synchroLight: Three-dimensional Pointing System for Remote Video Communication

Jifei Ou

MIT Media Lab

75 Amherst St.

Cambridge, MA 02139

jifei@media.mit.edu

Sheng Kai Tang

MIT Media Lab

75 Amherst St.

Cambridge, MA 02139

tonytang@media.mit.edu

Hiroshi Ishii

MIT Media Lab

75 Amherst St.

Cambridge, MA 02139

ishii@media.mit.edu

Abstract

Although the image quality and transmission speed of current remote video communication systems have vastly improved in recent years, its interactions still remain detached from the physical world. This causes frustration and lowers working efficiency, especially when both sides are referencing physical objects and space. In this paper, we propose a remote pointing system named synchroLight that allows users to point at remote physical objects with synthetic light. The system extends the interaction of the existing remote pointing systems from two-dimensional surfaces to three-dimensional space. The goal of this project is to approach a seamless experience in video communication.

Author Keywords

Video conferencing, Synthetic Light, 3D Position Tracking, Embodied Interaction

ACM Classification Keywords

H.5.m. [Information interfaces and presentation] User Interfaces

General Terms

Design, Human Factors

Introduction

Non-verbal cues play a significant role in inter-personal communication in the physical world. One of the most widely used cues is pointing. In collocated communication, the continuity of physical space enables participants a mutual understanding of spatial coordinates. One can point to any direction in the physical space to indicate the current subject that he is talking about, without verbally describing it. In standard video communication systems this simple task becomes extremely inaccurate and frustrating due to the orientation disparity of screens and cameras. (Figure 1).

We propose synchroLight, a three-dimensional pointing system, to bring the real world experience to the current video communication system. We were inspired by the phenomena that light could penetrate from one side of a glass and illuminate objects on the other side (Figure 2). In our system, the user takes a flash light and points to any position on the screen on her side. This 3-D coordination will be captured and transmitted to the remote side, where a projector simulates an illuminating spot at the corresponding spatial location. By doing so, users are able to intuitively point to physical objects in the remote location.

"Seamlessness"[1] is our goal of this work. The way people point at things in collocated interactions should have a seamless analogue in the remote space. We believe that "penetrating" light could be an effective metaphor for connecting collaborators in remote spaces.

Figure 1. Pointing to physical objects is distorted in current video communication systems.

Figure 2. Light could penetrate from one side of a glass and illuminate objects on the other side.

Related works

There have been many efforts pursuing a seamless remote collaboration experience in the field of

TeleTorchlight by Suzuki, G., Klemmer, S. [5]

Computer-Supported Cooperative Work (CSCW). ClearBoard [2] uses a calibrated projection of a remote user overlaid on a drawing glass to create an illusion that users from remote places are working on the two sides of the same glass. The HP HALO [3] system carefully configures the camera positioning, display dimension and physical spatial arrangement to re-create a perceptually continuous and identical space in a video conferencing room. The installation "rope in space" [4] properly positions real-time video displays and physical tugs in space to create an immersive experience of telematic tug-of-war game via force feedback technology.

There are also researches focusing on supporting remote assistance tasks by projecting annotation on the surface of physical objects. TeleTorchlight [5] and TeleAdvisor [6] are both portable devices that consist of a PICO projector and web camera, which allow a remote user to see and annotate local physical objects in real-time. Due to the lack of the embodiment of remote users, the non-verbal communication is not supported by these two systems. GestureMan[7] uses a tele-operated robot to move in space and point physical objects with a laser beam. However, the remote person's image, facial expression, etc. are eliminated in this system.

Design Guidelines
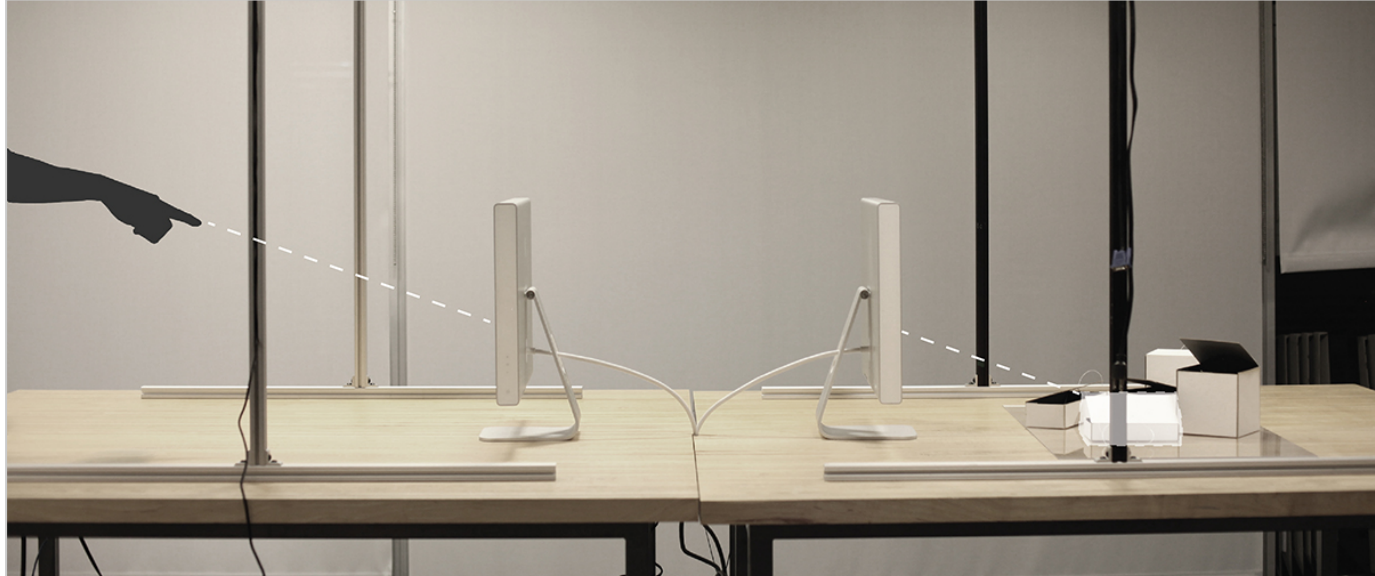We aim to design a remote communication system that

Figure 3. Illustration of interactions

allows users to seamlessly point to remote objects or space similar to the real world experience. To do so, we followed two design guidelines:

1. *Accuracy*. When people are discussing in the real world, the gesture of pointing is often approximate. Sometimes we need tools to improve the accuracy, for example, an extension stick or a light beam. A remote communication system should also deliver a high accuracy of pointing.

2. *Embodiment*. Current interfaces for remote pointing of physical objects are primarily built upon the paradigm of Graphical User Interfaces, which require users to perform actions on surfaces (tablets, touchscreens, etc.). In such systems the action of pointing is invisible to the remote users. Seeing the pointing gesture would help understanding between users on both sides.

Based on the guidelines, we propose using a visible light spot as a medium to enable users to point into a remote space. In this remote communication system, a local user turns on a typical flashlight and points to a desired spot on the display. On the other site, a projector simulates a light spot that is projected on the desired coordinates. The light metaphorically "penetrates" the local and remote displays and "reachs out" to the remote physical world. The trajectory of the simulated light is computationally calibrated to be perceptually coherent with our experience in the real world (Figure 3). The light spot provides a higher accuracy of pointing than with bare hands. Meanwhile, the embodiment of the pointing action enriches the experience of seamlessness during the communication.
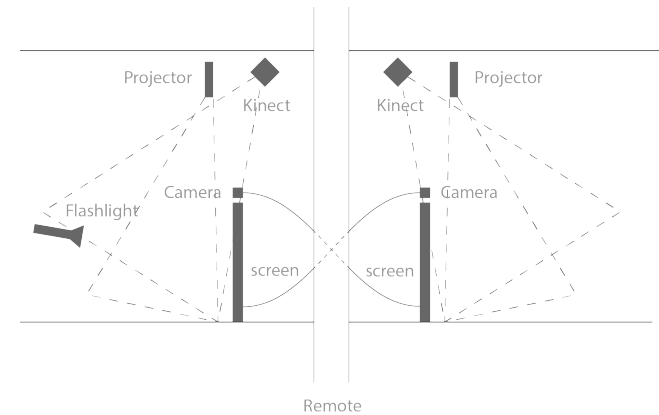

Figure 4. Ideal Hardware setup.

System implementation

Hardware The system is designed for two distributed networked locations. The current prototype consists of a Microsoft Kinect on one side and a projector on the other. The Kinect constantly tracks the three-dimensional position of the pointing hand and the vector of pointing. The projector simulates the penetrating light on the other side of the communication. Both sides have a computer display and a web camera for real-time video streaming. The current setup only allows one-way pointing. It is possible to implement the system that features two-way communication base on this setup (Figure 4).

The projector and Kinect are placed above the displays. In order to obtain wider physical range of detection and interaction, the heights of the Kinect and projector are 2.3 meters, while the displays' are 0.7 meter, relative to the floor. The current hardware setup only serves for proof-the-concept. As we are testing other possible

hardware setups, an instruction of real scenario setup will be provided in the future work.
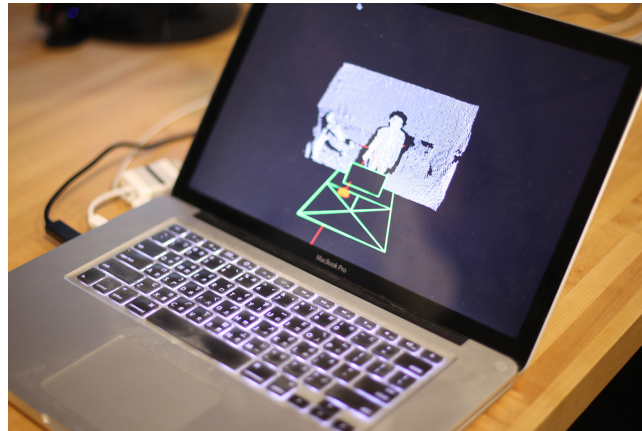

Figure 5. Calculation of intersection point.

Software Processing and Kinect library are used for the hand position tracking and light simulation. The procedure of tracking hand and generating light is as follows:

1. Use the Kinect to capture the body
2. Based on the body coordinates and orientation, create a virtual table
3. Capture wrist and elbow coordinates and generate a vector line based on them.
4. Calculate the intersection point between the vector line and the virtual desk (Figure 5).
5. Map the virtual desk to the physical remote one.
6. Transfer the coordinates of the intersection point to the remote site and project a light spot on it.

Discussion& Future Works

Using the "penetrating" light as a medium to point to remote objects has many advantages in video communication. It not only overcomes the inaccuracy of pointing in a face-to-face communication scenario, but also preserves our communication experience in the real world. In the synchroLight system, the pixels on two computer screens are the representations of the remote world, as well as a metaphor of a window through which light can freely travel.

In a preliminary test we asked 3 groups (each with 2 person) to use our system to point to remote objects. The result shows that synchroLight provides them a vivid experience of real world communication. The next step of our work is to run a user test on how much the accuracy of pointing can be improved with our system, compared to traditional video communication setup. Also we would like to compare the difference in experience between pointing with a flashlight and with bare hand.

Several technical improvements need to be done in the future:

1. Current synchroLight system supports only one-way pointing. A two-way pointing setup is desired to complete the system.
2. Re-calibrate the position of local camera and screen to archive more accuracy of pointing.
3. The mapping procedure from the virtual desk to the physical one need to be improved as well so that the system can automatically fit to any table in physical space.

4. We are looking forward using the LEAP[8] instead of Kinect for our system to reduce the complexity of system hardware setup.

We would like also to expand the metaphor of "penetrating" light spot to environmental light. The prototype for this idea is the distributed networkd workbench, whose environmental lights are synchronized. It can serve as an augmentation for remote urban planning[8] or storytelling. In this application, local user can create and move an invisible light source, which is essentially simulated by the projectors, with hand gestures. The light source illuminates both local and remote workbench and cast shadows of objects on the workbench that give users a perceptual illusion that the two workbenches were physically connected together. When the light source changes its position, the simulated shadow from both sides moves to corresponded positions. User can also pass the virtual light source to each other remotely by the hand gestures (Figure 6).
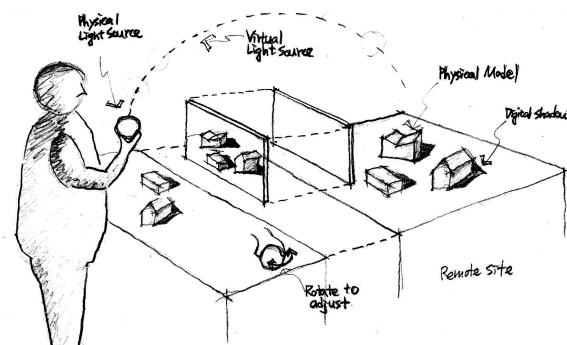

Figure 6. a illustration of remote networked workbenches with synchronized illumination.

References
[1] Ishii, H., Kobayashi, M., Arita, K., Iterative Design of Seamless Collaboration Media. *Communications in ACM, Vol. 37, Issue 8,* ACM Press (1994), 83-97.

[2] Ishii, H., Kobayashi, M., ClearBoard: a seamless medium for shared drawing and conversation with eye contact. In *Proc. SIGCHI Conference on Human Factors in Computing Systems,* ACM Press (1992), 525-532.

[3] Gorzynsky, M., Derocher, M., and Slayden M.A. The Halo B2B studio. In *Media Space 20+Years of Mediated Life*, S. Harrison Ed., Springer (2009).

[4] Stocker, Gerfried et al., *Ars Electronica 2009: Human Nature* Hatje Cantz Verlag(2009).

[5] Suzuki, G., Klemmer, S., TeleTorchlight: remote pointing and annotation using a mobile camera projector. In Proc. *14th international conference on Human-computer interaction with mobile devices and services companion*, ACM Press (2012), 35-40.

[6] Gurevich, P., Lanir, J., Cohen, B. Stone, R., TeleAdvisor: a versatile augmented reality tool for remote assistance. In Proc. *SIGCHI Conference on Human Factors in Computing Systems*, ACM Press (2012), 619-622.

[7] Kuzuoka, H., Oyama, S., Yamazaki, K., Suzuki K., Mitsuishi, M., GestureMan: a mobile robot that embodies a remote instructor's actions. In Proc. *ACM conference on Computer supported cooperative work*, ACM Press (2000), 155-162.

[8] https://leapmotion.com/

Underkoffler, J., Ishii, H., Urp: a luminous-tangible workbench for urban planning and design, In Proc. SIGCHI conference on Human Factors in Computing Systems, ACM Press (1999), 386-39.