# ONESPACE: Bringing Depth to Remote Communications

**David Ledo[1], Bon Adriel Aseniero[1], Saul Greenberg[1], Sebastian Boring[1,2] and Anthony Tang[1]**
{dledomai, b.aseniero, saul.greenberg, tonyt}@ucalgary.ca, sebastian.boring@diku.dk
[1]Interactions Lab
University of Calgary
2500 University Dr NW
Calgary, Alberta, T2N 1N4
[2]Department of Computer Science
University of Copenhagen
Njalsgade 128, Bldg. 24
2300 Copenhagen S

## 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 in the form of a shared stage.

### 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 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.

Although the original intent was to have a focus on business, the low cost of video conferencing systems led to a much broader audience, namely the domestic setting. However, the interactions in the domestic setting have typically been limited to people approaching a computer (e.g. desktop or laptop) and launching a video conferencing system. This provides people with a deep form of communication: one is able to perform a face to face conversation, and manipulate the camera to show the surroundings of the place and as a way to share experiences. With the increasing demand for long distance communications, some companies have tried to push video conferencing to a larger space: the living room. This allows people to see more information in the portal, from seeing just the head to fully being able to see people's bodies. This larger space provides individuals with a large area where people can move and communicate, yet unfortunately the video portal metaphor is unable to exploit this use of the space.

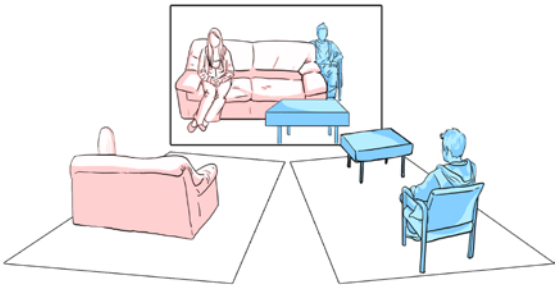

**Figure 1. OneSpace integrates two remote spaces (bottom right and left) into a single space (top) by presenting a virtual depth mirror of both spaces.**

In order for people to go beyond typical conversation and move towards shared experiences, there is a need for them to be able to engage in activities, to interact in a similar manner to being co-present. Consequently, in order to address remote interaction beyond talking heads, we propose the use of a "shared stage" interaction. More specifically, we explore the use of a depth-mirror metaphor where people: (1) share the same virtual space, (2) can make use of full-body interaction and (3) preserve spatial relationships to further encourage the use of the space. We illustrate this notion of shared-stage in our system, OneSpace (Figure 1).

## 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

Ledo, D., Aseniero, B.A., Greenberg, S., Boring, S. and Tang, A. (2013)
OneSpace: Shared Depth-Corrected Video Interaction. In Proc. ACM Conference on Human Factors in Computing Systems: CHI Workshop on The Future of Personal Video Communication: Moving Beyond Talking Heads to Shared Experiences. (Paris, France), 4 pages, April 27. Workshop held at the ACM CHI Conference, and organized by E. Oduor, C. Neustaeder, G. Venolia and T. Judge.

shared artifacts (or drawings). The metaphor being implied here is of a "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. More current examples, such as People in Books show how the shared stage can encourage immersive environments for co-located storytelling [5].

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! Hence, these apparent spatial relationships meaningfully affect how people interact with one another.

Although 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 ignore the depth relationships (Z-ordering) and present limitations for the use of the space. VideoPlace employed silhouettes, while HyperMirror used chroma-key effects, effectively always placing one space atop another. We realize a shared stage model, and build 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.

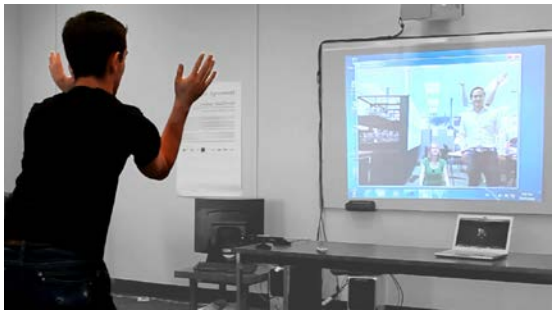

**Figure 2. OneSpace in action: physical movement is used as means of interaction.**

## ONESPACE

OneSpace integrates remote spaces through a shared depth-mirror metaphor, and acts as a technology that more fully enables the concept of the shared stage. Having depth integrated allows for respecting the location, distance and orientation between people and objects in the shared space: things and people who are closer to the mirror appear in front of those who are further away. As shown in Figure 2, people are able to interact through body movement and motion in the space and also by manipulating the physical objects in the space. OneSpace can fuse any number of real locations into a single virtual space (we have tested it with up to four environments).

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. These effects are meant as a way to create a disambiguation of the shared space and to encourage people's expressiveness:

*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

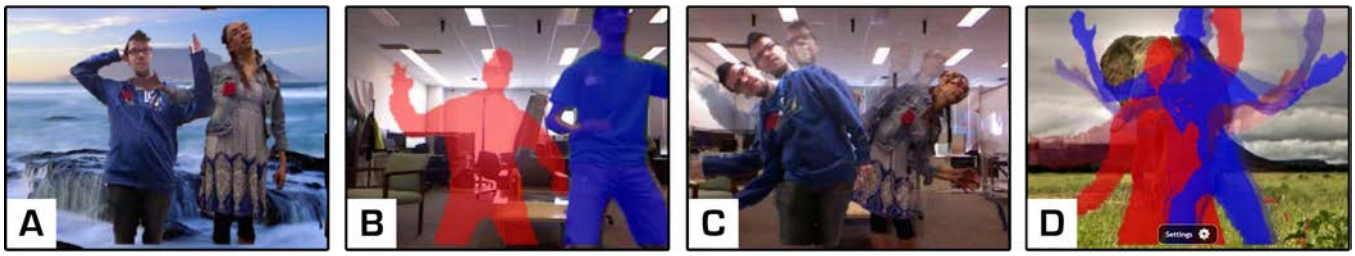


**Figure 3. Some of the effects applied in OneSpace: (a) shows a static background, (b) shows the shadow effect, (c) shows traces of movement and (d) shows a mixture of the three effects.**

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. Each site had its own whiteboard-sized display and Kinect camera. 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.

*Virtual physical and visual play.* While we expected that people would still use the system for conversation, we 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 the scene with video embodiments of people from the remote site on the shared "stage." Speech only occurred to coordinate 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 uncomfortable 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.

*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.

## IMPLICATIONS OF THE SHARED STAGE

We believe that the shared stage model brings forth a set of new interactions that were not possible with the current video portal communication model.

### Engaging Interaction

One inherently different aspect of the shared stage model is that it integrates multiple spaces in which people cannot directly identify the physical location other individuals are in. That is, because all the spatial relationships are preserved and the bodies are projected into a virtual space, one is not able to tell apart how many source clients are in the current conversation. This contrasts to having multiple windows in multi-client video portal communication and only being able to pay attention to one at the time.

Another interesting aspect resulting from our observations is how the disappearing of the sense of interpersonal touch could potentially encourage bonding between individuals who are not very close. Potentially, the act of physically getting closer virtually in the context of OneSpace might transition into these types of interactions in the real world.

### Focus on activities

Often, video mediated communication implies that people are either discussing matters of the workplace, or they have a strong connection with the other people they communicate with (family or close friends). Conversely, we believe that the use of technologies similar to OneSpace leads to

video conferencing becoming a more informal activity, in which people engage with the goal of performing a physical type of interaction. Going beyond our current conversation-based approach, we believe that the shared stage opens the possibility for many other applications:

*Physical activity*. The shared stage affords for groups of people performing activities together, such as dance, yoga or other forms of exercising. These activities can be performed in small, casual groups, or potentially as a large class, where an instructor guides the tasks. The notion of depth allows for better mechanisms to correct improper postures, as one can more accurately perform gestures, or through overlaying the body parts that need correction (as a result of the ghost-like effect caused by depth preservation).

*Games*. The shared stage opens up opportunities for games and entertainment. The aspect of mixed reality enables participants to be immersed in a changing world: the background, virtual objects, and virtual characters can become integrated and exploit the use of the three dimensional space. We can also imagine large physical objects (such as a baseball bat) to be integrated into the scene.

## 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. [4,11,12]). In particular, the "stage" of interaction is not only shared, but 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])—rather than being controlled by an unseen force (as would be with chroma-key approaches), people control these features, and use it in their interactions with one another. We see these types of playful interactions as an opportunity for further exploration. To support this kind of exploration, we consider other improvements, such as increasing the frame rate, or providing virtual toys that for play. Further, this kind of video space can provide a means to support physiotherapy, where the depth cues can aid teaching movements and poses.

## AUTHOR BIOGRAPHIES

**David Ledo** is a first year M.Sc. student in Human Computer Interaction at the University of Calgary under the supervision of Saul Greenberg. His research interests lie in ubiquitous computing. His goal is to create technologies that blend the physical and virtual worlds.

**Bon Adriel Aseniero** is also a first year M.Sc. student at the U of C's Interactions Lab, under the supervision of Sheelagh Carpendale and Anthony Tang. His research focuses on a blend between personal informatics and information visualization, and allowing people to gain insight from the representation of the data.

**Saul Greenberg** is a Full Professor at the University of Calgary. While he is a computer scientist by training, the work by Saul and his students typify the cross-discipline aspects of Human Computer Interaction, CSCW, and Ubiquitous Computing.

**Sebastian Boring** is an Assistant Professor in the HCC Group at the University of Copenhagen. His research focuses on novel interaction techniques around the display's boundaries and considering the surrounding of displays as new interaction canvas instead of only looking at the display itself.

**Anthony Tang** is an Assistant Professor at the University of Calgary. His research interests are situated in CSCW with a twist of Ubiquitous Computing. His current research investigates the integration of mobile devices in large display environments, exploration into personal informatics, and telecommunication technologies for collaborative work.

## 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. Comm. ACM.
**3.** Dourish, P. and Bly, S. (1992). Portholes: supporting awareness in a distributed work group. Proc. CHI '92, 541-547.
**4.** Ishii, H., and Kobayashi, M. (1992). ClearBoard: a seamless medium for shared drawing and conversation with eye contact. CHI '92, 525-532.
**5.** Follmer, S., Ballagas, R., Raffle, H., Spasojevic, M., and Ishii, H. (2012). People in books: using a FlashCam to become part of an interactive book for connected reading. Proc. CSCW '12, 685–694.
**6.** Junuzovic, S., Inkpen, K., Blank, T., and Goopta, A. (2012). Illumishare: sharing any surface. Proc. CHI '12, 1919-1928.
7**.** Krueger, M. W. (1991) Artificial Reality II. Reading, MA: 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 '09, 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.