2006 January 07

Towards a Video Collaboratory

Ronald M. Baecker, David Fono, and Peter Wolf

University of Toronto

Introduction

Wulf (1989) defined a collaboratory as a "'center without walls,' in which the nation's researchers can perform their research without regard to geographical location -- interacting with colleagues, accessing instrumentation, sharing data and computational resources, [and] accessing information in digital libraries." We shall present the ePresence system, an open source interactive real-time webcasting and multimedia archiving solution. We sketch how it could be extended and applied to enable research and collaborative use of video-as-data by a worldwide community of educational researchers, teachers, and learners. Particular emphasis will be placed on how the system supports representation, reflection, interaction and collaboration.

Facilitating the Use of Video in the Learning Sciences

Goldman et al. (2006, this volume) describe the learning sciences as “a distinctive branch of the multidisciplinary cognitive sciences, with distinctive emphases on the problems of education and learning,” and assert that an understanding of human learning requires insights from multiple disciplines including cognition, developmental psychology, educational psychology, linguistics, anthropology, education, and computer science. A central goal of the learning sciences is to produce enhanced descriptions and understandings of education and learning, using various kinds of data and various methods of deriving meaning from the data.

Learning science descriptions are often expressed in text, e.g., the teacher was "eloquent", the student was "puzzled", or the class was "disruptive". Descriptions may also be represented mathematically, e.g., the student made "two errors." Yet such abstractions are reductionist abstractions of what one or more observers have concluded is the meaning of the data; they do not have the richness of video records of the actual behaviour. As this volume demonstrates, there is therefore much interest and activity in the use of video as data.

This article begins by reviewing past and current research on tools to facilitate the use of video-as-data, with particular interest in the concept of the "collaboratory" and in the support of video data in collaboratories. We then introduce ePresence Interactive Media — an open source interactive real-time webcasting and multimedia archiving solution, present some sample uses of ePresence, and discuss its architecture and implementation. Although not originally designed for use as a video collaboratory, the system is modular and malleable enough to allow modest extensions that enable such use. We present these planned extensions in terms of support for representation, reflection, interaction, and collaboration.

Related Research

Since the publication of Wolf's paper, many collaboratory initiatives have been undertaken, primarily by organizations studying the physical sciences (e.g., Berman, 2000; Caspar et al., 1998; Van Buren et al., 1995). Much of the existing research on collaboratories describes lessons learned from these ventures, focusing in particular on the sociological conditions necessary to facilitate successful collaboration (e.g., Olson, Finholt, & Teasley, 2000; Schunn, Crowley, & Okada, 2002).

The National Research Council (1993) identified key technological needs that must be satisfied in order to achieve a fully functional collaboratory:

- data sharing (including electronic libraries, accessible archives, and a comprehensive retrieval system)
- software sharing (including interoperability of local software with remote data, and network-accessible storage of results)
- communication (including voice, video, text, data, and images in both synchronous and asynchronous modes), and
- the ability to control remote instruments.

In particular, the system must integrate these functionalities so that the borders between them are transparent to the collaboratory participant, who works with data, tools, and colleagues in unison. These requirements were conceived with the physical sciences in mind, but most are typically applicable to learning science collaboratories, although it would be rare to need to control remote instruments. However, most projects have been developed without any guiding plan, using off-the-shelf technologies not designed to work well together (Finholt, 2002).

The prospect of a video collaboratory faces additional barriers to an effective implementation. Video can be a primary data source for observational inquiry, permitting qualitative analyses of behaviors and processes (Smith and Blankinship, 1999). However, video-as-data is a novel application requiring novel tools (Nardi et al., 1996).

Video can also be used to share real-time depictions of shared work objects, and thus bring complex objects at one physical location into a virtual shared workspace to coordinate distributed teams (Whittaker, 1995). For example, a study by Gaver et al. (1993) emphasizes that in many cases real-time views of an object under study in a shared workspace is preferable to conversational views among fellow collaborators. Use of video in this manner can help to establish *common ground* among collaborators, a shared physical context that adds meaning to indexical utterances (Clark and Brennan, 1991). Thus, the development of WYSIWIS (What You See Is What I See) interfaces is one of strong imperatives in collaboratory design (Finholt, 2002).

Pea (2005) summarizes the challenges facing a video collaboratory as follows: video data and analyses must be universally accessible, collaborators must have access to video analysis tools that support discipline-specific analytic practices and are interoperable, and analyses and commentary must be available for public participation and collaboration.

Several ongoing and past projects have aimed to make video data universally accessible by developing large, public corpora of digital multimedia recordings. The Open Video Project (www.open-video.org; Marchionini & Geisler, 2002) is in the process of accumulating a shared digital video repository to serve as a test bed for research into information retrieval in multimedia libraries. The Informedia Digital Video Library (Hauptmann, 2005) and the CAETI Internet Multimedia Library (Wolf et al., 1997) are similar test bed repositories, together having archived over 1000 hours of educational video for K-12 students as part of ongoing research into constructing digital video libraries; the Informedia collection since been added to the Open Video Project. The Talkbank site (www.talkbank.org) maintains a large body of video and audio data, as well as transcription, coding, and annotation tools, designed to help researchers studying human and animal communication.

Many projects also aim to provide comprehensive retrieval functionality for existing libraries. For example, the Informedia Project uses combined speech, language and image understanding to provide intelligent search and selective retrieval within its database of videos (Wactlar et al., 1999; Christel et al., 2004). Users can search for "video paragraphs" based on text extracted from the soundtrack or captioning, and rapidly browse the results using a "video skimming" technique. Other projects explore the automatic capture and archival of live content (Brotherton and Abowd, 2004; Chiu et al., 2000; Hurst et al., 2001; Kientz et al., 2005; Moran et al., 1997; Mukhopadhyay and Smith, 1999).

Tools for video analysis and interoperability of these tools with the corpus of networked data form complementary challenges. Pea and Hay (2003) conducted a workshop with video researchers in the learning sciences to identify the functions that analytical tools should support: acquisition, chunking, transcription, way-finding, organization/asset management, commentary, coding/annotation, reflection, sharing/publication, and presentation.

VideoPapers (Beardsley, Cogan-Drew, and Olivero, 2006, in this volume) are online multimedia publications designed for educational researchers and practitioners. Using the VideoPaper Builder, users can synchronize videos with textual analysis and images, and publish the presentation to the web.

DIVER (Digital Interactive Video Exploration and Reflection) allows researchers in the behavioral sciences to easily edit and annotate collected footage, thus creating lightweight "dives" that illustrate some specific point or piece of evidence (Pea et al., 2004). Dives can subsequently be exported to a website, so that other researchers have a chance to observe and comment on them.

VideoTraces (Stevens, 2006, in this volume) offer a unique form of "show and tell" multimedia representations for conveying embodied knowledge. Users augment video with an interpretive layer of audio and visual gestures to create traces, which can then be exchanged with others or used for future reflection.

The importance of point-of-view in interpreting and constructing thick descriptions of video data has been particularly emphasized in the video ethnographic work of Ricki Goldman (Goldman-Segall, 1998). Goldman has also for the past two decades been designing a variety of interesting systems for video annotation and analysis (Goldman, 2006, in this volume; see also Harrison and Baecker, 1992).

Numerous technologies have also been developed to facilitate synchronous and asynchronous collaboration at a distance. Fishman's KNOW (Knowledge Networks on the Web) system (Fishman, 2006, in this volume) envisions a collaborative teacher professional development environment that combines hyperlinked curriculum documents, student work with teacher feedback, tools for personal weblogging and discussion, and several types of instructional video.

Several systems developed by Microsoft Research allow real-time interaction amongst individuals who are collaboratively watching a video webcast (Cadiz, et al., 2000; Jancke, et al., 2000; Rui, Gupta, & Grudin, 2003; White, et al., 2000). Finally, Grudin & Bargeron (2005) present technology to allow asynchronous collaboration around video archives, using a shared annotation system integrated with e-mail.

The system we are about to describe was based around the dual goals of enabling the real-time transmission through the Internet of synchronized multimedia including video, and of enabling the efficient archiving of such presentations to allow flexible browsing, navigation, and searching of the material. We shall introduce the system, then sketch how it could be extended to enable the use of video-as-data in a video collaboratory for the learning sciences.

ePresence Interactive Media

ePresence Interactive Media (Baecker, 2003; Baecker Moore and Zijdemans, 2003; Baecker Wolf and Rankin, 2004; Rankin Baecker and Wolf, 2004; Baecker et al., 2006) is a web-based streaming (webcasting) and collaboration tool for the large-scale broadcast of events over the Internet — from university lectures to demonstrations by master teachers to public health briefings to annual meetings to rock concerts. Events are streamed live and can later be easily deployed as browsable, searchable archives accessed through a customizable web portal. Webcasting itself is non-interactive, which is overcome by combining it with interactive features. For example, ePresence currently employs text chat as a mechanism for allowing interaction among remote participants, and between these individuals and the speaker via a moderator, and has also recently added VoIP support for voice questions and discussion.

For both live and archived events, ePresence provides a rich and engaging multimedia experience for viewers connecting over the Internet using desktop and mobile clients. During a live event, end-users have access to an audio-video feed, navigable slide images, and a text chat system. Live events can be quickly and easily archived, and made available to users via the portal. Archives are full-text searchable, and provide an interactive timeline and two-level table of contents for easy browsing and navigation. The concept of hierarchically structured video is based in part on work described in Baecker, et al. (1996) and Baecker and Smith (2003)

More specifically, ePresence currently includes support for:

- video, audio, slide, and live desktop demos;

- slide review;
- moderated chat, private messages, and the submission of questions via text and voice; and
- the automated creation of structured, navigable, searchable event archives.

ePresence also allows configurable live and archive interfaces through tailorable "skins", which allow site-specific control over the layout and typography of both interfaces, and the inclusion of corporate logos for purposes of "branding." The server-side software runs under Linux and Windows; media capturing and streaming engines run under Windows; client viewers exist for the IBM PC, the Macintosh, and Linux platforms. Media may be transmitted using Windows Media, Real Media, and MPEG-4. Webcasts may be received with bandwidths as low as a 56Kbits/second. The software is implemented with .NET and Mono technologies, is highly modular, and has been released open source and community source (Baecker 2005, see also http://epresence.tv).

User Interface

One interface to access live webcasts is illustrated by the screen snapshot in Figure 1. The video window and its controls are in the upper left; the slide window and its controls are in the upper right; the chat system is at the bottom. Slide controls allow a remote viewer to review any slide already presented by the speaker. The chat system supports public chat, private messages, and questions to the speaker. Web links can also be sent by the speaker and synchronized with the video. A "live demo" feature enables transmission of live 600X800 screen captured streams of live demos from the presenter's computer. There is an integrated registration and systems check procedure so that potential viewers can ensure technology compatibility in advance.

PLACE FIGURE 1 ABOUT HERE

The archives interface allows retrospective navigation and browsing through a webcast using a two-level outline of the logical structure of the talk and its slides and live demo sessions (Figure 2, right side). Slide titles are picked up automatically from PowerPoint in case it is used; the outline is input by the moderator during the talk and if need be updated afterwards using the ePresence Producer (see below). Archive viewers can also navigate by a timeline (Figure 2, bottom). We also allow searching based on key words in the slides when PowerPoint is used. (Both dependencies on PowerPoint are to be removed in the February 2006 release of ePresence Version 3.1.) Chapter titles appear darker in the table of contents, and as the upper tick marks on the timeline. Slides appear lighter in the table of contents, and as the lower tick marks on the timeline.

Example Uses of ePresence

An interesting case study of ePresence (Zijdemans, et al., 2005) has been its use by the Millennium Dialogue on Early Child Development (MDECD) project, part of the Atkinson Centre for Society and Child Development's steps towards establishing a learning community for child development based on a theoretical model for developing a learning society network (Keating & Hertzman, 1999; Matthews & Zijdemans, 2001; http://www.webforum2001.net/; http://www.acscd.ca).

PLACE FIGURE 2 ABOUT HERE

MDECD brought together eight experts from different areas in child development for a two-day conference in November, 2001. The meeting was attended by roughly 200 local participants and was webcast using ePresence to 20 remote North American groups. Over 600 public and private chat messages among the remote groups were exchanged. Table 1 shows how the composition of the chat messages changed over the two days. Note the increase in the percentage of messages related to the content of the sessions, from an average of 4% on day 1 to 13% on day 2, and in the percentage of social messages, from 8% on day 1 to 26% on day 2.[1]

PLACE TABLE 1 ABOUT HERE

Since the conference, the ePresence multimedia archive of the scientist presentations has served as a knowledge base and nurtured the learning community through ongoing activities:

- ongoing curriculum development and incorporation of the knowledge base into courses for graduate students and professionals
- creation of *Conversations on Society & Child Development* [see http://www.cscd.ca], an interactive CD/Web ePublication that provides an environment for accessing the knowledge and supporting exchange among researchers and those who want to apply the findings; and
- plans to translate the knowledge for use by parents, educators, and policy makers.

[1] White, et al. (2000) similarly report that text exchanges went from 27%:62%:11% content:technology:social messages to 60%:14%:26% over the last 3 sessions of an eLearning course.

A Second Example

On July 16, 2003, North Network with the assistance of Videotelephony Inc. used their videoconferencing network and the ePresence system to webcast a talk entitled “West Nile Virus: first Canadian Experiences.” Simultaneous talks were given by two regional infectious disease experts, and the North videoconference network was used to connect remote communities with these experts. The videoconferencing feed was bridged into ePresence, and used as the basis for a simultaneous webcast. The talks were recorded, digitized, and mounted on NORTH Network’s server for archive access (www.northnetwork.com). ePresence technology was also used to create a knowledge product (on Web and CD): *Just-in-Time’ Clinical Education During SARS and West Nile*, which was awarded “best innovation in use of technologies in health education” at the Canadian Society for Telehealth Conference.

A Third Example

May 9-11 2004 saw the Knowledge Media Design Institute hosting a major international conference entitled “Open Source and Free Software: Concepts, Controversies, and Solutions.” There were roughly 20 hours of lectures, panel discussions, and question and answer dialogues between the 250 local audience members, 25 remote audience members, and 30 speakers. Graduate student editors reviewed the proceedings in detail and added hundreds of chapter titles that together with the slide titles provide a rich table of contents into the multimedia proceedings (see Figure 2 and also http://www.epresence.tv/website_archived.aspx?dir=7?).

System Architecture and Implementation

The system is implemented using .NET and Mono technologies. The server software runs under Windows or Linux. Webcasts can be viewed on client personal computers running the Linux, Windows 98/2000/2003/XP, and Mac 9.x or OS/X operating systems, and the Internet Explorer, Netscape Navigator, Mozilla Firefox, Opera and Safari 1.2+ browsers, and using either Real Media or Windows Media live streaming. Archives may be produced in Real Media, Windows Media, and MPEG-4 formats.

The architecture of our highly modular system may be portrayed as in Figures 3 and 4. For further details, see Baecker et al. (2004) and Rankin et al. (2004).

PLACE FIGURE 3 ABOUT HERE

Interactive Live Webcasting (Figure 3)

An ePresence live webcast is created by a speaker, an operator, and a moderator. These can be different individuals or the same person depending upon the scale of the event. The ePresence Mobile Station (4) includes several live media encoding and capturing software applications (e.g., Windows or Real Media) controlled by the operator (3) or speaker (7) via a single unified remote control interface. The remote control interface has been developed for different internet-connected devices (Laptops, tablet PCs, and PDAs). The operator can perform the following operations either locally or remotely: initiate live broadcast, start or stop archiving session, control slides transmission, submit URLs, and initiate multiple live software demo sessions. The speaker may give a talk to a local audience or remotely via a telephone, VoIP, or videoconferencing (1). This allows us to webcast a meeting that is being held via videoconference. Web-based slide controlling and projecting (7) software allows having multiple distributed audiences listening and following the slide presentation in real time. The moderator interface (10) supports a local moderator who is watching the webcast, sending public announcements to a web audience, and submitting notes (chapter titles) to the archiving application. The moderator works as a communication "bridge" between the speaker and web audience transferring questions and comments on behalf of remote participants.

An ePresence webcast is typically viewed by both a local audience (2) and a live web audience (6). The web audience receives video and audio (13) of the speaker(s) from the streaming servers (5), a synchronized slide presentation stream (18) or a screen capture stream (14) from the presenter's computer, and web URLs (15). Remote viewers can also submit questions to the speaker (directly via text or voice or indirectly via the moderator), have public or private text based dialogs (16), and review the slides that have been already presented. The live interface (9) has been developed as a set of templates ("UI skins") that support different layouts, media formats, video resolutions and other features. The operator can choose the most suitable template depending on the content of the talk. Adopters of the ePresence system can easily develop their own skins using XML, HTML, and a choice of several scripting technologies.

PLACE FIGURE 4 ABOUT HERE

Archiving and Publishing a Webcast (Figure 4)

The webcast data (4) such as video (1), slides (3) and event streams (2) is automatically captured during the live webcast. The events stream data includes time stamp information of slides and chapters submitted during the live webcast. Event streams can be updated (5) after the webcast using the ePresence Producer application (12). The operator can add additional keywords to enhance search, update slide synchronization data, edit chapter and slide titles, and replay the event with all synchronized materials before publishing the archive. The ePresence Producer software also allows encoding the captured video in different popular streaming formats (6), automatic uploading to a streaming server (7), automatic creation and publishing of web archives (8), and production of multimedia CDs (9). The software provides a selection of archive templates. The published archive becomes automatically available on the ePresence website (10). It includes video player, slide frame, interactive timeline component, search tool, interactive table of contents, and threaded discussion board. Every archive exposes its keywords through the XML web services. This makes it easy to integrate the archives into different document repositories, "learning object" banks, and other searchable data storage systems.

ePresence as a Collaboratory Enabling the Use of Video in the Learning Sciences

ePresence was originally developed to enable the worldwide broadcast of presentations, interactive access to these broadcasts in real-time, and flexible retrospective access to structured archives of the presentations. Yet there is nothing in the technology that restricts the video channel to "talking heads" presenting lectures, or to a presentation in which a small video image is portrayed as an adjunct to a large slide image.

In Figure 5[2], we see a screen shot from an experimental video collaboratory version of ePresence developed by Russ Shick (2005) as part of his M.Sc. work. The application is collaborative video viewing (Cadiz et al., 2000) over the Internet of a structured archived video of a classic Canada-USA hockey match. Video structure consists of the periods of the game and within these periods interesting events such as goals, penalties, and "near goals." Multiple

[2] This image is included for illustration purposes only, and implies nothing about current or planned product or service offerings in the Bell Sympatico portal.

viewers are able to converse in two modalities (Schick et al., 2005) using both ePresence text chat and an experimental spatial audio voice-over-Internet system known as Vocal Village (Kilgore et al., 2003). Control over video playback is distributed among all viewers of the video.

More generally, this system could contain any video stream of relevance to educators and to the learning sciences, such as a teacher demonstrating a difficult concept in the analysis of an English text, a group of students discussing an ethical quandary, an animation of a law of physics, or the movement of organisms seen under a microscope.

PLACE FIGURE 5 ABOUT HERE

This implementation is encouraging, but we need to go further. If ePresence is to serve as a collaboratory for the learning sciences, then its capabilities for representation, reflection, interaction, and collaboration need to be enhanced. We shall now address each of these in turn.

Representation

ePresence represents video as a structured document with a 2-level hierarchy. The upper level encodes logical sections of a video, which we typically term “chapters”. The lower-level is usually used for representing associated slides or live demonstrations. But these are totally general indices into a video, so can be used in whatever way is desired by the producer of the multimedia event. For example, in the multimedia proceedings of the open source conference discussed above, the upper-level was used to provide detailed topic references to talks or panels at a rate of once every minute or two, and, in the hockey game collaboratory, the lower-level was used to encode interesting events.

For simplicity, we resisted the temptation to provide arbitrarily-linked hypervideo. Yet links from anywhere in one video to anywhere in any other ePresence multimedia archive can be introduced. We can currently reference slides, screen capture videos, and HTML pages. Other media formats (like Flash movies) will be added soon. We can reference the entire repository as a single object, which is used in our “Repository Search” feature. We can also reference arbitrary entry points or arbitrary sections of any video in order to enable chunks of material to function as “learning objects”.[3] We plan to extend this feature to allow referencing synchronized presentation media from other applications, such as email and online learning environments.

The system does currently impose the limitation that there be only one video active at a single time, although lifting this restriction through use of the open source code would be possible. This would be required to realize the notion of multiple video points of view of an event as proposed by Goldman (2006, in this volume).

Reflection

ePresence currently provides the ability for viewers to chat over a live event both publicly and privately and to send questions to the speaker. Reflection and note-taking could be realized by sending private messages to oneself, which could function as bookmarks or more generally as notes to oneself. The system stores all messages, although they have not as yet been made public in the archive. Doing this would be trivial, although we would need some way of distinguishing between those messages that are intended to be persistent, and those that are not, and some lightweight method to control access to certain messages. We need to allow notes to include hyperlinks as well as raw text.

[3] "A learning object... [is] any digital resource that can be reused to support learning." (Wiley, 2002).

More interesting than textual notes is the ability to annotate or scribble on the videos. This can easily be added to ePresence, leveraging insights from work such as Brotherton and Abowd (2004).

Interaction

We are currently investigating ways to extend interaction amongst users beyond basic live chat at the time of the event.[4] ePresence archives already have a rudimentary threaded discussion capability, but the interface lacks elegance and the feature has rarely been used. The discussion boards are fully separate and inaccessible from the live chat, which may explain the scarcity of usage. We intend to replace this functionality with a more robust discussion interface that tightly integrates synchronous and asynchronous communication, both during and after the event.

The extended interface that we are currently building consists of a single discussion view that supports chat as well as threading and basic annotation. The threading model is based in part on the Threaded Chat system (Smith, Cadiz, & Burkhalter, 2000). Users may contribute to discussions either by sending a chat message, or by posting a message in response to a previous message. When a response is posted, it also appears as a new chat message with a link to the corresponding thread, so that discussants only need to pay attention to a single stream of new content. The combination of “chat-like” and “message board-like” functionality within a single interface means that participants can adapt the style of their conversations to the circumstances at hand, including the number of discussants and the complexity of the topics involved, as well the circumstances of any concurrent webcast.

[4] This is the M.Sc. research of David Fono.

Messages may also be annotated using user-defined tags, as well as priority indicators. Tagging has been found to be a useful technique for collaboratively organizing a variety of media, and we expect that it will have a similar effect in the chat context. Using the interface's customizable visualization of chat history, users can navigate content marked with specific priorities and specific tags. Thus, participants can engage in discussions over a variety of topics within a single archive, even as that archive grows indefinitely.

The interface's appearance and behaviour will be the same for both live and archived events. This will serve to narrow the distinction between the two modes of operation, and thus encourage sustained interaction. Discussion amongst users should start with the event, move into the archives, and continue well beyond. We are also exploring means for participants to export discussions to third-party servers and interfaces, in order to further facilitate varied forms of collaboration around video archives. We expect that the final version of our chat interface will serve not only as an interface for communication amongst participants, but also as a portal for tracking the various discussions about the video that develop elsewhere.

Collaboration

Collaboration in ePresence is enabled by communication using the various chat and discussion capabilities. But one can only interact with individuals if you are aware of their existence as viewers of an event, or as possible viewers of an archive. ePresence currently provides a rudimentary display of all individuals watching a live event (Figure 1), but no display of all individuals who have watched an archive or are currently watching it.

In order to enable any kind of collaboration, we need to have awareness of who is potentially available to chat, to brainstorm, or to do focused work of some kind (Kraut et al., 1998). There is much current research on mechanisms for group awareness (mechanisms to enrich group awareness (see, for example, Rounding and Greenberg, 2000; Elliot and Greenberg, 2004; Gutwin and Greenberg, 2004). Work is just beginning on how to integrate such capabilities into ePresence (Baecker, et al., 2006).

Summary and Conclusions

We have presented a scalable, modular, extensible architecture for interactive multimedia webcasting and for providing access to structured archives of these webcasts. Various features to enable communication and collaboration over live and archived events have also been proposed and discussed. These features can be implemented and need to be implemented if the environment is to function well as a video collaboratory for educational researchers.

To facilitate this happening, as well as to allow ePresence to be molded by its adopters in many different directions, we have decided to release our software open source (DiBona, Ockman, & Stone, 1999; Benkler, 2002; Weber, 2004; Baecker, 2005). Applications of ePresence to date, mostly in the eLearning space, have all been carried out with differing goals, modes of usage, and measures of success. An open source implementation that puts maximum control in the hands of adopters should also enable a rich set of new ePresence applications in the learning sciences.

Acknowledgements

Guidance of early ePresence definition and scoping was shared with Dr. Gale Moore, Director of KMDI. Kelly Rankin has very capably anchored the business, marketing, and production side of ePresence, working side-by-side with system architect and developer Peter Wolf. Maciek

Kozlowski of Videotelephony Inc. has provided many technical insights that have informed these developments. We are thankful for research support to the Natural Sciences and Engineering Research Council of Canada (NSERC) especially through the NECTAR (Network for Effective Collaboration Technologies through Advanced Research) NSERC Research Network and to the Bell University Laboratories.

References

Baecker, R.M. (2003). A Principled Design for Scalable Internet Visual Communications with Rich Media, Interactivity, and Structured Archives. *Proc. CASCON 2003,* 83-96.

Baecker, R.M. (2005). Open Source Strategies for Educational Multimedia, *Proceedings of ED-MEDIA 2005*, June 27-July 2, 2005, Montreal, P.Q.

Baecker, R.M., Baran, M. Birnholtz, J., Chan, C., Laszlo, J., Rankin, K., Schick, R., and Wolf, P. (2006). Enhancing Interactivity in Webcasts Using VoIP, *Proc. CHI 2006*, to appear.

Baecker, R.M., Moore, G., & Zijdemans, A. (2003). Reinventing the Lecture: Webcasting made Interactive. *Proc. HCI International 2003*, Lawrence Erlbaum Associates, Volume 1, 896-900.

Baecker, R.M., Rosenthal, A., Friedlander, N., Smith, E., & Cohen, A. (1996). A Multimedia System for Authoring Motion Pictures, *Proc. ACM Multimedia 1996,* 31-42.

Baecker, R.M. and Smith, E. (2003). Modularity and Hierarchical Structure in the Digital Video Lifecycle. *Proc. Graphics Interface 2003*, Halifax NS, 217-224.

Baecker, R.M., Wolf, P. and Rankin, K. (2004). The ePresence Interactive Webcasting System: Technology Overview and Current Research Issues, *Proceedings E-Learn* 2004.

Beardsley, L., Cogan-Drew, D., and Olivero, F. (2006). VideoPaper: Bridging Research and Practice For Pre-Service and Experienced Teachers, in Goldman et al. (2006), this volume.

Benkler, Yochai (2002). Coase's Penguin, or, Linux and the Nature of the Firm, *112 Yale Law Journal* 369, 2002, 1-79.

Berman, H.M. (2000). Research Collaboratory for Structural Bioinformatics. *Protein Data Bank Annual Report, 1999-2000*. Rutgers University.

Brotherton, J.A. and Abowd, Gregory D. (2004). Lessons Learned from eClass: Assessing Automated Capture and Access in the Classroom. *ACM Transactions on Computer-Human Interaction*, *11* (2). 121 - 155.

Cadiz, JJ., Balachandran, A., Sanocki, E., Gupta, A., Grudin, J., Jancke, G. (2000). Distance Learning Through Distributed Collaborative Video Viewing. *Proc. CSCW 2000,* 135-144.

Caspar, T.A., Meyer, W.M., Moiler, J.M., Henline, P., Keith, K., McHarg, B., Davis, S., & Greenwood, D. (1998). Collaboratory Operations in Magnetic Fusion Research. *ACM Interactions* 5(3), 56-64.

Chiu, P., Kapuskar, A., Reitmeier, S., Wilcox, L. (2000). Room with a Rear View: Meeting Capture in a Multimedia Conference Room. *IEEE Multimedia October-December,* 48-54.

Christel, M., Papernick, N., Huang, C., & Moraveji, N. (2004). Exploiting Multiple Modalities for Interactive Video Retrieval. *Proc. International Conference on Acoustics, Speech and Signal Processing*, 1032-1035.

Clark, H.H., & Brennan, S.E. (1991). Grounding in Communication. In L.B. Resnick, J.M. Levine, & S.D. Teasley (Eds.). *Perspectives on Socially-Shared Cognition*. American Psychological Association.

DiBona, Chris, Ockman, Sam, and Mark Stone, Mark (Eds.) (1999). *Open Sources: Voices from the Open Source Revolution*, O"Reilly and Associates.

Elliot, K. and Greenberg, S. (2004). Building Flexible Displays for Awareness and Interaction. *Video Proceedings and Proceedings Supplement of the UBICOMP 2004 Conference*, (Sept 7-10, Nottingham, England) 6 minute video and 2-page paper.

Finholt, T.A. (2002). Collaboratories. In B. Cronin (Ed.). *Annual Review of Information Science and Technology*, Volume 36. American Society for Information Science and Technology.

Fishman, B.J. (2006). Linking On-line Video and Curriculum to Leverage Community Knowledge, in Goldman et al. (2006), this volume.

Gaver, W., Sellen, A., Heath, C., and Luff, P. (1993). One is Not Enough: Multiple Views in a Media Space. *Proc. CHI 1993*, 335-341.

Goldman, R. (2006). Orion, An Online Collaborative Digital Video Data Analysis Tool: Changing Our Perspectives as an Interpretive Community, in Goldman et al. (2006), this volume.

Goldman, R., Pea, R., Barron, B., and Derry, S. (Eds.) (2006). *Video Research in the Learning Sciences*, Lawrence Erlbaum, this volume.

Goldman-Segall, R. (1998). *Points of Viewing Chidren's Thinking: A Digital Ethnographer's Journey.* Lawrence Erlbaum.

Grudin, J., & Bargeron, D. (2005). Multimedia Annotation: An Unsuccessful Tool Becomes a Successful Framework. In K. Okada, T. Hoshi and T. Inoue (Eds.). *Communication and Collaboration Support Systems*. Ohmsha.

Gutwin, C., and Greenberg, S. (2004). The Importance of Awareness for Team Cognition in Distributed Collaboration. In E. Salas and S. M. Fiore (Editors), *Team Cognition: Understanding the Factors that Drive Process and Performance*, 177-201, Washington:APA Press.

Harrison, B.L. and Baecker, R.M., Designing Video Annotation and Analysis Systems, *Proceedings Graphics Interface '92*, Vancouver, B.C., 11-15 May 1992, 157-166.

Hauptmann, A. (2005). Lessons for the Future from a Decade of Informedia Video Analysis Research. *Lecture Notes in Computer Science*, 3568, 1-10.

Hurst, W., Maass, G., Muller, R., & Ottmann, T. (2001). The "Authoring on the Fly" System for Automatic Presentation Recording. *Proc. CHI 2001 Extended Abstracts,* 5-6.

Jancke, G., Grudin, J., & Gupta, A. (2000). Presenting to Local and Remote Audiences: Design and Use of the TELEP System. *Proc. CHI 2000,* 384-391.

Keating, D.P. & Hertzman, C. (Eds.) (1999). Developmental health and the Wealth of nations: Social, biological, and educational dynamics. New York: Guilford Press.

Kientz, J.A., Boring, S., Abowd, G.D., & Hayes, G.R. (2005). Abaris: Evaluating Automated Capture Applied to Structured Autism Interventions. *Proc. Ubicomp 2005*, 323-329.

Kilgore, R., Chignell, M., & Smith, P. (2003). Spatialized audioconferencing: What are the benefits? *Proc. IBM CASCON'03*, 135-144.

Kraut, R., Egido, C. & Galegher, J. (1988). Patterns of contact and communication in scientific collaboration. *Proc. CSCW'88*, 1-12.

Marchionini, G., & Geisler, G. (2002). The Open Video Digital Library. *D-Lib Magazine*, 8(12).

Matthews, D., & Zijdemans, A. S. (2001) Toward a Learning Society Network: How being one's brother's keeper is in everyone's self interest*. Orbit V31*, N4, 50-54.

Moran, T.P., Palen, L., Harrison, S., Chiu, P., Kimber, D., Minneman, S., van Melle, W., & Zelweger, P. (1997). I'll Get that Off the Audio: A Case Study of Salvaging Multimedia Meeting Records. *Proc. CHI 1997,* 202-209.

Mukhopadhyay, S., & Smith, B. (1999). Passive Capture and Structuring of Lectures. *Proc. ACM Multimedia 1999,* 477-487.

Nardi, B.A., Kuchinsky, A., Whittaker, S., Leichner, R., & Schwarz, H. (1996). Video-As-Data: Technical and Social Aspects of a Collaborative Multimedia Application. *Computer Supported Cooperative Work* 4, 73-100.

National Research Council (U.S.), Committee on a National Collaboratory (1993). *National Collaboratories: Applying Information Technology for Scientific Research*. National Academy Press.

Olson, G.M., Finholt, T.A., & Teasley, S.D. (2000). Behavioral Aspects of Collaboratories. In S.H. Koslow & M.F. Huerta (Eds.). *Electronic Collaboration in Science*. Lawrence Erlbaum Associates.

Pea, R.D. (2005). Video-As-Data and Digital Video Manipulation Techniques for Transforming Learning Sciences Research, Education and Other Cultural Practices. To appear in J.

Weiss, J. Nolan & P. Trifonas (Eds.). *International Handbook of Virtual Learning Environments*. Kluwer Academic Publishing, in press.

Pea, R.D., & Hay, K. (2003). *Report to the National Science Foundation: CILT Workshop on Digital Video Inquiry in Learning and Education, November 25-26, 2002*. Stanford Center for Innovations in Learning.

Pea, R.D., Mills, M., Rosen, J., Dauber, K., Effelsberg, W., & Hoffert, E. (2004). The Diver Project: Interactive Digital Video Repurposing. *IEEE Multimedia* 11(1), 54-61.

Rankin, K., Baecker, R.M., and Wolf, P. (2004). ePresence; An Open Source Interactive Webcasting and Archiving System for eLearning, *Proceedings E-Learn 2004*.

Rounding, M. and Greenberg, S. (2000). Using the Notification Collage for Casual Interaction. *ACM CSCW 2000: Workshop on Shared Environments to Support Face-to-Face Collaboration.* Philadelphia, Pennsylvania, USA, December.

Rui, Y., Gupta, A., & Grudin, J. (2003). Videography for Telepresentations. *Proc. CHI 2003*, 457-464.

Shick, R. (2005). A Study of Student Conversation in Text and Audio During Webcast Lectures, M.Sc. Thesis, Dept. of Computer Science, University of Toronto, July 2005.

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., 822-829.

Schunn, C., Crowley, K., & Okada, T. (2002). What Makes Collaborations Across a Distance succeed? The Case of the Cognitive Science Community. In P. Hinds & S.B. Kiesler (Eds.), *Distributed Work*, MIT Press.

Smith, B.K. & Blankinship, E. (1999). Imagery as Data: Structures for Visual Model Building. *Proc. Computer Support for Collaborative Learning* 99, 549-557.

Smith, M., Cadiz, J.J., & Burkhalter, B. (2000). Conversation Trees and Threaded Chat. *Proc. ACM CSCW 2000*, 97-105.

Stevens, R. (2006). Capturing Ideas in Digital Things: A New Twist on the Old Idea of Inert Knowledge. In , in Goldman et al. (2006), this volume.

van Buren, D., Curtis, P., Nichols, D.A., & Brundage, M. (1995). The AstroVR Collaboratory: An On-Line Multi-User Environment for Research in Astrophysics. In R.A. Shaw, H.E. Payne, and J.J.E. Hayes (Eds.). *Astronomical Data Analysis Software and Systems IV*, ASP Conference Series, Volume 77, Astronomical Society of the Pacific.

Wactlar, H.D., Christel, M.G., Gong, Y., & Hauptmann, A.G. (1999). Lessons Learned from Building a Terabyte Digital Video Library. *IEEE Computer*, Feb. 1999, 66-73.

Weber, Steven (2004). *The Success of Open Source*, Harvard University Press.

White, S.A., Gupta, A., Grudin, J., Chesley, H., Kimberly, G., & Sanocki, E. (2000). Evolving Use of a System for Education at a Distance. *Proc. HICSS-33.*

Whittaker, S. (1995). Rethinking Video as a Technology for Interpersonal Communications: Theory and Design Implications. *International Journal of Human-Computer Studies* 42(5), 501-529.

Wiley, David. A. (2002). Connecting Learning Objects to Instructional Design Theory: A Definition, a Metaphor, and a Taxonomy. Section 1.1. of the online version of *The Instructional Use of Learning Objects* (Bloomington, IN: Agency for Instructional Technology), http://www.reusability.org/read/

Wolf, W., Liang, Y., Kozuch, M., Yu, H., Phillips, M., Weekes, M., & Debruyne, A. (1997). A Digital Video Library on the World Wide Web. *Proc. ACM Multimedia 1997*, 433-4.

Wulf, W.A. (1989). The National Collaboratory: A White Paper. In J. Lederberg & K. Uncaphar (Eds.). *Towards a National Collaboratory: Report of an Invitational Workshop at the Rockefeller University*. National Science Foundation.

Zijdemans, A., Moore, G., Baecker, R.M., and Keating, D.P. (2005). ePresence Interactive Media and WebForum 2001: An Accidental Case Study on the Use of Webcasting as a VLE for Early Child Development. In Weiss, J., Nolan, J., and Trifonas, P. (Eds.), *International Handbook of Virtual Learning Environments*, Kluwer, in press.

BAECKER FIGURES AND TABLES

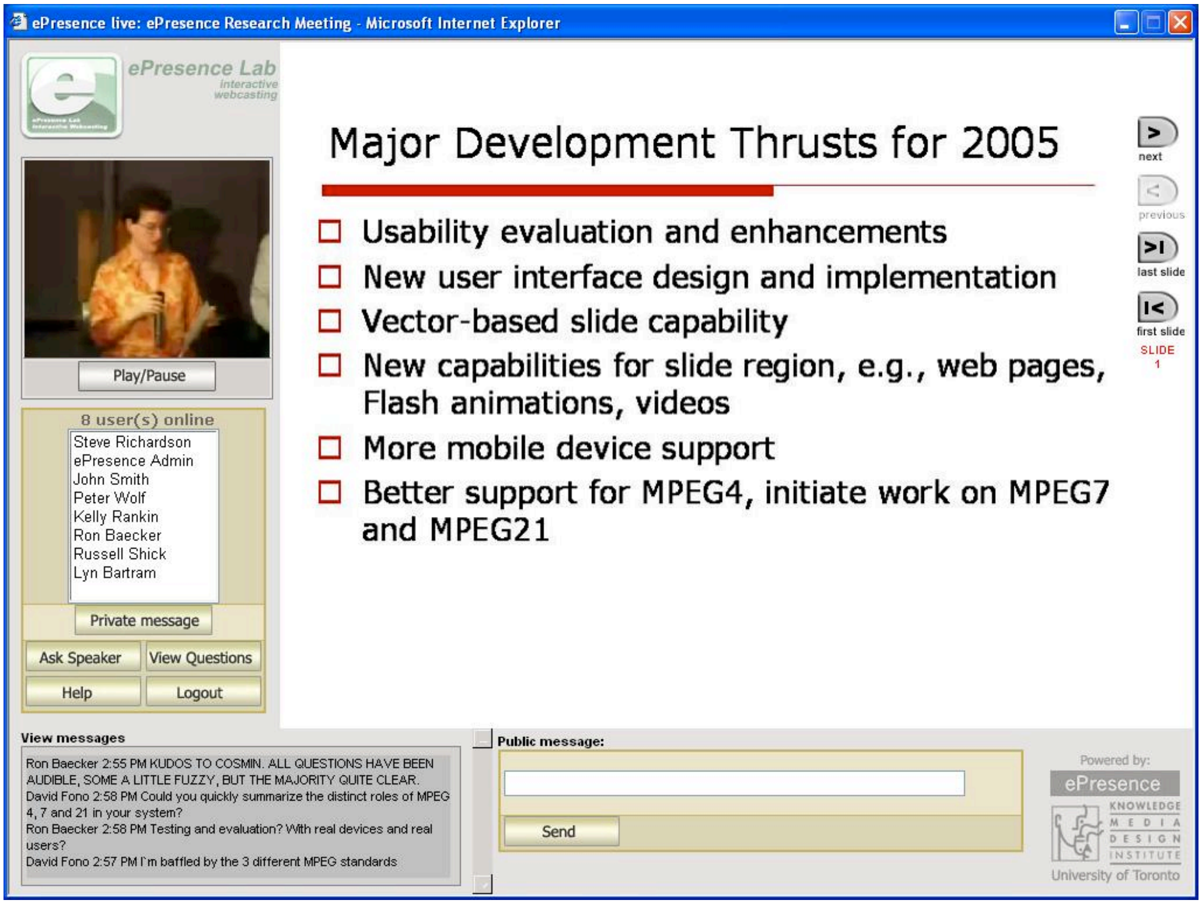

Figure 1. A screen shot from a live webcast.

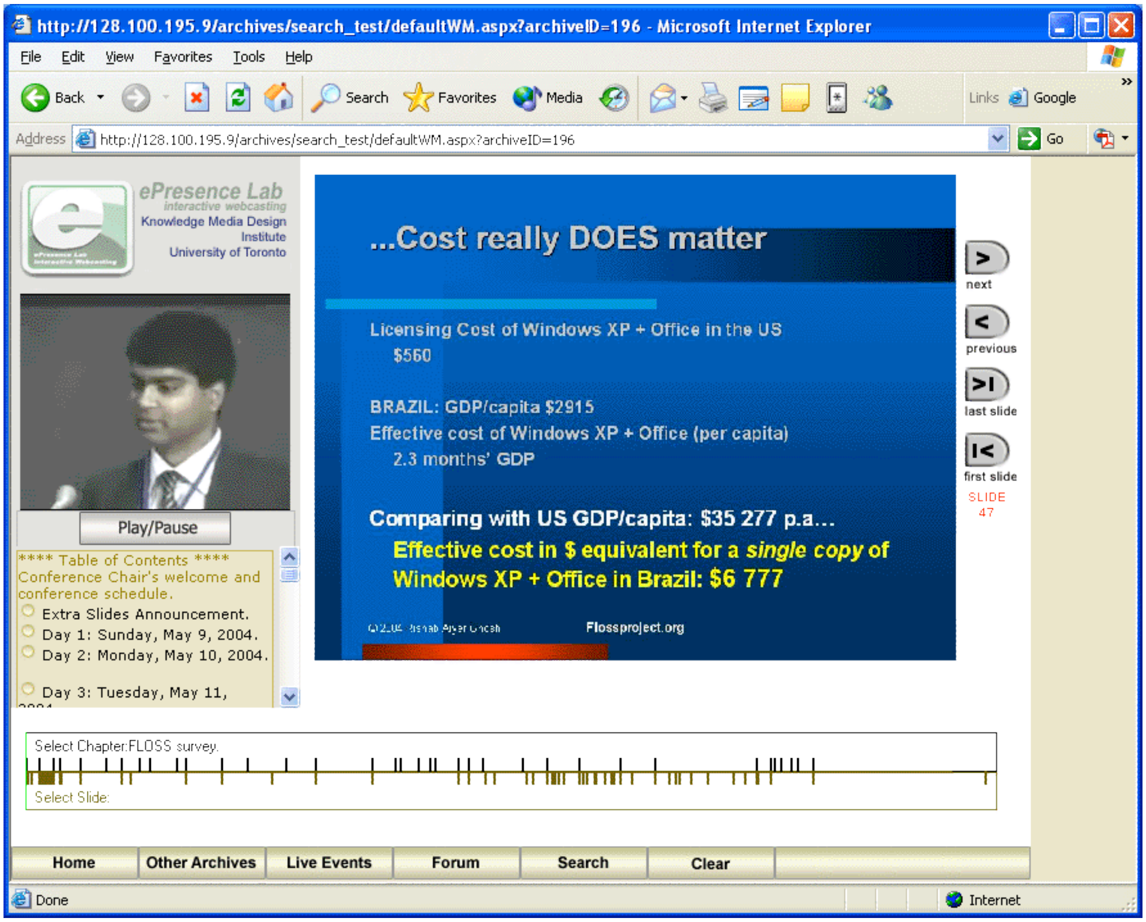

Figure 2. A screen shot from the interface to an archive constituting the multimedia proceedings of a conference on open source and free software.

	a.m. Day1	p.m. Day1	a.m. Day2	p.m. Day2
Content-related	11	5	13	16
Technology-related	116	112	44	41
Administrative	38	21	13	10
Social	30	1	28	30
Other	18	19	13	14

Table 1: Categorizing chat messages over the four half-days of the webcast of WebForum 2001

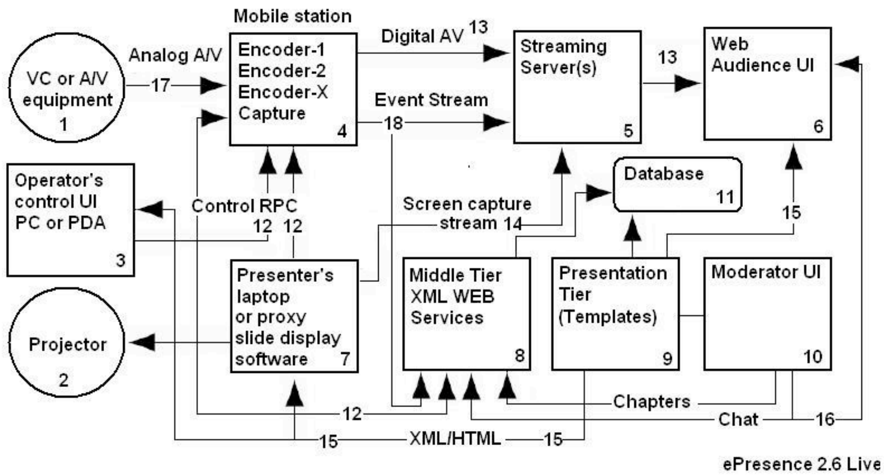

Figure 3: Current ePresence system architecture for live webcast

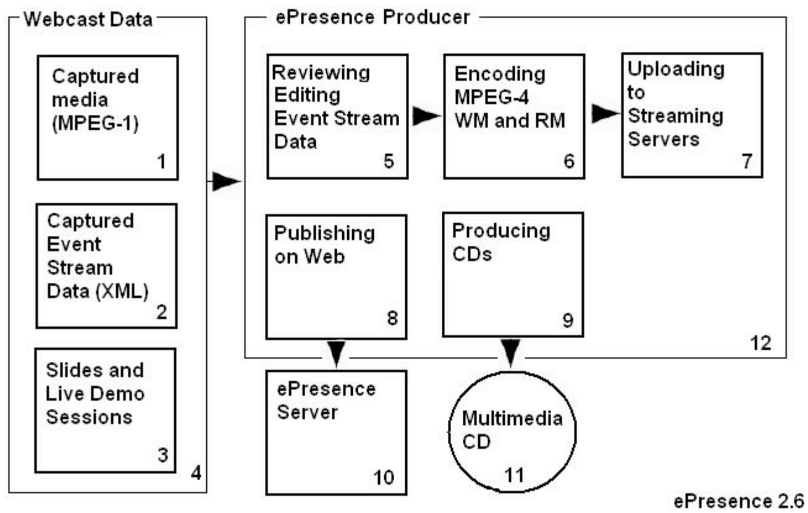

Figure 4: Producing an archive of an ePresence webcast

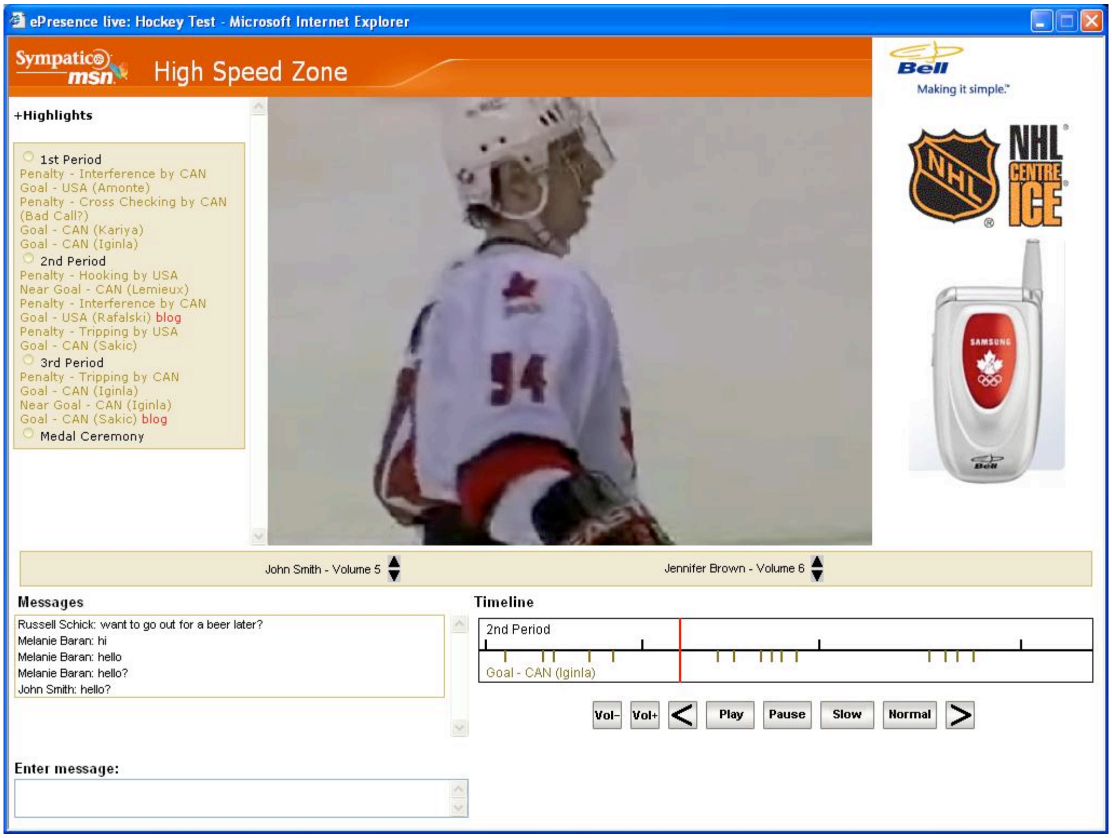

Figure 5: A consumer prototype of an ePresence-based video collaboratory.