Improving Remote Collaboration through Side-by-Side Telepresence

Paul Tanner
Carnegie Mellon University
Pittsburgh, PA 15213 USA
subcog@gmail.com

Varnali Shah
Carnegie Mellon University
Pittsburgh, PA 15213 USA
varnali@gmail.com

CHI 2010, April 10–15, 2010, Atlanta, Georgia, USA.
ACM 978-1-60558-930-5/10/04.

Abstract
Virtually all teleconferencing solutions are designed to facilitate face-to-face interactions. While face-to-face is suitable for meetings or conversations, we see many real-world situations where people choose to sit in other configurations. Face-to-face telepresence inaccurately simulates these alternate interaction styles. In this paper we describe a side-by-side telepresence concept, which is more appropriate for side-by-side style interactions, such as collaborative writing or training. We explore the differences between face-to-face and side-by-side telepresence, and discuss our prototype side-by-side telepresence workstation.

Keywords
CSCW, Videoconferencing

ACM Classification Keywords
H.5.3 Synchronous Interaction, H.1.2 Human Factors, H.5.1 Video

General Terms
Design, Human Factors

Introduction
Current efforts to improve telepresence tend to focus on either creating high-definition face-to-face

experiences, expanding the number of users that can be involved in a session, or expanding the experience to use more of the senses. While these directions can enhance the immersion of a meeting and allow greater numbers of users to participate, these improvements assume that face-to-face contact is always the most appropriate physical layout for human interaction.

We are proposing a new side-by-side telepresence layout that we believe is more appropriate for certain kinds of interactions. Combined with screen capture technologies, the proposed side-by-side telepresence layout can properly simulate a work environment where users sit next to each other to work on collaborative tasks.

Limitations of Face-to-Face Telepresence

Traditional telepresence is excellent for simulating face-to-face communication. Users can share a casual video-chat, or teams can meet remotely in a high-end telepresence conference room. The physical layout places the focus on the users, and thus de-emphasizes documents or artifacts that the users are working with. This is typically good for meetings where people express ideas, report progress, or assign tasks.

We have observed however that face-to-face telepresence is inappropriate for certain kinds of common interactions, such as collaborating on a document with a partner, an example that has become common since the release of *Google Docs*. In these cases, we noticed that face-to-face video conferencing significantly distracts from the task, as users are pushed into focusing on each other instead of the artifact. Users are constantly making accidental eye contact, which can create a feeling of awkwardness.

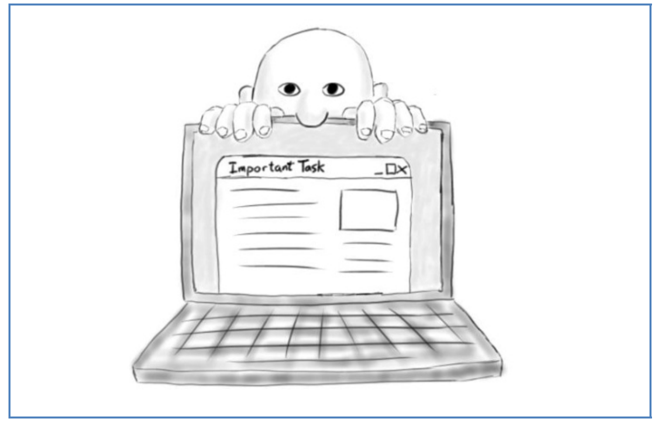


figure 1. The feeling of uneasiness that face-to-face video conferencing can cause while a user works on a task.

We came to understand that the face-to-face video chat was actually an incorrect simulation of many collaborative interactions. When two people in the same room work together on a task, document, or other kind of artifact, they tend to sit side-by-side, not face-to-face.

Benefits of Side-by-Side Interactions

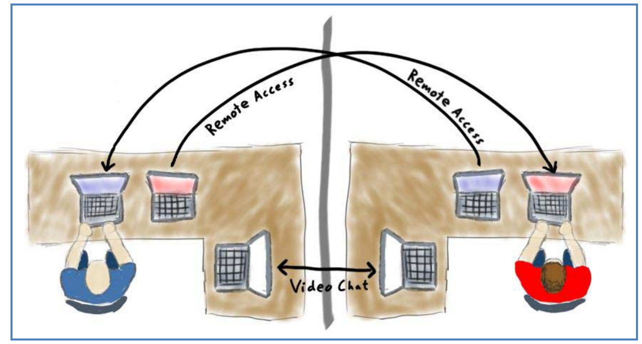
There are benefits to collaborating side-by-side. Both partners are focused on the artifact in front of them, rather than on each other. They are free to communicate verbally, but eye contact is only made when users turn to see their partner. This allows users to transition their focus freely and naturally between the artifact and their partner.

Side-by-Side Telepresence Concept

We began to consider alternative physical layouts for telepresence systems to better simulate a side-by-side working environment. We chose to simulate a physical layout that we were familiar with: Two people sitting side-by-side, each with a laptop in front of them. This layout focuses the users on the laptops, while allowing the users to seamlessly shift their focus to their partners as frequently as is needed. Each user can see their partner's screen, and even reach over and interact with it when necessary.

We prototyped our concept by putting together two separate workstations, each consisting of three laptops and a webcam. We used free software, including a remote-access software solution called *TeamViewer*, and *Gmail*'s embedded video chat client. For each workstation, the first laptop was used as the user's primary screen, with a second laptop placed next to it. *TeamViewer* connects the user's second laptop to their partner's first laptop so that the user can view and interact with their partner's screen. The partner's workstation is connected similarly, so that both users can view and interact with each system. The third laptop at each station was connected to a webcam and ran a simple video chat. We positioned each of the third laptops to the sides of the users, to capture a profile view of each user. The two workstation layouts are mirrors of each other, so that one user sits on the left and the other sits on the right.

In this setup, users have nearly identical views of the artifacts in front of them. This solves common telepresence questions of where and how to display work artifacts. This setup facilitates both WYSIWIS (What you see is what I see) and WYSSITYS (What you see is what I think you see) interactions [1]. WYSIWIS is supported as both users have identical views of the work on both laptops. WYSIWITYS is supported as both users can visually see the general angle that their



figures 2 and 3. Two remote workstations prototyped to simulate a side-by-side work environment.

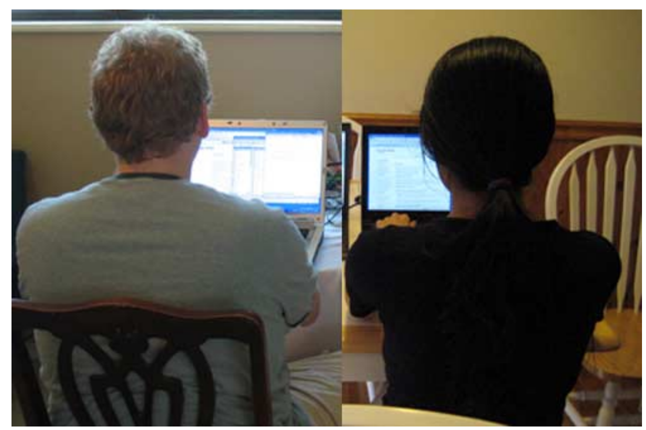


figure 4. We want to create a solution that feels like users are sitting next to each other.

partner is looking to gain awareness of which screen the partner is looking at. Because users can see each others' mouse cursors, users can also direct their partner's focus to specific on-screen elements.

Preliminary Evaluation with Users

We ran an informal user test, to better understand the strengths and weaknesses of the side-by-side concept. We recruited users to try out our system and give us feedback on their experience.

The first scenario was an open-ended collaborative task. We asked our users to organize a Christmas party for a fictitious company, including choosing a venue, entertainment, and food. The first pair of users expressed that they loved the screen-sharing, but were mildly frustrated that they had to turn so far to the side to see their partner. They indicated that they would have preferred a face-to-face video chat. Upon further discussion, we realized that we had miscommunicated the scenario, and they had thought that they were supposed to be having a meeting, which would be more suitable for face-to-face interactions.

We decided to test this scenario again with a second set of participants. This time, we explicitly stated that the scenario was about collaborative working and not a planning meeting. To reinforce the collaborative-work nature of the scenario, we also added an additional requirement: that the pair turns in a document describing the event plans as a final deliverable. We also tweaked the workstations, moving the video conference laptops forward slightly, so that they could be seen in each user's peripheral vision, to make it slightly easier for partners to see each other. These changes appear to have been effective, as the second pair responded much more positively to the test, expressing that the experience was highly immersive, and pointing out tasks that they would use such a system for in their own work.

figure 5. Users collaborating remotely through the prototype workstations.

The second scenario was an open-ended training task. We asked a pair of users to choose a computer-based task that one partner knew how to do and the other did not. The pair selected "creating a blog" as their task.

After the task, they expressed satisfaction with how similar the experience was to sitting next to someone. They also explained that having two computers in front of them encouraged the learner to work through the instructions from the trainer on her own system, helping to reinforce the learning.

We hope to conduct more formal experiments in the future, as we continue to iterate our workstation and expand the kinds of interactions it supports.

Discussion

Although we did not conduct a formal comparison between our prototype and face-to-face solutions, we were able to observe aspects of the system that work well and we received feedback from our users. Our users found our setup to be intuitive and immersive, and they could use the system with almost no training. This can be largely attributed to properly simulating a physical work layout. This builds a sense of co-location for users, which allows them to interact naturally, without having to be trained on how to use a new interface. "When collaborators are co-present and have smooth access to views of shared work spaces, they can easily identify each others' focus of attention, monitor facial expressions, body orientations and actions, and assess whether their utterances have been adequately comprehended [2]".

Another advantage we found in our system over traditional face-to-face telepresence is that because users are not facing each other by default, it is extra clear when users shift focus to each other. We achieve a similar effect that Sellen described with his Hydra system (although with only two users): "When person A turns to look at person B, B is able to see A turn to look towards B's camera" [3]. The act of shifting focus is an event that emphasizes that focus has been shifted.

Although our prototype was not designed to facilitate exact eye-contact, and we did not focus on the exact angles of the cameras, or sizes and positions of the monitors, our users still indicated that they felt like they were interacting naturally, and did not complain about problems with eye-contact. In regards to eye-contact, it appears that our system is more forgiving than face-to-face telepresence systems. We have two hypotheses to explain this:

1. Because we were using a standard (low resolution) video-chat, it may be that users were less perceptive of small differences in the angles of gaze of their partners. It's possible that using a high-definition video-conference would cause users to notice problems in eye-contact.
2. Because users are not facing each other by default, it may be that eye contact is rarely sustained anyway. It's possible that users mostly glance toward each other to confirm they have their partners' attention, and that neither sustained or exact eye contact is expected. We have not yet pursued further research in this area.

Although our prototype is far less expensive than a professional telepresence conference room, the cost of hardware (3 computers per user) may still be prohibitive to many people. We can identify ways to lower the cost in terms of hardware, power consumption, and desk real-estate.

Although our prototype used six laptops, we believe that a workstation can be created that uses a single CPU for all three screens, although we have not explored this further and it may require custom software to run two separate sets of mouse and keyboard inputs, as well as a video chat and two simultaneous screen-sharing sessions. We envision a software solution that is as simple as a chat client, where a user can see which friends are online, and connect to a partner with a click or two.

Future Directions

In the future we want to explore expanding the side-by-side telepresence model in several directions.

- Including additional participants. We believe we can make side-by-side telepresence work well with 3 participants, and we are hoping to find ways to expand it to 5 or 6.
- Combining side-by-side and face-to-face models to simulate round table meetings or classroom environments. We see potential for creating an environment where a user can collaborate side-by-side with partners while interacting face-to-face with a teacher or another team.
- We also see an opportunity to include a tablet component, so that users can sketch ideas as they collaborate.

Potential Applications:

- Collaborative document writing
- Software training or support
- Remote classrooms

References

[1] Smith, R. B. 1992. What you see is what I think you see. SIGCUE Outlook 21, 3 (Feb. 1992), 18-23.

[2] Naomi Yamashita, Keiji Hirata, Toshihiro Takada, Yasunori Harada, Yoshinari Shirai, Shigemi Aoyagi. Effects of Room-sized Sharing on Remote Collaboration on Physical Tasks. IPSJ Journal, Digital Courier, Vol. 3, 2007.

[3] Sellen, A., Buxton, B., and Arnott, J. 1992. Using spatial cues to improve videoconferencing. CHI '92, pp. 651-652, 1992.