

Simplifying Remote Collaboration Through Spatial Mirroring

Fabian Hennecke¹, Simon Voelker², Maximilian Schenk¹, Hauke Schaper²,
Jan Borchers², Andreas Butz¹

¹University of Munich (LMU), HCI Group
80333 Munich, Germany

²RWTH Aachen University
52056 Aachen, Germany

{fabian.hennecke, andreas.butz}@ifi.lmu.de
{voelker, borchers}@cs.rwth-aachen.de

Abstract. Even though remote collaboration through telepresence is supported by a variety of devices and display environments, it still has some inherent problems. One of these problems is the definition of a unified spatial reference system for the shared workspace in combination with an immersive representation of the collaborator. To mitigate this problem we propose a technique we call spatial mirroring. It is based on a virtual collaboration environment using two curved displays and aims to eliminate possible communication errors due to left/right misunderstandings. We explain the working principle and ideas behind spatial mirroring, and present two consecutive user studies in which we were able to verify its benefits.

Keywords: Remote collaboration; Telepresence; Orientation problem; Spatial reference; Mirroring.

1 Introduction

Over the last decade remote collaboration and telepresence have been widely researched and became available to many users (e.g. Skype¹). A first vision of remote collaboration [8] was sketched in the Starfire video prototype [14]. Since then many projects integrated the remote collaboration system into the actual working environment of the user and created the illusion of being there [7, 11]. Recently the MirageTable [1] combined two differently-oriented display areas with a stereoscopic projection of a remote collaborator allowing three-dimensional collaboration. Beside this, additional horizontal displays have also been used to visualize a remote collaborator's interactions based on video overlays on an entirely shared workspace [6]. Even a video representation of remote collaborators on additional vertical displays all around a horizontal display was proposed [4] and its effect on collaboration was evaluated [14]. However, these approaches suffer from a problem, which can even arise in

¹ <http://www.skype.com>

co-located collaboration settings: it is a priori unclear which of the different spatial reference systems of each collaborator someone refers to [16]. Collaborators can share the same point of view, which minimizes reference problems, but this also destroys the feeling of ‘being there’ in a remote collaboration scenario. Contrary to this the collaborators can be positioned on different sides of the shared workspace [10], which offers a good immersion but suffers from different spatial reference systems. While this problem can be mitigated with gazes and gestures in a co-located setting it is hard to properly transfer these additional collaboration information into a remote system. As speakers tend to use their own spatial reference system more often during collaborative tasks [12] this can become a problem during remote collaboration.



Fig. 1. Different views of the same table: (A) User 1: normal view, (B) User 2: normal view, (C) User 2: horizontally mirrored, (D) User 2: horizontally mirrored and content on the table additionally horizontally mirrored (spatially mirrored). Using a spatially mirrored visualization reduces misunderstandings about left and right and reduces selection times and error rates.

We propose a technique called spatial mirroring to overcome such spatial reference problems in remote collaboration environments. Spatial mirroring is basically a seamless combination of the two metaphors ‘over a table’ and ‘through a glass window’ described by Ishii et al. [5] and is based on the collaboration concept of the PerspectiveTable [3]. Based on two curved displays [17, 18] we utilize a double mirroring effect to align the spatial reference systems for both users in terms of left and right. Instead of ‘my left’ and ‘your left’ spatial mirroring creates ‘our left’ across the entire working area. In two consecutive user studies we were able to verify the benefits of spatial mirroring compared to a non-mirrored visualization.

2 Spatial Mirroring

A seamless virtual remote collaboration environment including differently oriented working areas allows for modifying the spatial reference system of a user without making this explicit or obvious. This way both users can refer to the same ‘left’ and ‘right’, while objects on the table still remain upside down for the opposite collaborator. (see Fig. 1.a and d). This avoids confusion and explanations about whose ‘left’ or ‘right’ they reference to.

Our setup for spatial mirroring is therefore based on a remote collaboration environment similar to the PerspectiveTable [3]. It is also based on two curved displays located in different cities. Using this setup spatial mirroring is realized with a mirroring along the central vertical axis of the entire video image using a simple pixel shader on one of the displays (see Fig. 1.b to c). While this is sufficient for content located vertically between both collaborators, an additional mirroring step is needed

for content in the shared horizontal workspace to eliminate the problem of mirror-inverted appearance caused by the first mirroring (see Fig. 1.c). We are aware that this isn't the only possible combination of mirroring effects to achieve a spatial mirroring. But as we already used shaders to correct the image distortion caused by the curved displays we could implement the first mirroring step without much extra cost.

We opted for a virtual avatar like the PerspectiveTable [3] instead of a real video image as this allows us to make spatial mirroring go unnoticed. A mirrored video image of a known person can be easily identified. In case of a virtual avatar only the inverted handedness of the remote collaborator might reveal the mirroring.

3 Evaluation of Spatial Mirroring

To evaluate spatial mirroring we conducted two consecutive studies. The first study explored the basic effects on users' performance during simple selection tasks while the second study investigated the effects on actual collaboration. Participants of this second study solved a puzzle with a remote collaborator. The results of the first study show that users performed significantly better in terms of speed and error rate using spatial mirroring. The video analysis of the second study revealed that users also adapt to the mirrored reference system without problems. Both studies were conducted with curved displays of similar specifications to BendDesk and Curve [17, 18].

3.1 Study 1.

In this study a virtual avatar at the opposite end of the remote collaboration table asked the participants to select one out of ten virtual playing cards. The cards' position was described relatively to the outer left or outer right cards using different instruction types (see Fig. 2.).

Participants and Design.

19 participants took part in the study (9 females, 10 males, aged between 19 and 59, all right-handed). We used three different instructions (*audio*, *pointing* and *audio+pointing*) to describe the cards' position and two different visualizations (*normal* and *mirrored*) of the collaboration environment as within-group factors. The combination of these factors was counterbalanced with an 8×8 Latin square resulting in 64 actual selection tasks per participant. As both visualizations are the same for the instructions audio and pointing we interpreted the results of the normal visualization separately for both visualizations.

Task and Procedure.

The participants were seated centered in front of the display. They had a training phase prior to the study where they were asked to select one out of two virtual playing cards using all instruction types of the study. There was no feedback if the selection was right or wrong during the entire study.

Instruction Type: We used three different instruction types: audio, pointing, and audio+pointing. For audio the participants were asked to select a particular card (“third card from the left”). In the pointing condition the avatar only pointed at that particular card while he combined both instructions in the audio+pointing condition (see Fig. 2.).

Visualization: Two different visualizations were used: normal and mirrored. In the normal condition the avatar referred to his left and right of the table, which could lead to a mismatch between audio and pointing in the combined condition (see Fig. 2.a). In the mirrored visualization the avatar referred to the mirrored spatial system and used left and right accordingly (see Fig. 2.b).

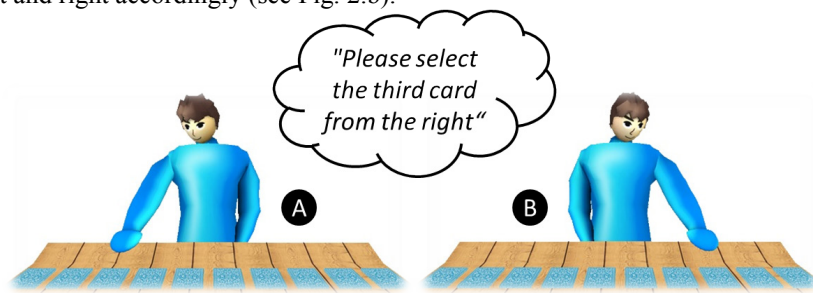


Fig. 2. Task design of our first study: A virtual avatar asked the participant to select a specific card from a row of ten cards using three types of instructions: audio, pointing, audio+pointing. The display was (A) not mirrored or (B) spatially mirrored across these three conditions.

Measures.

We measured two types of objective data: selection time and errors. Selection time was defined as the time between the mention of the location information (end of the instruction) and the actual selection of a card by the participant via touch. It was also logged if the selection was correct. In the case of audio and pointing these measures were interpreted separately for normal and mirrored based on the same input data. After each session, participants answered an online questionnaire.

Statistical Tests and Analysis.

Objective data was analyzed using a repeated measures analysis of variances (ANOVA). In cases in which the assumption of sphericity was violated, we used a Greenhouse-Geisser correction.

Results.

As selection times for audio and pointing were identical due to the study design the selection times of the visualizations only differed for the audio+pointing instruction. Interestingly we still found a significant main effect of the visualization ($F_{1,18} = 12.628, p < .05$) with mirrored (1953 ms) being faster than normal (2095 ms) across all instruction types. Besides instruction also had a significant effect on selection

times ($F_{2,17} = 19.212$, $p < .05$). Post-hoc tests revealed significant differences between audio (2275 ms) and pointing (1642 ms) and also between audio+pointing (2156 ms) and pointing. Audio+pointing was even slightly slower than audio only in the normal visualization. We believe this is caused by the inconsistency between the audio and gestural instruction for audio+pointing in the normal visualization, which might have created a higher cognitive load.

There were also significant main effects of the visualization ($F_{1,18} = 222.493$, $p < .05$) with the spatial mirror (99%) being more correct across all instructions compared to normal (43%). This difference is caused by the large number of errors due to differing audio+pointing instructions in terms of left and right in the normal visualization – only 29% of these selections were correct while 100% of them were correct in the mirrored visualization. The instruction also had a significant effect ($F_{2,36} = 18.691$, $p < .05$) on the correctness of selection. Post-hoc tests revealed that pointing (100 %) performed best compared to audio+pointing (64 %) and audio (49 %) (all $p < .05$) across both visualizations. Besides we also found a significant interaction effect between visualization and instruction ($F_{1,074,19,340} = 82.899$, $p < .05$). Looking at each instruction separately shows that the mirrored visualization achieves nearly 100% correctness for all instructions while the normal visualization suffers from nearly 0% correctness in the audio only and 29% in the audio+pointing scenario.

3.2 Study 2

The second study consisted of a puzzle solving task with a remote collaborator similar to the study of Gergle et al. [2]. Two participants had to solve a 4×4 puzzle on the shared virtual table sitting remotely face-to-face while on screen was spatially mirrored to unify the collaborative working environment.

Participants and Design.

16 participants (8 females, 8 males, aged between 21 and 31, 14 right-handed) were split into 8 teams (two participants per team). These teams were divided into two groups based on the visualizations already used in the first study: normal (no mirroring and no mentioning of the purpose of the study) and mirrored with an explicit explanation of the mirroring prior to the study. The remote collaborator was visualized with an animated virtual avatar based on Kinect tracking data for both groups.

Task and Procedure.

The participants received a short introduction and training by solving a simple 3×3 puzzle on their own using direct touch. Afterwards the remote connection and an audio channel were established and the participants started to solve a 4×4 puzzle which showed an abstract tribal face in black and white (see Fig. 3). To enforce collaboration during the study one third at each end of the virtual table was only available to the corresponding user at that side. This way the users had to ask their collaborators for a specific tile if they could not access it.

Measures.

We recorded each session at both displays in order to analyze the participants' behavior and interaction with the display. We counted the spatial descriptions of different puzzle tile positions during a session (left/right, front/back) and its according spatial reference system. We also counted pointing, description of the tile content and tile movement for each participant. Each description was also categorized into one of three categories: successful (e.g., participant received a tile after asking for it and describing its position), not-successful (e.g., collaborator moved the wrong tile), and descriptive (e.g., "I will drag this tile over here."). The participants answered an online questionnaire based on 5 point Likert scales after solving the puzzle.

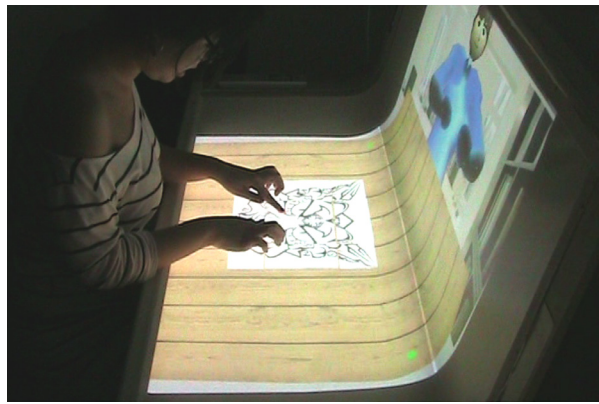


Fig. 3. A participant during the second study with a nearly complete puzzle and her remote collaborator in front of her.

Statistical Tests and Analysis.

Due to the limited number of participants and the subjective video analysis we didn't run statistical tests with the data and use frequencies instead to report our findings. They balance numeric differences in terms of left/right references, which might occur due to the explanation prior the mirrored condition compared to the normal condition. We combined the subjective ratings of the participants into 3 bins: 'Agree' ('1' & '2'), 'Neutral' ('3'), and 'Disagree' ('4' & '5').

Results.

We recorded a total of 117 pointing and movement references and descriptions during the study. Some Participants also pointed at tiles or moved them to describe other tiles next to them. This happened to similar extents in both groups and therefore doesn't constitute a difference.

Normal: Surprisingly only 20% of the participants used their own reference system for left/right instructions, of which 88% were successful. This is possibly due to the rather unfamiliar remote collaboration setting [12]. In most cases they added the re-

spective reference system (e.g., “on your left side”) to the instruction and if they didn’t their collaborator asked. Interestingly only 50% stated afterwards that they mirrored left and right in order to describe a tile’s position. In terms of front and back, 65% used their reference system, which worked quite well (67% success, only 25% reported communication problems regarding “in front” and “in the back”).

Mirrored: All participants of this group used the unified spatial reference system regarding left and right: 62% were requests (88% successful) and 38% were descriptive. We found no usage of the pre-mirrored collaborator’s reference system in this group. The self-assessment of the participants backs these results: 88% stated that they didn’t have problems to determine what was ‘left’ and ‘right’. Nevertheless, 25% reported communication problems regarding ‘left’ and ‘right’ after the study. 80% of the front/back references were made according to the users’ own point of view; they were mainly used for descriptive purposes (71%). However participants reported communication problems (25%) and problems to determine what is in the front and back (50%) after the study.

Further Findings: We noticed that subjects used an object’s relative position between other objects to describe its position during the entire study. Some participants even used the display shape as a reference (e.g., “in the curve”). Contrary to our expectations participants only used a visual description of specific tiles after having a rough idea of the entire image and didn’t describe specific visual properties of a tile. Finally, 88% agreed that a uniform point-of-view can help to avoid problems.

4 Summary and Conclusion

We presented spatial mirroring as a technique to avoid misunderstandings during remote collaboration about different interpretations of left and right. We implemented a spatial mirror based on a three-dimensional virtual collaboration environment similar to the Perspective Table [3]. We conducted two consecutive user studies to gather first insights on spatial mirroring. The first study showed that a spatially mirrored visualization not only allows faster selections but also minimizes selection errors. We assume that the faster selection times are due to a lower cognitive load during audio+pointing instructions as the instructions do not conflict during the mirrored visualization. The same applies for the selection error: The difference between the instructions for audio+pointing in the normal visualization led to wrong selections, while the mirrored visualization was completed nearly without errors. The second study involving a more realistic collaboration showed that a spatial mirror can help to ease spatial descriptions between collaborators during remote collaboration.

Based on these results we conclude that spatial mirroring can be used to ease remote collaboration by providing a unified spatial reference system. It reduces selection times due to fewer spatial reference conflicts and helps to avoid communication problems caused by misunderstandings about left and right. Nevertheless, one has to keep in mind that the position of the content within a document on the horizontal working area is also horizontally inverted to avoid mirror-inverted content. It will be interesting to investigate if and how this affects remote collaboration with text-based

content [8]. Beside this allowing to fluently change the seating position around the virtual table could help to enhance the feeling of ‘being-there’. Finally the generalization to more than two users will be very interesting as it can be achieved by modifying the avatar movements to fit the virtual seating arrangement in each case but also opens up a lot of new questions about unified spatial references for all collaborators.

References

1. Benko, H., Jota, R., and Wilson, A. MirageTable: freehand interaction on a projected augmented reality tabletop. Proc. of CHI’12, 199–208.
2. Gergle, D., Kraut, R., & Fussell, S. R. Language efficiency and visual technology: Minimizing collaborative effort with visual information. Journal of Language and Social Psychology, 23 (2004), 491-517.
3. Hennecke, F., Voelker, S., Schenk, M., Schaper, H., Borchers, J. and Butz, A. PerspectiveTable: Blending Physical and Virtual Collaborative Workspaces. Workshop 'Blended Interaction: Envisioning Future Collaborative Interactive Spaces' in conjunction with CHI’13, <http://hci.uni-konstanz.de/blendedinteraction2013/>.
4. Hirata, K., Harada, Y., Takada, T., et al. The t-Room: Toward the Future Phone. NTT Technical Review 4, 12 (2006), 26–33.
5. Ishii, H. and Kobayashi, M. ClearBoard: a seamless medium for shared drawing and conversation with eye contact. Proc. of CHI’92, 525–532.
6. Junuzovic, S., Inkpen, K., Blank, T., and Gupta, A. IllumiShare: sharing any surface. Proc. of CHI’12, 1919–1928.
7. Kauff, P. and Schreer, O. An immersive 3D video-conferencing system using shared virtual team user environments. CVE 2002, 105–112.
8. Matsushita, M., Iida, M., and Ohguro, T. Lumisight Table: A Face-to-face Collaboration Support System That Optimizes Direction of Projected Information to Each Stakeholder. Proc. of CSCW’04.
9. Minsky, M. Telepresence. OMNI.
10. Olson, M.H. and Bly, S.A. The Portland experience: a report on a distributed research group. Computer-supported Cooperative Work and Groupware 34, 2 (1991), 81–98.
11. Sadagic, A., Towles, H., Holden, L., Daniilidis, K., and Zeleznik, B. Tele-immersion Portal: Towards an Ultimate Synthesis of Computer Graphics and Computer Vision Systems. 4th Annual International Workshop on Presence, (2001).
12. Schober, M. F. Spatial perspective-taking in conversation. Cognition, 47 (1993), 1-24.
13. Schwarz, T., Hennecke, F., Lauber, F., and Reiterer, H. Perspective + Detail – a visualization technique for vertically curved displays. Proc. of AVI’12, 485–488.
14. Tang, A., Pahud, M., Inkpen, K., Benko, H., Tang, J., and Buxton, B. Three’s company: understanding communication channels in three-way distributed collaboration. Proc. CSCW 2010, 271–280.
15. Tognazzini, B. The “Starfire” Video Prototype Project : A Case History. Proc. of CHI’94, 99 – 105.
16. Tversky, B. and Hard, B.M. Embodied and disembodied cognition: Spatial perspective-taking, Cognition, 110, Issue 1 (2009), Pages 124-129.
17. Weiss, M., Voelker, S., Sutter, C., and Borchers, J. BendDesk: Dragging Across the Curve. Proc. of ITS’10, 1–10.
18. Wimmer, R., Hennecke, F., Schulz, F., Boring, S., Butz, A., and Hußmann, H. Curve: Revisiting the Digital Desk. Proc. of NordiCHI’10, 561–570.