e).

Questions, comments and discussion can be sent to other remote participants and the in-room display using the BackTalk persistent chat tool (Fig. 4f). Fono and Baecker [18] report that such a system, by incorporating features such as threading, annotation, and filtering, can create an environment for rich, collaborative, and ongoing conversation while maintaining the intuitiveness of more typical chat clients.

Remote participants are also shown their own chosen representation—the video image from their webcam or uploaded still image—so they see how the speaker and others will see them (Fig 4b). When remote participants are granted permission to enter the voice conference, their status bar, which is typically gray (Fig. 4b) and says "watching," turns green and says "on air." This would be the case whether they raise their hand or are "called on" by the speaker. It is also possible, in small classes or seminars where it may be desirable for all students to speak, or where certain students may have more to contribute than others, for the moderator or speaker to designate certain remote students as "VIPs" who can enter into and out of the voice conversation without having to "raise their hand."

Fig. 4. The remote viewing interface: (a) webcast video; (b) user controls; (c) active speaker; (d) remote participant list; (e) presentation content; (f) chat

For awareness of others watching remotely, there is a full list of remote participants (Fig 4d), with small visual representations. The system automatically detects and displays a list of the remote participants who are currently speaking. The first participant in the list is given a larger representation above the full user list (Fig. 4d) as the "Active Speaker" (Fig. 4c). Showing only a single speaker at a time works well because speech from remote participants typically only overlaps very briefly, and only during turn transfer. Just as a group will place focus on a person who is speaking in a traditional lecture, and a video camera will focus in on the current speaker in a traditional webcast, remote speakers are given a similar focus in our system, enabling remote participants to be aware of current speaker's identity, whether local or remote.

3.5 The Moderator Interface

As speakers have many styles for engaging students, we provided flexibility in the moderator interface. Moderators can control voice permissions, tweak automatic speech detection settings (for current speaker identification), and remove certain remote clients from being displayed on the in-room awareness display (e.g., to hide those connected locally). The moderator can also control frame rates and capture resolutions of individual remote participants' webcams. This enables the moderator to optimize the system based on available bandwidth (or even on the basis of speaker interest in certain remote participants), to divide bandwidth unequally between users (e.g., to give a larger share to students who may have more to contribute), or even to prevent certain users from using their webcam if this is necessary for any reason.

4 A Field Trial in a Multi-Campus Class

In fall 2006, we deployed this experimental system in a 4th year Computer Science evening course dealing with creating new software ventures. The course was offered to students on two campuses of the same university, located 25 kilometers apart. Approximately 60 students were enrolled at the "main" campus (where the instructor was based), and 15 at the "satellite" campus. The class met weekly for 3 hours, and consisted primarily of lectures by the instructor and weekly guest speakers.

These presentations were, for the most part, delivered in a lecture hall on the main campus, though there were occasional presentations to the entire class by a teaching assistant at the satellite campus, and two of the guest speakers delivered their presentations from California and Korea, respectively. Students had the option of attending the course at whichever campus was most convenient for them, or, when this was not possible, to sign in and "attend" from any Internet-accessible location.

At the main campus, the room was configured as shown in Figure 2. There were two video cameras at the main campus, and two staff members who operated the cameras, moderated the discussion and selected the camera shot for the webcast.

Students at the satellite campus sat in a smaller room where the webcast presentation (using the remote participant interface described above) was displayed on a large screen at the front of the room. At the satellite campus, there was a single camera, and one staff member was responsible for setting up the technology and operating the camera. To provide for VoIP interaction, there was also a wireless, handheld microphone at the satellite campus. At this campus, students indicated by raising their hand that they wished to speak, and then the microphone was brought to them.

A comprehensive evaluation of student and speaker experience is underway. Due to space constraints, we discuss here only the initial reactions of students and presenters.

4.1 Evaluation Methods

Four questionnaires were administered to all students at periodic intervals. The first of these gathered baseline demographic data and student attitudes toward and experience with technology, using an instrument described in [20]. The remainder assessed student experience in the course and with ePresence.

Interviews lasting 30-60 minutes were conducted with 7 students, with 3 students interviewed multiple times during the term. Interviews were also conducted with 5 of the guest speakers, and the teaching assistant at the satellite campus. All interviews were fully transcribed and preliminarily coded using qualitative coding software.

Field observations were conducted at both campuses. Three independent observers conducted eleven 1-3 hour observation sessions and recorded detailed field notes that were later typed and expanded. Four observation sessions were conducted at the satellite campus, seven at the main campus. One observer visited both campuses.

4.2 Student Experience

Preliminary analysis of questionnaire data indicates that there were some significant differences in technology-related attitudes between this group and those in the Caruso & Kvavik study. A higher percentage of the students in this class have higher selfreported skills ratings, and prefer a higher level of information technology within their courses compared with those surveyed in the prior study. This may be partly due to many students being computer science majors. There are also some differences in the student profiles between the two campuses. For example, the students at the "main" campus were on average two years older and had substantially more work experience in fields related to information technology than those at the "satellite" campus.

Students at both campuses were satisfied equally with the smoothness of the performance of ePresence. Evidence that students at both campuses learned from the course is that they, on average, indicated agreement or strong agreement with the statement "I learned a lot from taking this course," (M_{Main} = 5.5, $M_{Satellite}$= 5.6, on a 7-point Likert scale). Students at the "satellite" campus indicated that it was slightly more difficult to participate (M= 3.57 out of a possible 7), than those at the "main" campus (3.09), but this difference was not statistically significant.

4.3 Presenter Experience

Presenters' experience with the system was also generally positive in that it allowed them to deliver their presentations and interact with remote students, but they also indicated some concerns about being able to get adequate information about remote audience members, and about interacting naturally.

Most speakers indicated that they periodically looked at the awareness display, but that it was good mostly for very high-level information about their audience. They reported being able to see the approximate number of students and whether or not they were walking around the room. They could not see the facial expressions of individual students or assess their level of comprehension.[4]

Another issue one presenter noticed was that the camera location was at the back of the room, while the awareness display was near the front. This meant that giving students at the satellite campus the impression that one was looking at them necessarily meant not looking at their images. This presenter indicated that, once he figured that out, he made conscious effort to look at the camera periodically, and to make comments to try to engage the students:

> *And then I would try to say something that made them realize that I wasn't just randomly looking straight into the camera so that would indicate that I was intentionally looking into the camera to look at them (GS1).*

Our observational data suggest, however, that while some speakers made a special effort to engage students at the satellite campus and those logged in remotely, not all of them did so. Nonetheless, most speakers did feel that the technology did help them to reach a larger audience than they would otherwise have been able to, and that it did not detract from their interaction with and ability to engage the local audience.

[4] This was likely due in part to the low-resolution video the awareness interface currently supports, which was designed mostly for individually located remote users, with one camera per person, rather than one camera for a small group of students as was done here.

4.4 Technical Issues

As might be expected, there were numerous technical issues with this experimental system deployed temporarily (and physically set up anew each week) in an environment designed primarily for face-to-face lectures. In some cases, problems were unavoidable. Network traffic and quality of service are unpredictable, and small "hiccups" can have a large effect if they interrupt critical moments in VoIP dialogue. There were several occasions where the satellite campus was switched out of the "conference mode" and into the "webcast" to afford smoother transmission. This had the unfortunate side-effect of causing a delay when students at the satellite campus wished to participate, and both students and speakers indicated in interviews that they found this distracting and an impediment to natural interaction.

Some problems could be overcome with a more permanent facility designed for the effective capture and transmission of audio and video. Examples of these problems are the awkwardness of passing a single microphone around the room for questioners, and the difficulty of seeing adequate detail in visual images of remote participants. Solutions for these problems include arrays of fixed-location microphones and higher-resolution cameras, possibly with automated control capabilities.

5 Summary and Conclusions

We have presented the design of a novel combination of webcasting, conferencing, and awareness technologies. Our prototype implementation of the concept worked sufficiently well for both students and presenters to convince us that the concept is sound. Further positive evidence comes from a current use in a research seminar where some participants and speakers are remote from the seminar room.

We are currently re-architecting and re-implementing the system for better performance and better scalability for both wecbast and conference participants. Research goals include more automation, including remote camera control, better remote audio capture, and less need for intervention by the moderator. A future version should allow the moderator to select the location where remote students will appear on the awareness display (to capitalize on presenters' spatial memory). We are also analyzing data from this study in more detail. A future in-depth study of the pedagogical impacts of the use of ePresence involving both instructors and learners may help to identify innovations — pedagogical as well as technical — that will better support teaching and learning.

Acknowledgements. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada through a grant to the Network for Effective Collaboration Technologies through Advanced Research (NECTAR).

References

1. Patcha, A. & Scales, G. (2005). Development of an internet based distance learning program at Virginia Tech. Poster, 6th Conf. Information Technology Education, Newark, NJ.
2. Kupritz, V.W. & McDaniel, S.K. (1999). Pedagogical reflections from an instructional technology workshop, *Educational Technology & Society 2 (4)*, 119-121.
3. Heeler, P. & Hardy, C. (2005). A preliminary report on the use of video technology in online courses. *Journal of Computing Sciences in Colleges, 20 (4)*, 127-133.
4. Mizukoshi, T., Kim, & Lee, J. Y. (2000). Instructional technology in Asia: focus on Japan and Korea. *Educational Technology Research and Development 48 (1)*, 101-112.
5. Baecker, R.M., Baran, M., Birnholtz, J., Chan, C., Laszlo, J., Rankin, K., Schick, R., & Wolf, P. (2006). Enhancing Interactivity in Webcasts using VoIP, *Proc. CHI 2006*, 235-238.
6. Causey, R., Birnholtz, J.B., & Baecker, R.M. (2006). Increasing Awareness of Remote Audiences in Webcasts, *Proc. CSCW2006*, Conference Supplement, Nov. 2006, 59-60.
7. Birnholtz, J.P. (2006). Back to School: Design Principles for Improving Webcast Interactivity From Face-to-Face Classroom Observation, *Proc. DIS 06*, 311-320.
8. Baecker, R.M. (2003). A Principled Design for Scalable Internet Visual Communications with Rich Media, Interactivity, and Structured Archives. *Proc. CASCON 2003*, 83-96.
9. Baecker, R.M., Moore, G., & Zijdemans, A. (2003). Reinventing the Lecture: Webcasting Made Interactive. *Proc. HCI Int'l 2003*, Lawrence Erlbaum Associates, Volume 1, 896-900.
10. Baecker, R.M., Wolf, P. & Rankin, K. (2004). The ePresence Interactive Webcasting System: Technology Overview and Current Research Issues, *Proc. Elearn 2004*, 2532-2537.
11. Rankin, K., Baecker, R.M., & Wolf, P. (2004). ePresence; An Open Source Interactive Webcasting and Archiving System for eLearning, *Proc. Elearn 2004*, 2888-2893.
12. Baecker, R.M., Fono, D., Blume, L., Collins, C., & Couto, D. (2007). Webcasting Made Interactive: Persistent Chat for Text Dialogue During and About Learning Events. *Proc. HCI International 2007*, July 2007, to appear.
13. Isaacs, E., Morris, T., & Rodriguez, T.K. (1994). A forum for supporting interactive presentations to distributed audiences. *Proc. CSCW 1994*, 405-416.
14. Jancke, G., Grudin, J., & Gupta, A., Presenting to Local and Remote Audiences: Design and Use of the TELEP System. *Proc. CHI* 2000, 384-391.
15. Chen, M. (2001). Design of a virtual auditorium. *Proc. ACM Multimedia 2001*, 19-28.
16. Chen, M. (2003). Visualizing the pulse of a classroom, *Proc. ACM Multimedia* 2003, 555-561.
17,Mark, G., Grudin, J., & Poltrock, S.E. (1999). Meeting at the Desktop: An Empirical Study of Virtually Collocated Teams, *Proc. ECSCW 1999*, 159-178.
18. Fono, D. & Baecker, R.M. (2006). Structuring and Supporting Persistent Chat Conversations, *Proc. CSCW 2006*, Nov. 2006, 455-458.
19. McCarthy, J.F. & Boyd, D. (2005). Digital backchannels in shared physical spaces: experiences at an academic conference. *Proc. CHI 2005*, 1641-1644.
20. Caruso, J.B. & Kvavik, R.B. (2005). ECAR study of students and information technology, 2005: Convenience, connection, control, and learning. Retrieved February 13, 2007 from http://www.educause.edu/LibraryDetailPage/666?ID=ERS0506.