Ledo, D., Aseniero, B.A., Greenberg, S., Boring, S. and Tang, A. (2013)
OneSpace: Shared Depth-Corrected Video Interaction. In Proc. ACM CHI 2013 Extended Abstracts - CHI'13. (Paris, France), 6 pages, April 27 - May 3. A CHI Work in Progress. Includes Poster.

ONESPACE: Shared Depth-Corrected Video Interaction

David Ledo
Interactions Lab
University of Calgary
2500 University Dr NW
Calgary, Alberta, T2N 1N4
dledomai@ucalgary.ca

Bon Adriel Aseniero
Interactions Lab
University of Calgary
2500 University Dr NW
Calgary, Alberta, T2N 1N4
b.aseniero@ucalgary.ca

Saul Greenberg
Interactions Lab
University of Calgary
2500 University Dr NW
Calgary, Alberta, T2N 1N4
saul.greenberg@ucalgary.ca

Sebastian Boring
Department of Computer Science
University of Copenhagen
Njalsgade 128, Bldg. 24
2300 Copenhagen S
sebastian.boring@diku.dk

Anthony Tang
Interactions Lab
University of Calgary
2500 University Dr NW
Calgary, Alberta, T2N 1N4
tonyt@ucalgary.ca

Copyright is held by the author/owner(s).
CHI 2013 Extended Abstracts, April 27–May 2, 2013, Paris, France.
ACM 978-1-4503-1952-2/13/04.

Abstract
Video conferencing commonly employs a video portal metaphor to connect individuals from remote spaces. In this work, we explore an alternate metaphor, a shared depth-mirror, where video images of two spaces are fused into a single shared, depth-corrected video space. We realize this metaphor in OneSpace, where the space respects virtual spatial relationships between people and objects as if all parties were looking at a mirror together. We report preliminary observations of OneSpace's use, noting that it encourages cross-site, full-body interactions, and that participants employed the depth cues in their interactions. Based on these observations, we argue that the depth mirror offers new opportunities for shared video interaction.

Author Keywords
Video communication; media spaces.

ACM Classification Keywords
H.5.3. Group and Organization Interfaces: Computer-supported cooperative work.

Introduction
Enabling synchronous interaction between people separated by physical distance has long been a principal concern for CSCW research. The core vision underlying considerable work in this space is to support interaction

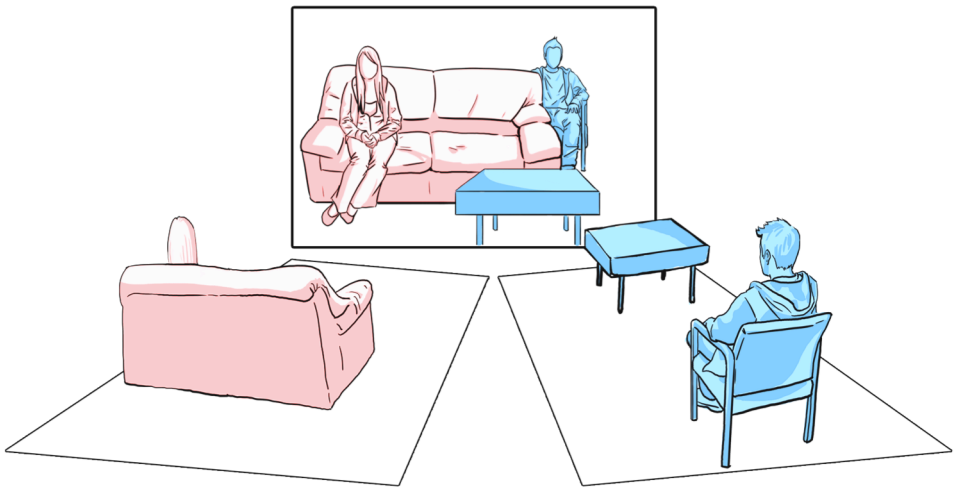


Figure 1. OneSpace integrates two remote spaces (bottom right and left) into a single space (top) by presenting a virtual depth mirror of both spaces.

with remote people as if they were co-present. To support "face-to-face" conversation and meetings, the most common approach has been to employ a media space, where an audio-video link is established between two remote spaces (i.e. video conferencing) [2]. We call this the "video portal" metaphor, as the system connects two virtual spaces through a virtual portal.

Our interest here is revisiting an alternate metaphor, that of a mirror [9]. The primary problem with the original implementation was that depth cues were not preserved—that is, one scene was always "in front" of another. Here, we explore a revision: a "depth-mirror", which still looks like a mirror, except that it preserves the *depth* cues for each location. As illustrated in Figure 1, people see themselves and interact with others in a shared video scene that looks like a mirror; in this mirror, objects and people are overlaid with correct depth cues. We were inspired by the video-based interfaces introduced by Krueger [7], and more recently popularized by video game systems, where people interact with a mirrored video image of themselves. This approach creates a virtual "stage" for interaction, and as we will see, fundamentally changes how people interact with one another.

Our preliminary observations show that the depth-corrected feed encourages a broad range of rich, playful interactions that go beyond a traditional chroma-key implementation without proper depth cues [8]. The depth cues provide people with a shared, negotiated stage for their shared interactions, where the negotiation occurs merely through one's closeness with the video scene (just like in a mirror—only one person can be "in front" at once).

Related Work

Researchers have long used video as a means to allow people to interact with one another as if they were in a collocated space.

Conversation through a portal. A traditional media space employs an audio/video link with the remote space. Here, the video link is a portal or tunnel that connects remote spaces, primarily for conversation [2].

Shared workspaces for tasks. Rather than focusing specifically on conversation, video has also been used to fuse two separate workspaces into a single shared workspace for task work. These generally project a video feed from the remote workspace onto the local space (e.g. [6, 11]). The result is a single "workspace" that allows people to interact through shared artifacts (or drawings). The metaphor being implied here is of a



Figure 2. OneSpace in action

"shared workspace," where all parties are effectively "sitting on one another's laps." Of interest is that the metaphor changes how people interact: here, the interaction allows for gesture, rather than solely through conversation. MirrorFugue [13] explores this interaction within a musical context, where the focus is on the placement and movement of fingers over a shared/mirrored piano keyboard.

Shared stage. Krueger's original Videoplace work realized a vision to connect remote spaces through full-body silhouettes that were simultaneously projected onto a large wall-sized display [7]. HyperMirror [9] also explores this concept of a shared stage, through a mirror metaphor. Here, video captured from remote spaces are fused through chroma-keying effects, with the resulting fused image (akin to a mirror) projected onto a large display. This mirror metaphor encouraged self-reflection, and accordingly, a more relaxed conversational environment. Hill et al. [4] also explored this metaphor, using virtual embodiments instead of video.

Both shared workspace and shared stage models fuse remote spaces together rather than keep them separate, as in the video portal model. Whereas the apparent spatial relationships between the remote spaces are fixed in a video portal model (i.e. people remain in their respective locations), shared spaces afford dynamic reconfigurations of these spatial relationships. The shared models allow people to "move around" with respect to one another, allowing for different spatial dynamics to emerge. For instance, Morikawa et al. [9], in observing people interact through HyperMirror, report that people felt closer to those who were seen to be close *in the shared mirror space* rather than those who were physically co-present! Thus, these apparent spatial relationships meaningfully affect how people interact with one another.

Thus, the shared stage model allows the dynamics of these spatial relationships play out. One fundamental problem with previous implementations is that while they preserve the apparent planar relationships on screen (i.e. X-Y relationships), they generally gloss over the depth relationships (i.e. Z-ordering). VideoPlace employed silhouettes, while HyperMirror used chroma-key effects, effectively always placing one space atop another. Our work also realizes a shared stage model, and builds on HyperMirror's implementation by also adding depth information to the video feed. As we will see, this substantially changes the space of possible interactions. We note that others are concurrently pursuing somewhat similar work (e.g. InReach[1]).

OneSpace

OneSpace integrates remote spaces through a shared depth-mirror metaphor. Having depth integrated allows for respecting the location, distance and orientation between people and objects in the shared space. OneSpace can fuse any number of real locations into a single virtual space (we have tested it with four environments) while respecting the spatial relationships of people and objects in the virtual space: things and people who are closer to the mirror appear in front of those who are further away. People are able to interact through the manipulation of physical objects in the space, and through body movement and motion in the space (as shown in Figure 2).

[1] InReach: http://fluid.media.mit.edu/node/179

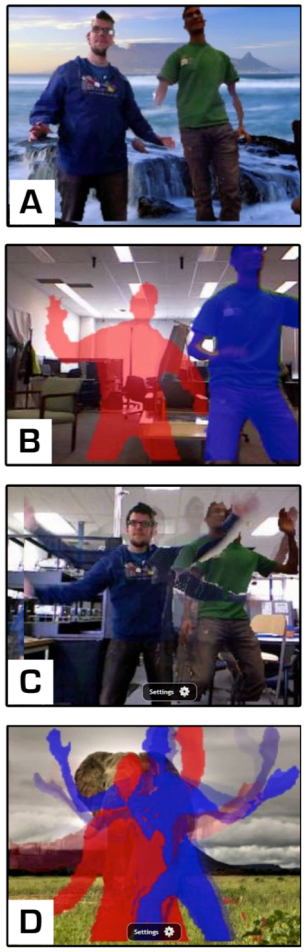


Figure 3. Some of the effects applied in OneSpace: (a) shows a static background, (b) shows the shadow effect, (c) shows traces of movement and (d) shows a mixture of the three effects.

Implementation

We implemented OneSpace as a distributed application using a client/server architecture. We make use of thin clients that send the RGB and depth data collected from connected Microsoft Kinects. The server merges this data before sending it back to the clients to be displayed. In our current setup, we use whiteboard-sized displays to show the output. OneSpace is implemented in C# WPF with the Kinect SDK.

OneSpace's server integrates the color video frames it receives from clients. On a per-pixel basis, it uses the depth information to extract the front-most color pixels to create a new video frame which is then sent back to all the clients for display. This process provides people with a mirrored image of themselves, and preserves the spatial relationships of every person and object in each space, allowing for occlusions and overlaps to occur in the final video frame. We apply standard image processing techniques to smooth the depth information, to help the resulting image appear smoother and more seamless.

Krueger's VIDEOPLACE provided a number of video effects on people's video embodiment [7] that allowed people to engage in expressive, video-based "embodied" interaction. Inspired by the opportunities for interpersonal interaction enabled by these video filters, we also designed a number of effects for OneSpace, as illustrated in Figure 3:

Environment Effects. OneSpace can use four different kinds of scenes as the surrounding environment for the interactions: (a) it can use the scene from one of the sites; (b) it can use a static image as background; (c) it can employ a pre-recorded 3D scene (with both color and depth information); and (d) it can loop a video that contains depth information, to encourage interactions with scenes in motion, similar to Looking Glass [1]. These changes of ambiance are important: they can create the illusion of presence in the other person's environment (when using the scene of the site as background), or can create a virtual "third place" to which people are transported together.

Shadows and traces. As with Krueger's original implementation, we can also draw foreground objects as silhouettes, allowing people to interact as shadow puppets rather than as video embodiments. We can also apply a trace effect, where ghostly trails of people's motions are overlaid atop one another. These effects encourage unique forms of interaction and playfulness, where people's bodies can be merged into one.

Preliminary Observations of Use

We made OneSpace available to several members of our institution to understand the kinds of interactions OneSpace afforded. For these tests, we connected two remote spaces through a Gigabit Ethernet connection. Each site had its own whiteboard-sized display and Kinect camera, and the two spaces were connected through a separate audio link. Typically, these tests involved groups of four people—two people per site. We only described the basic technical features of the system and did not guide their interactions. Participants had never been exposed to the system before. They were asked to use of it for 30 minutes however they wanted. This allowed us to see the kinds of experiences they created within the OneSpace environment.

Virtual physical and visual play. While we expected that people would still use the system for conversation, we



Figure 4. Participants using OneSpace to simulate a fight.

were surprised to see very little conversation at all (although there was a lot of laughter). Instead, interaction focused on the shared scene being displayed on-screen, with participants focused on how their video embodiment (i.e. their reflection) interacted with/shared the scene with video embodiments of people from the remote site on the shared "stage." Where speech did occur, it was to coordinate or guide these interactions.

These scenes were striking, as we saw our participants engage spatially with one another in ways that they would not if they were actually physically co-present. That is, they allowed their visual embodiments to interact and virtually "touch" one another in ways that would be unusual or un-comfortable in real life. For instance, a common interaction (perhaps a statement about our society) was to enact mock fist-fights with participants from the remote site. These fist fights made use of the depth-cues—for example, a punch might begin from "behind" a user, and follow through into the foreground. Here, the target would feign being hit in that direction. Perhaps as a response to these fist-fights, our participants also hugged one another, as the system would create the visual effect of these interactions in the mirror without actual physical contact. Notably, none of these participants had gotten into fist-fights or hugged one another in real life before. Figure 4 shows an example of these interactions.

Staging visual interaction. Participants also carefully staged the visual interaction with one another. In many of the fist-fights, people who were "not involved", would move out of the scene. In other cases, we observed several participants playing "headless horseman" with one another. Here, two people would stand "atop" one another in the scene, with one person leaning his head back, while the other would lean his head forward. The resulting scene would produce a humorous combination "person" with the body of one person, and the head of another. Here, the depth cues allow for interactions that would not be otherwise possible with a chroma-key solution.

We see here then that people are negotiating the use of the "stage" in two ways: in the first, people who are not involved move out of the way, while in the second, correcting the shared scene for depth allows people to alternate who takes "the stage." This stage is a flexibly negotiated space, since it merely means moving closer to the camera. Yet, it is not binary, as it would be in a chroma-keyed approach: as we saw in the "headless horseman" example, this stage is a blended area, where people can choose what "part" of their body is in front. The feedback provided by seeing one's own embodiment enables this active negotiation.

Engagement and enjoyment. Participants clearly enjoyed using our system. Much as in Social Comics [8], participants took pleasure in making one another laugh through the shared visual scene, and to create scenes that would be absurd, unusual or even impossible to enact in real life. The size of our display and capture area allowed for full-body interaction, and the shared depth-mirror metaphor allowed our participants to exploit spatial relationships. We saw them engaging in play, and immersing themselves in the activities that they created. For these reasons, we believe our system to be particularly useful for play environments and also useful to bring people together to have fun.

Conclusions and Future Work

In this paper, we introduced OneSpace, a system that performs depth-corrected integration of multiple spaces. The system supports a number of variations on the visual output, including static and 3D scenes, as well as silhouette and trace effects. Based on our preliminary observations of the system, we see how people understand and appropriate the depth-mirror metaphor for physical and visual play. We have seen that this metaphor encourages forms of shared interactions that go beyond current efforts in video conferencing, and presents a unique set of opportunities for shared video interaction across remote spaces.

Standard video conferencing will likely remain the dominant form of interaction across remote spaces. However, we have seen that OneSpace's shared depth mirror metaphor blends spaces in a way that is fundamentally different from the video portal approach (e.g. [5,11,12]). In particular, the "stage" of interaction is shared, and because it is based on depth cues, it becomes a space negotiated by one's proximity to the camera. Thus, people interact through the system in a qualitatively different manner from prior systems (e.g. [4,9]), people control these features, and use it in their interactions with one another.

There are several application areas that we want to explore with OneSpace. We believe that the playful interactions can create an interesting space for play between children. First, Yarosh et al. [14] state that a distributed children's play space should blend the representations of remote children. As OneSpace can already do this, we are interested in seeing if Yarosh's expectations are correct. Second, we believe that OneSpace can provide a means to support physiotherapy, where the depth cues can aid teaching movements and poses. Both these application areas would also serve as case studies that provide a better understanding of the affordances provided by the shared depth-mirror.

References

[1] Aseniero, B. A. and Sharlin, E. (2011). The looking glass: visually projecting yourself to the past. Proc. ICEC '11, 282-287.

[2] Bly, S., Harrison, S. and Irwin, S. (1993). Media spaces: bringing people together in a video, audio and computing environment. Communications of the ACM 28-46.

[3] Dourish, P. and Bly, S. (1992). Portholes: supporting awareness in a distributed work group. Proc. CHI '92, 541-547.

[4] Hill, A., Bonner, M. N., and MacIntyre, B. (2011). ClearSpace: mixed reality virtual teamrooms. Proc. HCI Intl '11, 333-342.

[5] Ishii, H., and Kobayashi, M. (1992). ClearBoard: a seamless medium for shared drawing and conversation with eye contact. Proc. CHI '92, 525-532.

[6] Junuzovic, S., Inkpen, K., Blank, T., and Goopta, A. (2012). Illumishare: sharing any surface. In Proc. CHI '12, 1919-1928.

[7] Krueger, M. W. (1991) Artificial Reality II. Addison-Wesley.

[8] Lapides, P., Sharlin, E., and Sousa, M. C. (2011). Social comics: a casual authoring game. Proc. BCS HCI '11, 259-268.

[9] Morikawa, O. and Maesako, T. (1998). HyperMirror: toward pleasant-to-use video mediated communication system. Proc. CSCW'98, 149-158.

[10] Mueller, F., Gibbs, M. R., and Vetere, F. (2009). Design influence on social play in distributed exertion games. Proc. CHI 2009, 1539-1548.

[11] Tang, J.C. and Minneman, S. (1991). VideoWhiteboard: video shadows to support remote collaboration. Proc. CHI '91, 315-322.

[12] Tang, J.C. and Minneman, S.L. (1991). Videodraw: a video interface for collaborative drawing. ACM Trans. Inf. Syst. 9(2), 170–184.

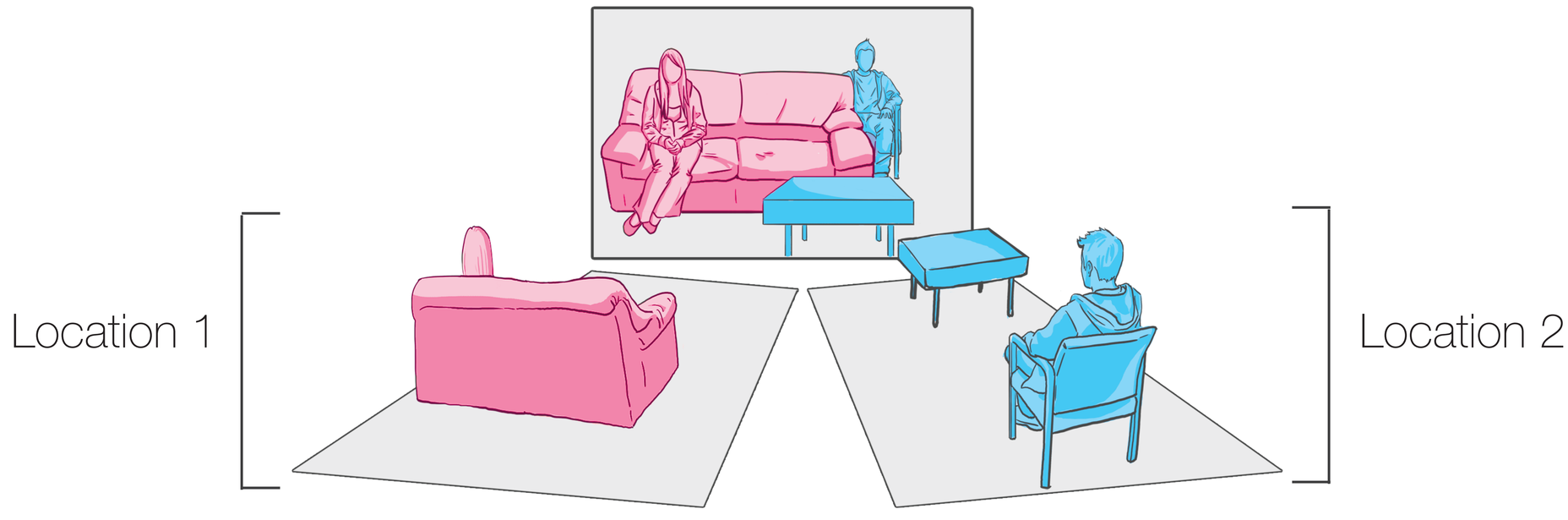
[13] Xiao, X. and Ishii, H. (2011). MirrorFugue: communicating hand gesture in remote piano collaboration. Proc. TEI '11, 13–20.

[14] Yarosh, S., Inkpen, K.M., and Brush, A.J. (2010). Video playdate: toward free play across distance. Proc. CHI '10, 251–1260.

OneSpace

Shared, Depth-Corrected Video Interaction

David Ledo, Bon Adriel Aseniero, Saul Greenberg, Sebastian Boring, Anthony Tang



OneSpace is a video conferencing system that creates a depth mirror. It therefore integrates multiple locations into a single view.



OneSpace allows for an exploration of the three dimensional virtual space, and preserves spatial relationships between people: horizontal, vertical, and depth.

Our system was implemented using thin clients that sent video feeds from the different locations. Our server would integrate these images together and show them to the client. The result was that people found themselves reflected amongst a virtually augmented space.



Environment Effects

One can modify the scene by using the different sites, static images, depth-recorded images and videos.

This allows for an ambiguation of the space and encourages exploration.



Shadows & Traces

People can be turned into silhouettes for shadow puppetry.

We can also show traces of movements



Playful Interactions

In our preliminary explorations, we found that people used the system as a way of performing virtual, physical play and staging interactions; creating an engaging and entertaining experience.

iLab
UNIVERSITY OF CALGARY FACULTY OF SCIENCE
Alberta Innovates Technology Futures
Graphisme, animation et nouveaux médias
grand
Graphics, Animation and New Media