
Enhancing Interactivity in Webcasts with VoIP

Ron Baecker

Melanie Baran

Jeremy Birnholtz

Clarence Chan

Joe Laszlo

Kelly Rankin

Russ Schick

Peter Wolf

Knowledge Media Design Institute

University of Toronto

40 St. George St. #7228

Toronto, ON M5S 2E4 Canada

rmb@kmdi.toronto.edu

Abstract

This Interactivity demonstration presents a novel coupling of webcasting (streaming) with audioconferencing in which Voice over Internet Protocol (VoIP) communication is used in webcasts to enhance interactivity, engagement, and the sense of presence among viewers and presenters.

Keywords

Webcasting, streaming, audioconferencing, VOIP, interactivity, engagement, presence

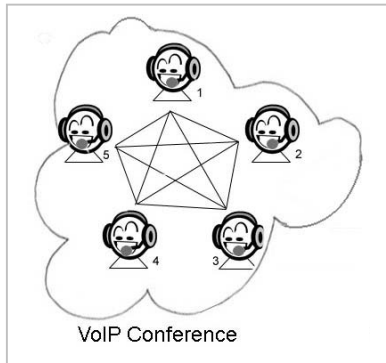
ACM Classification Keywords

H5.1 Multimedia Information Systems, H.5.2 User Interfaces, H.5.3 Group and Organization Interfaces

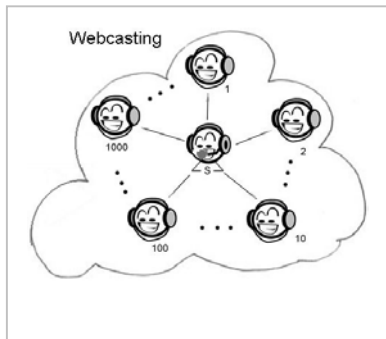
Introduction

Media spaces are distributed collaborative environments linking teams in work, learning, and play. Media spaces seek to enrich participant engagement and the sense of presence through the use of media such as audio, animation, video, and immersive 3D display.

Two major media space technologies employing video are videoconferencing and webcasting. *Videoconferencing* is most effective for connecting small numbers of individuals or groups or rooms together to speak



and view one another in almost real time (response time delays may be up to a quarter second). Videoconferencing is frequently used in group and committee meetings, or for one-one interactions. Each participant must have at least a webcam. The technical quality of the conference depends critically on bandwidth availability. User experience can degrade rapidly as the number of conference participants rises, since it becomes harder to support video and audio from all participants. Successful videoconferencing environments often employ dedicated rooms and specialized equipment such as echo cancellation devices.



Webcasting (streaming) systems allow broadcasting over the Internet to large numbers of individuals or groups or rooms. No video comes back from viewers, but they can interact with a speaker or performer via text chat. Webcasts can be viewed using almost any modern PC with all modern web browsers. Dial-in modem speeds are satisfactory, but high-speed access is preferable. To guarantee smooth delivery of audio and video, streaming engines buffer transmission between sender and receiver. Receivers thus receive the stream with what typically is a 10 to 25 second delay.

This Interactivity demonstration presents novel uses of Voice over Internet Protocol (VoIP) communication to enhance interactivity, engagement, and the sense of presence among webcast viewers and presenters. We first report on an implementation of "virtual viewing rooms" that allows students watching a webcast to converse during and after the lecture. We then discuss and will demonstrate at CHI2006 how we have enabled webcast viewers to ask questions using VOIP and to converse with a speaker and with other viewers. Both

projects extend the ePresence Interactive Media open-source cross-platform webcasting/archiving system [1].

Virtual Viewing Rooms

One of the primary disadvantages of standard webcast presentations is their lack of support for communication within the remote audience. This limitation prevents remote attendees from interacting before, during, or after a presentation as is possible when attending a presentation in person. Such interactions are often an important part of the experience of attending a lecture, allowing members of the audience to discuss the content of a presentation and to form and maintain social relationships. Without these opportunities to communicate, webcast lectures can feel disengaging and impersonal. One approach to this problem is to combine webcasting with peer-to-peer communication tools such as text chat and VoIP audio conferencing.

To investigate the impacts of peer-to-peer communication on remote users, we implemented such a system and conducted a user study on an undergraduate computer science class [6]. 27 students were placed into 9 study groups of three students each. Groups were exposed to three experimental conditions: a standard webcast lecture, where students viewed a pre-recorded lecture and could not communicate with their study groups; a condition in which students could communicate with their study group using text chat; and a condition in which students could speak to each other using a VoIP audio channel as well as text chat.

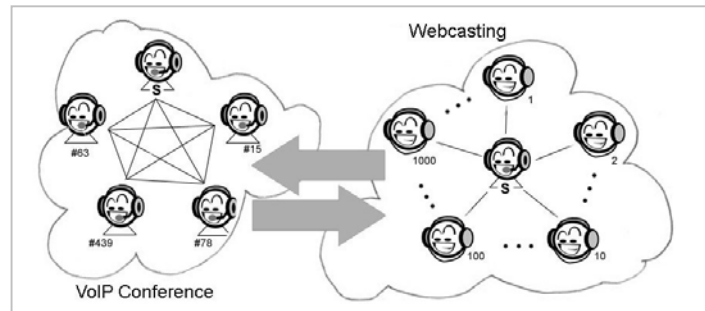
The results of the study showed that students clearly preferred having the ability to communicate with their group. Students reported feeling most isolated and least engaged when no communication was allowed

with their group, and felt most engaged when they could use the audio channel. There was also more social conversation with audio than text.

Voice Communication with the Speaker

Based on what we had learned from that project, we created a new voice interactivity module to enable dialogue between the speaker and remote audience members, who could then become more active participants.

An online 'moderator' for the webcast is physically connected to a Flash Communication Server that links them to the moderator's machine. The viewer may then question and talk to the presenter; other remote viewers can also join the discussion. The audio is typically fed to the local audience with loudspeakers and also delivered back into the webcast so that other remote viewers can hear it.



We used this version of the ePresence Interactive Media system (see <http://epresence.tv>) in a Canada-wide distributed course on CSCW conducted from September 2005 through December 2005. It was also used in webcasts of readings by Canadian writers to schools located throughout the country.

User impact

In addition to these experiences, there is significant evidence suggesting that the webcast user experience could be improved by the addition of live interaction between audience and speaker via audio.

Improved understanding

Webcasts are frequently used in lecture settings where the presenter is teaching difficult or abstract concepts to a virtual classroom full of students. Developing a shared understanding of the concepts under discussion, or "common ground", can be difficult in this setting because of the difficulty of the concepts and because participants may come from different cultural or disciplinary backgrounds [3]. While the combination of text chat with traditional webcast audio systems that support only "one to many" or communication are useful, they do not facilitate natural conversation and negotiation in ways that are necessary for effective negotiation of common ground.

Specifically, Clark and Brennan point to the importance of co-temporality and audibility between speakers attempting to negotiate common ground. Olson and Olson [5] also point to the importance of "rapid feedback" in communication contexts where common ground is poor and misunderstandings are likely. With rapid feedback such misunderstandings can be quickly detected and corrected. Our VoIP tool improves upon existing webcasting tools in two specific ways.

In the first place, it allows for improved interaction between the speaker and a specific audience member, in that both can speak simultaneously to the entire group and it is possible to more rapidly come to a shared un-

derstanding of, for example, what exactly an audience member means in asking a question.

Second, by facilitating conversation involving *multiple* members of the audience simultaneously, it better supports the entire group's negotiation of a shared understanding of the concepts being presented. In the case where, for example, a presenter does not understand a question being asked by an audience member, allowing for the voice participation of multiple audience participants allows the audience to help correct the presenter's interpretation of the question, or to voice their own lack of understanding.

Community Building

There has also been recent interest in using webcasting and elearning technologies to foster "learning communities." The technologies presented here foster the development of social communities, as measured preliminarily with trust – arguably a critical component of any community [4]. This is accomplished first by allowing for informal interaction both before and after the presentation, which has been shown to aid in the development of trust in distributed groups [7].

Audibility itself has also been shown to contribute to the development of trust [2]. Thus, the capacity to hear the voices of their classmates and contribute to an ongoing dialogue arguably increases students' sense of belonging.

User Testing

We are currently putting together plans to conduct field tests of this new ePresence functionality. We plan to measure the effect of the VoIP features on users' sense of community, their ability to understand each other,

and on how trust develops between participants in a webcast lecture. We also plan to test these features in conjunction with multiple awareness tools, to better understand how to facilitate effective dialogue and discussion between participants. We plan to present preliminary data from these tests at the conference.

Acknowledgements

Thanks for support to the Natural Sciences and Engineering Research Council of Canada and to the Bell University Laboratories of the University of Toronto.

References

- [1] Baecker, R.M. (2003). A Principled Design for Scalable Internet Visual Communications with Rich Media, Interactivity, and Structured Archives. *Proceedings of CASCON 2003*, 83-96.
- [2] Bos, N. D., Olson, J. S., Gergle, D., Olson, G. M., & Wright, Z. (2002). Effects of four computer-mediated channels on trust development. *Proc. CHI 2002*, 135-140.
- [3] Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, R. M. Levine & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127-149). American Psychological Assoc.
- [4] Jarvenpaa, S.L., & Leidner, D.E. (1999) Communication and trust in global teams. *Organization Science*, **10**, 791-815.
- [5] Olson, G. M., & Olson, J. S. (2001). Distance matters. *Human-Computer Interaction 15(2-3)*, 139-179.
- [6] Schick, R., Baecker, R.M., and Scheffel-Dunand, D. (2005). Bimodal Text and Speech Conversation During On-line Lectures, *Proceedings of ED-MEDIA 2005*, June 27-July 2, 2005, Montreal, P.Q.
- [7] Zheng, J., Veinott, E., Bos, N., Olson, J. S., & Olson, G. M. (2002). Trust without touch: Jumpstarting long-distance trust with initial social activities, *Proc. CHI 2002*, 141-146.