Social Computing

Toward Connected Shared Experiences

Pablo Cesar
Centrum Wiskunde & Informatica: CWI

Rene Kaiser
Institute for Information and Communication Technologies, JOANNEUM RESEARCH

Marian F. Ursu
University of York and Goldsmiths, University of London

To move beyond the current "talking heads" paradigm and enable truly natural and immersive shared experiences, video-mediated communication systems must be able to understand the context of the shared activity as well as the social layer of interaction.

With the increasing number of heterogeneous devices accessing the Internet and many people using multiple devices—sometimes simultaneously—to communicate, the notion of connected shared experiences is gaining momentum within the social networking and social computing communities. In this column, the authors highlight a paradigm shift in video-mediated communication from "talking heads," in which we see each other's faces but not much else, to more natural and intimate interactive environments. In such environments, the system's ability to adapt to the context and to social cues is critical to achieving a satisfactory quality of experience. The authors outline fundamental system requirements and call for the development of new use cases.

Christian Timmerer, column editor

People are social beings: whether engaging in work, play or social conversations, we seek to interact with others. Current technologies support various types of interaction among individuals at different locations—for example, remote learning, connected healthcare, and multiparty videoconferencing—but are limited in both coverage and flexibility. However, research in this area has flourished in recent years and is moving communication from static interpersonal exchanges to dynamic shared experiences. The challenge is to export promising results from the lab to the world.

The following provides an overview of the present state of technology in this field through a few examples and discuss future possibilities and challenges. We highlight a key research goal: to provide adaptive communication systems that react not only to changes at the network and application layers, but also at the social layer of interaction.

Connected Shared Experiences

Linking multiple people across physical spaces is not a new idea. As Figure 1 shows, connected shared experiences were envisioned soon after the telephone's invention in 1876. Since then, academic and industry efforts to achieve immersive group communication have met with varying degrees of success.[1] Recent developments include the rapid adoption of video-mediated communication technologies by home users[2] and children reading with long-distance family members via background-subtracted video communication.[3]

We have made great strides in the past century and a half but still have a long way to go. The next step, as discussed in a workshop at the 2013 ACM SIGCHI Conference on Human Factors in Computing Systems (http://clab.iat.sfu.ca/beyondtalkingheads), is to move beyond the "talking heads" paradigm that characterizes most video chat systems, like Skype and Google Hangouts, to provide rich communication support for everyday activities such as holiday get-togethers, outdoor sports events, shared meals, birthday parties, storytelling, or just "hanging out."

Figure 1. Early conceptions of connected shared experiences combined recently invented telephony with motion-picture-like projections. (a) George du Maurier's "Telephonoscope" (1878). (b) Postcard depicting "Correspondence Cinema" (1910).

Videoconferencing is no longer confined to users gathered in traditional meeting rooms, but it is quickly moving to the living room. The combination of powerful computing devices, widespread broadband availability, and optimized video streaming protocols is enabling mass adoption of video for communication purposes. Camera-equipped laptops and free services such as Google Hangouts let people connect with one another in more natural and immersive ways than were possible only a few years ago.

Figure 2. Examples of connected shared experiences. (a) Screenshot from a prototype distributed board game system enabling remote play between users at different locations. (b) Linked performances in two dance studios at Falmouth University, UK (www.youtube.com/user/Vconect).

Three examples highlight recent advances in providing connected shared experiences.

- *Gaming*. Technology can help nurture bonding among physically separate family members by enabling them to play games as if they were together at the same location, as Figure 2a shows.[4] Developing such distributed games requires intelligent management of limited home computing power and network resources.
- *Performing arts*. 3D tele-immersion,[5] which involves transmitting 3D reconstructions of participants in different places in real time, enables live co-performance. This will encourage new forms of expression in dance, music, theatre, and other arts. Low latency is a key requirement to ensure seamless interaction.
- *Television broadcasting*. Social TV connects people watching the same program or event via different modalities.[6] For example, those watching a soccer game can share the experience with other home viewers as well as attendees.[7] Conveying emotions between fans at different locations in real time is essential.

These examples suggest some fundamental requirements for truly connected shared experiences: natural and seamless interaction, high realism and immersion, and adaptability to users' needs and available resources.

Research Challenges

These requirements present researchers with numerous challenges.

- *Context awareness*. Current communication systems are unaware of the activity, the communication context (work, family interaction, gaming, learning, etc.), and the strength of interpersonal ties. Thus, they're unable to adequately adapt to different situations and environments and so improve users' quality of experience (QoE).
- *Cue processing*. Although affordable high-definition webcams and sensors, including motion-capture cameras, are available, current communication systems don't take advantage of their capabilities. For example, facial expressions and body language could be captured to infer conversation patterns and interaction quality among participants in real time.
- *Automatic adaptation*. Communication systems aren't able, based on the context and cues alone, to ensure that individual participants always see and hear what is most relevant to them at a given moment. Such expertise, in fact, does not even exist yet. Research is building a body of knowledge that informs how to construct such adaptation policies—what is known as *communication orchestration*.[8]
- *QoE evaluation*. Currently, metrics aren't available to adequately measure the QoE of distributed gatherings. Such metrics should consider various factors such as social presence, naturalness, level of immersion, and "togetherness."[4] These evaluations are key to assessing automatic adaptation processes.

Solving these challenges will open new opportunities for the realization of connected shared experiences. Still, the obstacles ahead are significant. Real-time content analysis algorithms that enable communication systems to understand what's going on are very difficult to construct, given the inherent noise. We're still in the infancy of comprehending the effect of each remote activity and the level of intimacy on parameters affecting QoE. Standards and methodologies for assessing QoE are nonexistent. Automatic adaptation mechanisms that consider the network, application, and social layers are rather basic and limited to specific use cases. This article is a call for the research community to focus on these fundamental research questions that will in turn enable new use cases, like distributed performing arts reaching large interactive audiences at home, as in Figure 2b.

Technological advances such as more efficient video-streaming protocols, more powerful sensors, faster processors, and improved networking abilities aren't sufficient to move beyond today's talking heads paradigm and enable connected shared experiences. Video-mediated communication systems must also be able to understand the nature of the shared activity and social cues to make such experiences natural and immersive.

Key challenges to realizing this vision include context awareness, cue processing, QoE monitoring, and reasoning processes that dynamically adapt to environmental changes at the networking layer (such as longer end-to-end delays), application layer (turn changes during board game play), or social layer (the fact that someone *doesn't* talk, for example). To achieve such automatic adaptation, new reasoning approaches and algorithms need to be developed and studied.

Acknowledgment

This work was supported in part by funding from the European Community's Seventh Framework Programme (FP7/2007-2013), grant agreement no. ICT-2011-287760.

References

1. P.A. Chou, "Advances in Immersive Communication: (1) Telephone, (2) Television, (3) Teleportation," *ACM Trans. Multimedia Computing, Comm., and Applications*, vol. 9, no. 1, 2013, article no. 41.
2. D.S. Kirk, A. Sellen, and X. Cao, "Home Video Communication: Mediating 'Closeness,'" *Proc. 2010 ACM Conf. Computer Supported Cooperative Work* (CSCW 10), 2010, pp. 135–144.
3. S. Follmer et al., "People in Books: Using a FlashCam to Become Part of an Interactive Book for Connected Reading," *Proc. ACM 2012 Conf. Computer Supported Cooperative Work* (CSCW 12), 2012, pp. 685–694.
4. I. Kegel et al., "Enabling 'Togetherness' in High-Quality Domestic Video," *Proc. 20th ACM Int'l Conf. Multimedia* (MM 12), 2012, pp. 159–168.
5. A. Arefin, et al., "Classification and Analysis of 3D Teleimmersive Activities," *IEEE MultiMedia*, vol. 20, no. 1, 2013, pp. 38–48.
6. P. Cesar and D. Geerts, "Understanding Social TV: A Survey," *Proc. 2011 Networked & Electronic Media Summit* (NEM Summit 11), 2011; http://homepages.cwi.nl/%7Egarcia/material/nem-summit2011.pdf.
7. C. Hesselman et al., "Sharing Enriched Multimedia Experiences across Heterogeneous Network Infrastructures," *IEEE Comm. Mag.*, vol. 48, no. 6, 2010, pp. 54–65.
8. M.F. Ursu et al., "Orchestration: TV-Like Mixing Grammars Applied to Video-Communication for Social Groups," *Proc. 21st ACM Int'l Conf. Multimedia* (MM 13), 2013, pp. 333–342.

Pablo Cesar *leads the Distributed and Interactive System group at Centrum Wiskunde & Informatica (CWI), Amsterdam, the Netherlands. Contact him at p.s.cesar@cwi.nl.*

Rene Kaiser *is a key researcher at the Institute for Information and Communication Technologies, JOANNEUM RESEARCH, Graz, Austria. See contact information at http://about.me/rene.kaiser.*

Marian F. Ursu *is a professor of interactive media at the University of York and a professor of computing at*

Goldsmiths, University of London, UK. Contact him at marian.ursu@york.ac.uk.

Computer's Social Computing column is closely connected with the Special Technical Community on Social Networking (STCSN). Further information about the STCSN, as well as featured articles and monthly bulletins, can be found at www.computer.org/stcsn. Come and join us now!

social computing, video-mediated communication, videoconferencing, video chat, Internet/Web technologies