Remote Collaboration with Spatial AR Support

Nobuchika Sakata, Yuuki Takano, and Shogo Nishida

Osaka University, Machikaneyama 1-3, Toyonaka City, Osaka, Japan
{sakata,takano,nishida}@nishilab.sys.es.osaka-u.ac.jp

Abstract. Typical view sharing system has same camera alignment that camera take images from back of remote instructor. We change this alignment to camera take images from front of remote instructor for preventing occlusions caused by a body of remote instructor self. Also as visual feedbacks, a mirror image of remote instructor is indicated in display of remote instructor side. Eventually remote instructor can confirm own instruction in the display. Therefore due to displaying the mirror image of remote instructor and changing camera alignment, we proposed and implement a novel remote collaboration system which prevents occlusion problems caused by instructor body self when he/she sends clear instructions by whole body gesture and allows instructor to use direct manipulation.

Keywords: Remote collaboration, Occlusion, Augmented Reality, View sharing system, Spatial AR.

1 Introduction

Work conducted by a local worker under the instructions of a remote instructor is called remote collaboration [1-3]. Using a telecommunication terminal, the remote instructor and the local worker transmit and receive sounds and videos to accomplish their work since they cannot share voices and views directly. On the other hand, a worker and an instructor sometimes communicate regarding objects and places in real work spaces in local collaborative works. To conduct such communication smoothly, a support system sends the remote instructions including the place of the local worker.

Especially, some studies focus on the situation in which a remote instructor provides an instruction to a local worker with real objects, for example, repairing machinery. In these studies, a tabletop display is adopted to capture the gesture of the instructor and a projector is adopted to indicate the gesture image to the real-world directly [9-12]. With these devices, it becomes easy that a local worker realizes an instruction intuitively with watching the projected image of instruction gesture on the work environment. This study focuses on remote collaboration in which a local worker works with real objects using a remote instructor. The goal of this study is to achieve an interaction that allows a remote instructor to provide a local worker with clear and accurate instructions by means of various gestures. A view sharing system between remote instructor and local worker are often used in this type of remote collaboration. In particular, we focus to occlusion problems when making clear

M. Kurosu (Ed.): Human-Computer Interaction, Part II, HCII 2014, LNCS 8511, pp. 148–157, 2014.

instructions by gestures. To solve this problem, we apply spatial augmented reality technique to view sharing system for remote collaboration..

2 Related Work

Some researches study the support of the instruction to the local worker by the remote instructor as a remote collaboration. Some of these research focus on the remote collaboration with real-world objects. The teleoperated laser pointer is adopted in some research as a pointing tool for remote collaboration [4-6]. Cterm [4] and GestureLaser [5] are device placed in a work space, and WACL [6] is a wearable device. Each of these is compact size and consists of a camera, a microphone, a speaker and a laser pointer which can be controlled remotely. The instructor can pan and tilt the laser pointer on the camera to point at real-world objects. GestureMan [7] is a system equipped with not only a teleoperated laser pointer but also a robot head and a robot arm. The robot head and the robot arm trace the motion of the remote instructor.

Kondo [8] develops view sharing system between an instructor and a worker for remote collaboration. This system is constructed from the video-see-through Head Mounted Displays (HMD) and motion trackers. The system allows two users in remote places to share their first-person views each other.

To achieve the instruction considering embodiment in the remote collaboration, some researches display the image or the shadow of the instructor on the work environment [9-12]. These systems allow to transmit the embodiment and awareness to the remote worker by sharing their arms and gestures each other on the displayed image. These research show the remote communication becomes smooth by considering embodiment and transmitting the awareness information or gestures. Therefore, the instruction via work field images including target object is effective for the remote collaboration with real-world objects. Moreover, considering embodiment and transmitting gesture or awareness information is important in the instruction with real-world objects. However, above systems focus on the system placed on the work environment. There has been little researches which proposes the instructor system, the remote interaction for the instructor and deploying spatial AR techniques.

3 View Sharing System Using Instructor Mirror Image

In typical view sharing system for remote collaboration, as conveying remote instruction to local worker, researchers studied indicating only arm image of instructor and line drawing [14-15], whole body or upper of instructor [13][16] and some instructions in VR space rebuilt for remote instructor. Especially, we focus to Kuzuoka works [17] that upper body image can help to understand instructor's gestures and measure intelligibility of worker and instructor. Along to the principle, we use images of instructor's upper body to support nonverbal communication.

In typical view sharing system for remote collaboration, HMD and table top display are used for displaying situation of remote instructor and local worker sites each other. Remote instructor and local worker cannot take different field of view because

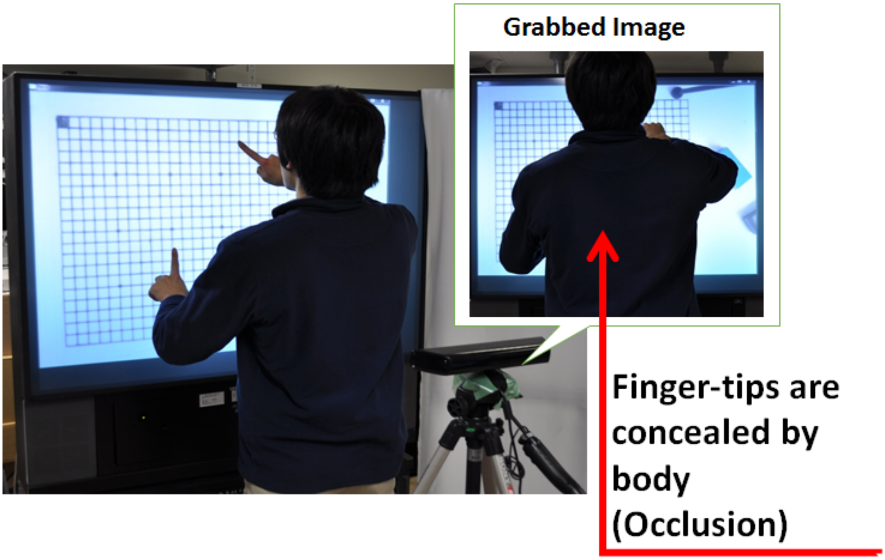


Fig. 1. Occlusion problems in typical view sharing system

view of remote instructor and local worker are perfectly matched in remote collaboration systems such as both remote instructor and local worker wears HMD. Finally, remote instructor cannot observe worker sites freely, and then performance of remote collaboration is decreased in tasks such as searching and picking. Using table top display in such remote collaboration are often used to compensate the previous problem that remote instructor cannot observe working site. In such remote collaboration system, instructions for real objects are conducted by gestures of only finger and arm. However, gesture of whole body, face and other body parts cannot be used in such kind of system due to deployment of the display placed horizontally. In our proposed system we use old fusion wall type LCD panel as output device for instructor to use gestures of whole body.

Typical view sharing system for remote collaboration has almost same camera alignment that camera take images from back of remote instructor as shown in fig 1. We change this alignment to take camera images from front of remote instructor for preventing occlusions caused by a body of remote instructor self as shown in bottom of fig 2. Also as visual feedbacks, a mirror image of remote instructor is indicated in display of remote instructor side as shown in fig 3. Eventually remote instructor can confirm own instruction in the display with this Spatial AR technique. Therefore due to displaying the mirror image of remote instructor and changing camera alignment, we proposed and implement a novel remote collaboration system which prevents occlusion problems caused by instructor body self when they sends clear instructions by whole body gesture and allows instructor to use direct manipulation. We suppose to unveil relationship among tasks efficiently, whole body gesture and even face expressions in remote collaboration regarding objects and places in real work spaces.

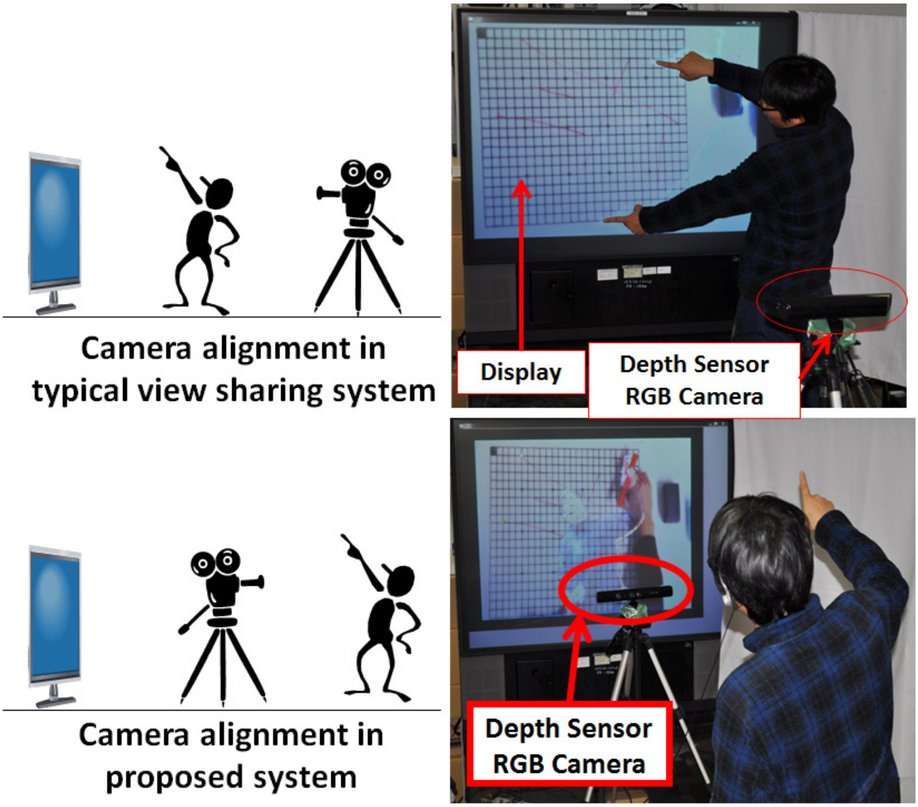


Fig. 2. Camera alignment in typical view sharing system and proposed system

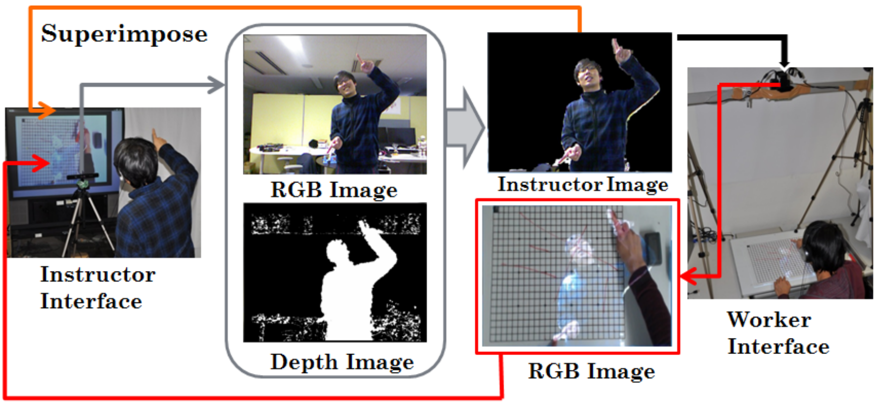


Fig. 3. System diagram

3.1 System Overview

Our proposed system is composed of instructor and worker interfaces. Fig 4 shows appearance of instructor interface. It composed of Flat panel display (Mitsubishi, 55P-FD100) to indicate video image of worker side, and Microsoft Kinect to capture instructor's body movement. The Kinect is deployed between the Flat panel display and standing position of instructor. Kinect can capture depth and RGB image of instructor at the same time. Image of only instructor can be extracted from RGB image according to the depth image. After that extracted instructor image is sent to worker interface as shown in Fig 3. Also instructor cannot touch the display directly with deploying the Kinect between the display and standing position of instructor. Instructor cannot recognize where instructor is pointing to exact. It causes a lack of direct touch and decreasing usability. To compensate those issues, transparent image of instructor body is superimposed to image on the display as shown in upper right of Fig 4.

Fig 5 shows worker interface. The worker interface is composed of a projector (Mitsubishi, LVP-DX95) and RGB camera (Logicool, HD Pro Webcam C920). The RGB camera capture circumstance of worker side. Image of instructor upper body is overlaid on working place with the projector as shown in Fig 5. This interface is not special and this style is typical Procams (Projector and Camera systems). Because of that we do not focus to improvement of worker interface in this paper.

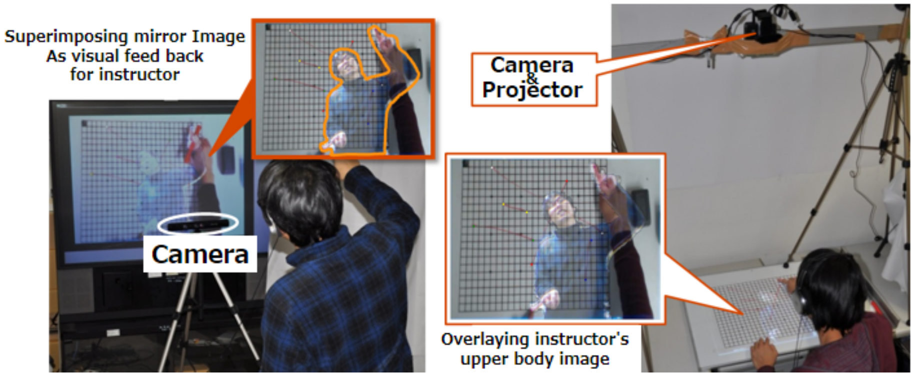


Fig. 4. Appearance of instructor interface (Left) and worker interface (Right)

Fig 3 shows process of instructions. Worker interface capture working place including worker's hand and objective parts of this task. The captured image stream is sent to instructor interface via TCP/IP network. At this time, we apply keystone effect to configure and compensate distortion. And then the corrected image indicate on the large display located in front of the instructor. Observing the image stream of worker side, the instructor make instructions by finger pointing and whole body gesture.

Also, image of instructor upper body can be extracted according to depth image of Kinect. Those image is sent to worker side, and then the instructor upper body is overlaying on work place. Finally, instructor's finger pointing and whole body gesture are duplicated in worker side. Referring duplicated nonverbal channel, the worker conduct tasks. Simultaneously, the image is superimposed to the display in instructor side as visual feedback for the instructor. Also aural communication can be used each other as full duplex via Skype.

3.2 User Study and Result

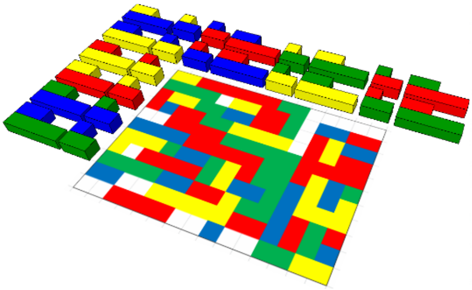
We conduct 2 type user study. After that we called them Experiment 1 and 2. Task 1 can be accomplished by pointing out only 1 place. Also we set that occlusion problems do not occur so much as the experiment condition of Experiment 1. Aim of Experiment 1 is to confirm that the proposed system can mark almost same performance as well as typical view sharing interface as shown in upper of fig 2.

In Experiment 2, subject should be pointing out 2 places to accomplish. We set Experiment 2 condition that occlusion problems occur so much. Aim of Experiment 2 is to confirm that the proposed system can perform much more than typical view sharing interface.

In Experiment 1, we set a task that subjects place some blocks on gridded paper as shown in left of Fig 5. The blocks are painted by random color pattern not to distinguish at once. Also the gridded paper is painted by random color pattern. It means that we let instructor use finger pointing rather than aural instruction. In experiment 2, we set a task that subject draw a line from two point indicated by remote instructor on a gridded board (Right of Fig 5). Some base points are painted as large black circle on the gridded board. The base points avoid that remote instructor spend time for searching two points which convey to a local worker.

12 subjects (ages 21-25, 11 male and 1 female) conduct Experiment 1 and 2 to consider order effects. As a result, typical view sharing is faster than the proposed interface in Experiment 1. Significant difference is found in task completion time of Experiment 1. In Experiment 2, the proposed interface is faster than the typical view shared interface. Significant difference is found in task completion time of Experiment 2.

Also we conduct questionnaire after Experiment 1 and 2. In experiment 1, we cannot find significant difference among almost all evaluation items excepting "Which condition do you transmit the instruction easier?". In terms of "Which condition do you transmit the instruction easier?", the proposed interface is more easier than typical interface. In Experiment 2, the proposed interface obtain good impressions and obtain significant differences in "Which condition do you transmit the instruction easier?", "Which condition do you conduct the instruction precisely?" and "Which condition do you feel burdens during instructions?". However, we cannot find significant difference in "Which condition do you conduct the instructions faster?"



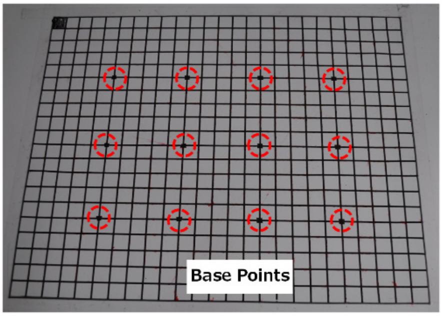


Fig. 5. Gridded board (Left). Blocks and gridded paper (Right).

3.3 Discussion

In Experiment 1 which needs one pointing instructions, the typical existed interface mark shorter completion time than the proposed interface. As quantities evaluation, we cannot find significant difference among almost all evaluation items excepting "Which condition do you transmit the instruction easier?".

In Experiment 2 which needs two pointing instructions, the proposed interface marks shorter completion time than typical interface. As quantities evaluation, the proposed interface provides better impressions than the typical interface. Especially, quite large difference is obtained in evaluation item of "Do you feel burden when instructions?" because instructors should take unnaturally posture during instruction in the typical interface. It can say the proposed interface using mirror image can compensate those occlusion problems. Also the proposed interface can provide same impression when conducting one pointing instruction.

4 Sharing Face Expression among Worker and Instructor Using Mirror Image

In previous chapter, we proposed, implemented and evaluated view sharing system using instructor mirror image as application of remote collaboration with installing spatial AR techniques. As a result, we can compensate occlusions problems during two pointing instructions. Also we propose a method of sharing face expression among worker and multiple instructors using mirror image as other application of remote collaboration with installing spatial AR technique. We assume that application can mark good performance in remote collaboration between two or three instructors and one field worker.

As shown in fig 6, two or three instructor are considering how to instruct. Then, local worker can watch the conversation among instructors with observing face expression, gestures and body. The local worker might obtain much nonverbal communication comparing to a method of left and middle of Fig 6. Finally, understanding the conversation deeply, the local worker can conduct tasks smoothly.

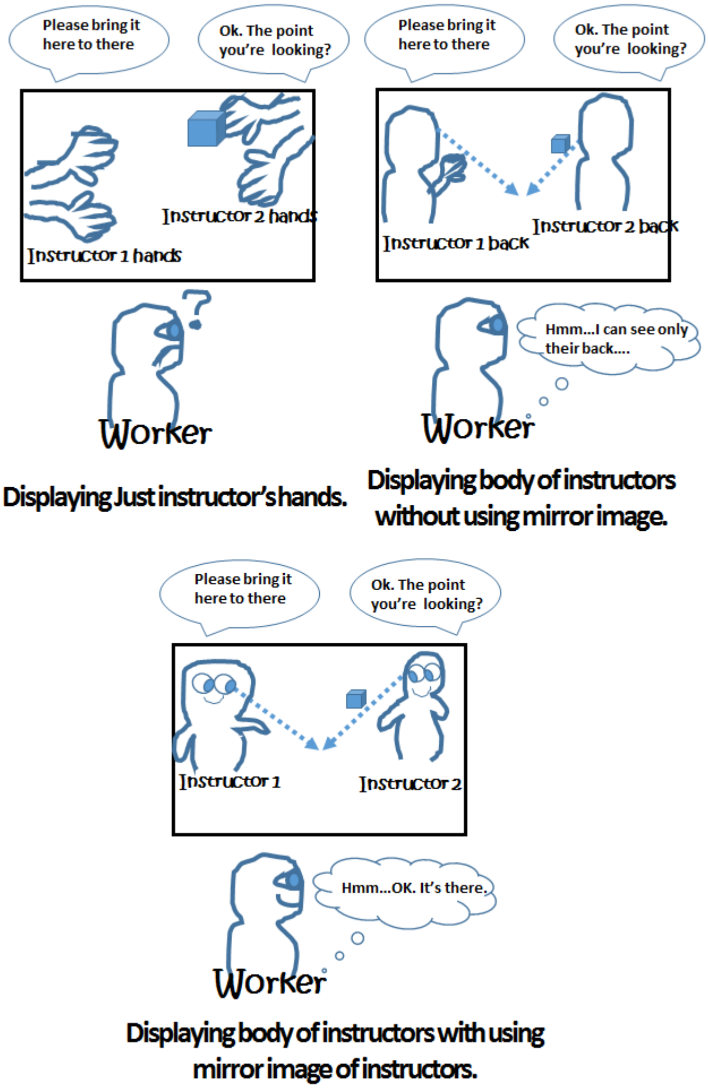


Fig. 6. Advantage of sharing face expression with using mirror image of instructors

5 Conclusion

We propose, implement and evaluate new view sharing system for remote collaboration. We change this alignment to camera take images from front of remote instructor for preventing occlusions caused by a body of remote instructor self. Also as visual feedbacks, a mirror image of remote instructor is indicated in display of remote instructor side. Eventually remote instructor can confirm own instruction in the display. Therefore due to displaying the mirror image of remote instructor and changing camera alignment, we proposed and implement a novel remote collaboration system which prevents occlusion problems caused by instructor body self when he/she sends clear instructions by whole body gesture and allows instructor to use direct manipulation. Also we propose a method of sharing face expression among worker and multiple instructors using mirror image as other application of remote collaboration with installing spatial AR technique.

References

1. Kuzuoka, H.: Spatial workpace collaboration: Asharedview video support system for remote collaboration capavility. In: Proc. CHI 1992, pp. 533–540 (1992)
2. Fussell, S.R., Setlock, L.D., Kraut, R.E.: Effects of head-mounted and scene-oriented video systems on remote collaboration on physical tasks. In: Proc. CHI 2003, pp. 513–520 (2003)
3. Kraut, R.E., Miller, M.D., Siegal, J.: Collaboration in performance of physical tasks: Effects on outcomes and communication. In: Proc. CSCW 1996, pp. 57–66 (1996)
4. Mikawa, M., Matsumoto, M.: Smooth and easy telecommunication using CTerm. In: Proceedings of IEEE SMC 1999, pp. 732–737 (1999)
5. Yamazaki, K., Yamazaki, A., Kuzuoka, H., Oyama, S., Kato, H., Suzuki, H., Miki, H.: Proceedings of the Sixth Conference on European Conference on Computer Supported Cooperative Work, pp. 239–258 (1999)
6. Sakata, N., Kurata, T., Kato, T., Kourogi, M., Kuzuoka, H.: WACL: Supporting telecommunications using wearable active camera with laser pointer. In: Proceedings of the Seventh IEEE International Symposium on Wearable Computers, pp. 53–56 (2003)
7. Kuzuoka, H., Furusawa, Y., Kobayashi, N., Yamazaki, K.: Effect on Displaying a Remote Operator's Face on a Media Robot. In: Proceedings of ICCAS 2007, pp. 758–761 (2007)
8. Kondo, D., Kurosaki, K., Iizuka, H., Ando, H., Maeda, T.: View sharing system for motion transmission. In: Proceedings of the 2nd Augmented Human International Conference (March 2011)
9. Kirk, D., Crabtree, A., Rodden, T.: Ways of the hands. In: Proc. 9th European Conference on Computer-Supported Cooperative Work, France, pp. 1–21 (September 2005)
10. Tang, A., Pahud, M., Inkpen, K., Benko, H., Tang, J.C., Buxton, B.: Three's company: understanding communication channels in three-way distributed collaboration. In: Proc. ACM Conference on Computer Supported Cooperative Work, Savannah, USA, pp. 271–280 (February 2010)
11. Yamashita, N., Kuzuoka, H., Hirata, K., Aoyagi, S., Shirai, Y.: Supporting fluid tabletop collaboration across distances. In: Proc. Annual Conference on Human Factors in Computing Systems, Vancouver, Canada, pp. 2827–2836 (May 2011)
12. Izadi, S., Agarwal, A., Criminisi, A., Winn, J., Blake, A., Fitzgibbon, A.: C-Slate: A Multi-Touch and Object Recognitio System for Remote Collaboration using Horizontal Surfaces. In: IEEE Workshop on Horizontal Interactive Human Computer Systems, Rhode Island, USA, pp. 3–10 (October 2007)
13. Uemura, K., Sakata, N., Nishida, S.: Improving Visibility of Gesture Image with Scaling Function for Tabletop Interface in Remote Collaboration. Transactions of the Virtual Reality Society of Japan 17(3) (2012) (in Japanese)
14. Grevich, P., Lanir, J., Cohen, B., Stone, R.: TeleAdvisor: A Versatile Augmented Reality Tool for Remote Assistance. In: CHI 2012 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 619–622 (2012)

15. Hiura, S., Tojo, K., Inokuchi, S.: 3-D Tele-direction Interface using Video Projector. In: 30th Int'l Conf. on Computer Graphics and Interactive Techniques (ACM SIGGRAPH 2003) Sketches & Applications, San Diego, California, July 27-31. ACM, New York (2003)
16. Kirk, D., Crabtree, A., Rodden, T.: Ways of the hands. In: Proc. 9th European Conference on Computer-Supported Cooperative Work, France, pp. 1–21 (September 2005)
17. Kuzuoka, H., Yamashita, J., Yamazaki, K., Yamazaki, A.: Agora: A Remote Collaboration System that Enables Mutual Monitoring. In: Proc. CHI 1999, pp. 190–191. ACM Press (1999)