ORIGINAL ARTICLE

An interactive whiteboard for immersive telecollaboration

Thomas Nescher · Andreas Kunz

Abstract In this paper, we present CollaBoard, a collaboration system that gives a higher feeling of presence to the local auditory and to the persons on the remote site. By overlaying the remote life-sized video image atop the shared artifacts on the common whiteboard and by keeping the whiteboard's content editable at both sites, it creates a higher involvement of the remote partners into a collaborative teamwork. All deictic gestures of the remote user are shown in the right context with the shared artifacts on the common whiteboard and thus preserve their meaning. The paper describes the hardware setup, as well as the software implementation and the performed user studies with two identical interconnected systems.

Keywords Remote collaboration · People on content · Shared workspaces · Mixed presence groupware · Digital whiteboard

1 Introduction

Working together efficiently over distance has become a key factor for successful business, in particular for companies which have expertise spread all over the world. This implies a need for means supporting distant teamwork. Thus, collaboration by using video- and data-conferencing systems becomes increasingly popular. Today, typical collaboration

T. Nescher · A. Kunz (✉)
ETH Zurich, Institute of Machine Tools and Manufacturing
CLA G9, Tannenstrasse 3, 8092 Zurich, Switzerland
e-mail: kunz@iwf.mavt.ethz.ch

T. Nescher
e-mail: nescher@iwf.mavt.ethz.ch

software does not only provide an audio- and video connection, but also a whiteboard functionality. However, working together is much more than just seeing each other and sharing a common whiteboard. In collaboration processes, the focus of attention is task-centered [5] and collaborators use deictic gestures for referring to shared artifacts.

The importance of body language such as (deictic) gestures, pose, gaze, and facial expressions of collaborators was already described earlier [8, 10, 11, 13, 20]. However, commercially available conferencing systems do not allow, e.g., a natural use of deictic gestures during combined video- and data-conferencing, since they display video and shared workspace in separate application windows, a so-called side-by-side representation (see Fig. 1).

The development of conferencing systems that enable rich use of body language in remote collaboration processes is subject to ongoing CSCW research. This paper fits in this research area and presents a conferencing system that gives distant collaborators the sensation of being virtually next to each other, and that allows collaborators using natural deictic gestures, pose, and gaze for communication by realizing a "people on content" representation.

2 Related work

An early work in this particular field of CSCW research was, e.g., ClearBoard [8]. Other systems solved the problem of a limited shared screen space by using projection technology (VideoWhiteBoard [21], ClearBoard [9], Distributed Designers' Outpost [3]) or specific software (Clear-Face [7] for TeamWorkStation, and Vis-a-Vid Transparent VideoFaceTop [19]). However, with all these systems, the camera for video conferencing or presence awareness was placed beside or virtually behind the screens, acquiring a
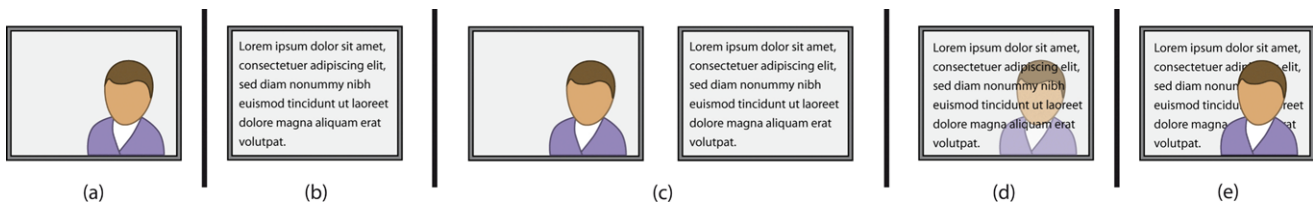
Fig. 1 Types of representation: (**a**) People only, (**b**) Content only, (**c**) People & Content, (**d**) Content on People, (**e**) People on Content

front or side view of the user. This camera alignment results in a "content on people" rather than in a "people on content" for the underlying metaphor "looking through and (collaboratively) drawing on a big glass board" [8].

A recently presented, very elaborated conferencing system that supports collaborator's natural use of gestures is VideoArms [23, 24]. The system acquires people interacting on shared workspaces (large interactive displays) by means of a camera that is on-axis with the display device. Therefore, the context of hand gestures is preserved, e.g., deictic gestures pointing out a shared artifact can be correctly interpreted by the remote collaborator. VideoArms is limited due to the deployed segmentation algorithm. While the system transmits a video showing hands and arms, it fails to transmit the collaborator's upper body which would mediate additional consequential and inconsequential communication such as pose and gaze that is not presented by the arms. Alternative segmentation algorithms are available [1], but entail time-consuming calculations, which lead to stuttering live video embodiments.

E-Chalk enhanced with SIOX [4] provides a very pleasing live video embodiment by using an elaborated segmentation technique that includes information from a depth-sensing camera. In order to mitigate any occlusion, the developers of the system chose to display the live video embodiment in a slightly translucent way. Note, that E-Chalk was developed to support distant learning, and not remote collaboration sessions. As a consequence, the system only provides a unidirectional live video embodiment (showing the teacher), and artifacts are not meant to be editable at both conferencing sites.

3 Contribution

In a team, a group of people collaborates in order to solve a given task. This group might have a local audience, but might also be geographically distributed and thus relies on a net-based collaboration (see Fig. 5). In order to overcome the limitations of VideoArms and E-Chalk with SIOX, and to master a typical net-based team session situation, we introduce a system that provides video- and data-conferencing with live video showing the full upper bodies of remote collaborators. The artifacts in the shared workspace can be edited by all interconnected stations. Placing the remote video atop the shared workspace, the deictic gestures can be transferred in correct relation to the shared artifacts, and thus they preserve their meaning.

The system consists of an audio- and video connection together with a shared whiteboard. In order to keep the shared whiteboard editable at both locations, video and content are acquired and transmitted separately. While the content of the shared whiteboard is transferred as application data, the person in front of each whiteboard is acquired by a camera from the perspective of the local audience, i.e., the camera stands opposite to the screen. Like for the local audience, also the acquired video image would consist of a person standing in front of the whiteboard's content. This content usually is highly inhomogeneous and also dynamic, which complicates the person's segmentation from the background. Thus, it is necessary to blank the shared workspace for the camera, while keeping it visible for the local audience and for the person standing next to the whiteboard. A time-multiplexing method, i.e., the insertion of black frames in certain time intervals, is not suitable since it requires a hard synchronization of the camera. Furthermore, this insertion of black frames would result in visible image flicker that could irritate the user. Thus, we take benefit from the fact that LC-screens already emit linearly polarized light that can be blanked for the camera by attaching linear polarization filters to it [12, 22].

Having now a static background that is independent of the actually displayed content on the whiteboard, we segment the foreground (user) by using an illumination-invariant method [14]. Previously captured sequences (images) without any user in front of the whiteboard are compared to the actual image. Using a statistical criterion that measures the collinearity of the actual color and the expected background in the color space, the segmentation is realized. The resulting image is then compressed and transferred as a video stream to the remote conferencing site. Here, the video stream is overlaid atop the shared content, which still is editable. In order to verify the efficiency of the realized setup, we performed user studies with the systems, having two collaborators together with a local audience at each site.

4 Hardware implementation

Two identical hardware setups were realized, each consisting of a large display, an interaction module, a camera with linear polarization filter, lighting bar, audio system, and a computer with network access.

The 65″ widescreen LC-display (SHARP LCD PN 655E) emits linearly polarized light, which can be easily blanked for the camera by using a linear polarization filter. The filter's orientation is twisted by 90 degrees compared to the orientation of the light coming from the screen. It is mounted directly in front of the camera so it cannot capture the content of the whiteboard anymore. Instead, the camera sees a very dark gray background. However, the user randomly reflects the ambient light in the room and thus is still visible to the camera.

In order to use the display as an interactive whiteboard, we use a commercially available interaction overlay (DViT PA 365 from Smart Technologies [18]). It employs four cameras for tracking, which are integrated in the corners of the overlay's frame. In addition, this module also provides pens with different colors, as well as an eraser. Since the DViT overlay is designed to be easily attachable to an LC-screen, the overlay's glass pane is atop the plastic frame of the LC-screen. Thus, the overlay's glass pane is in a certain distance to the matrix of the LC-screen. This results in a noticeable parallax distortion, which has a significant impact on the interaction precision of the user [16]. Small objects (targets) cannot be hit easily and the user has to continuously correct the position during his interaction.

Another drawback of the interactive overlay is the depolarizing effect of the glass pane. Now, the linearly polarized light coming from the LC-screen is partially depolarized, resulting in the effect that the whiteboard's content becomes slightly visible for the camera. Unwanted "shadows" occur, which cause problems for the segmentation algorithm. In this case, the segmentation algorithm cannot work properly anymore and artifacts in the video image could occur.

In order to eliminate the problems mentioned in the above, the display as well as the interactive overlay was completely disassembled. Within the reconstruction of the LC-screen, all spacers were removed and now the LC-matrix was only held in place by the interactive overlay. By this modification, the distance between image plane and interaction plane was reduced from 15 mm to 7 mm (see Fig. 2). We also replaced to glass of the interactive overlay by a high-transparent, non-depolarizing glass. By this, we could eliminate the visual artifacts in the segmented video image.

Finally, two identical setups were installed at two locations at the institute, which had neither a visual nor an acoustic direct contact. In order to not bias the succeeding user studies, the rooms were of same size and had the same infrastructure.
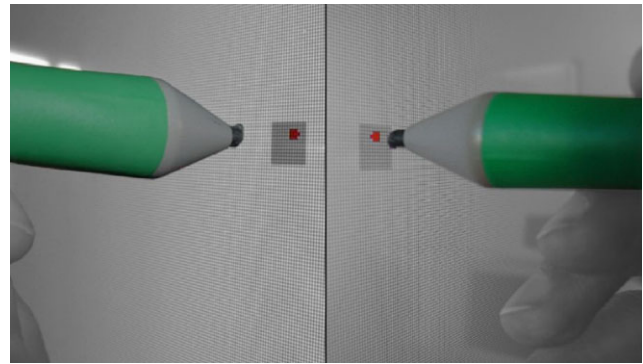

Fig. 2 Reducing the distance between image plane and interaction plane (*left*: original distance, *right*: modified screen)

5 Software

Classical video conferencing systems transfer the video image and the content of the whiteboard time-sequentially, or they display the video and the whiteboard's content on different screens. Thus, the video and the content do not have to be perfectly synchronized all the time. A continuous, real-time synchronization is complex to implement and requires a higher bandwidth. For instance, in most existing whiteboard applications, a stroke that has been drawn with a pen is only transmitted after the pen is removed again from the whiteboard's surface.

In the CollaBoard case, the video is part of the whiteboard content and therefore the whiteboard content must be synchronized in real-time between the remote peers. For example, if a local peer draws a line on the whiteboard, the video at the remote location shows him moving a pen on the whiteboard. In order to give the remote peer the sensation that a line is being drawn, the line must show up in real time while it is being drawn and in sync with the video. We call this feature real time strokes.

For the audio link between the CollaBoards, Skype audio-conferencing software is used [17]. The video connection is established by running a customized version of the open source software ConferenceXP [15]. ConferenceXP was modified and extended for the CollaBoard case. A module for video segmentation was implemented and the video rendering capability was modified such that the peer's video stream is shown in the foreground on top of all other applications.

Video segmentation essentially means that the background from the video is removed (i.e., made transparent) on the fly. Hence, only the person standing in front of the screen remains opaque in the video stream that is sent to the remote peer. Video segmentation quality is improved by applying morphological filters (see Figs. 3 and 4).
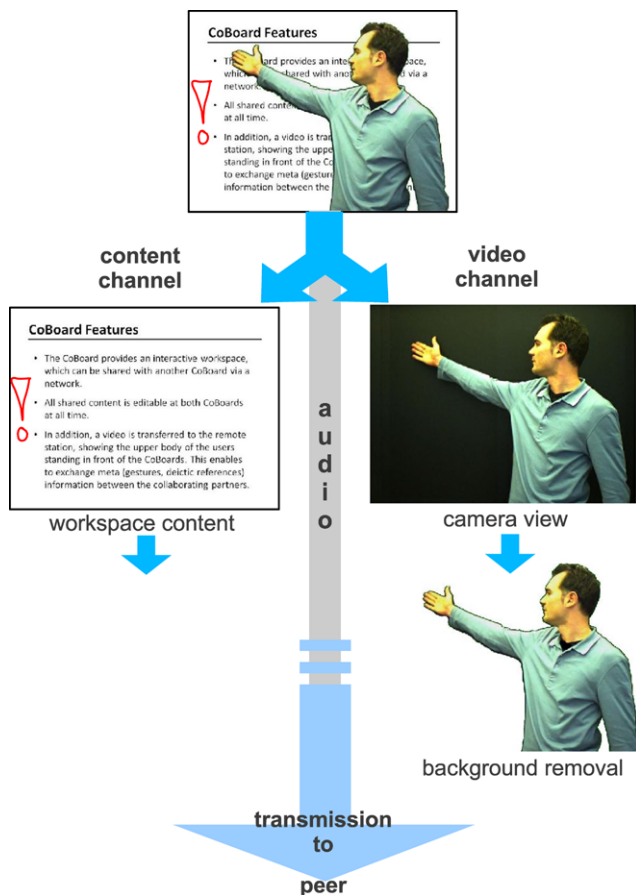
Fig. 3 Software working principle: content, audio, and video channel are split up and sent to the remote peer

5.1 CollaBoard whiteboard software

Many client-server or peer-to-peer based whiteboard applications provide synchronization of strokes across the network. But synchronization is done at the stroke level so that only finished strokes or user interactions have to be transmitted. Although this makes the synchronization significantly easier, it is not helpful for the CollaBoard and also introduces several problems. What happens for instance, if two peers are drawing something at the same time at the same place? Both will not notice that their partner is drawing as well, because the transmission takes place only after they will have lifted the pen. Such a stroke-based synchronization can be regarded as a time-multiplexed synchronization method. Strokes are sent to the server and then distributed to all the clients. The last stroke arriving at the server stays on top of previous strokes. In order to integrate support for real time strokes, there are basically two possibilities.

One possibility is to artificially cut strokes into smaller strokes, i.e., during a user interaction. Now, a movement is artificially split up into several smaller movements (the pen is virtually removed and set again). The shorter the time interval between the invisible splits is, the more strokes are created by one user interaction, and the more transmissions take place. Therefore, the interaction is transmitted almost in real-time, if the interval is small enough. The problem with this approach is that it generates a lot of overhead. The stroke database is filled rapidly, and erasing is much more expensive, as searches in the stroke database become more time-consuming. Furthermore, the stroke's smoothness is lost.

For our CollaBoard whiteboard application, we chose a different way to support real time strokes. The transmission of strokes is split up into a two-phase process. During a user interaction, the coordinates of the movement are recorded by a real-time transmission module and sent to the remote peer (the update rate of the transmitted coordinates can be reduced for slow network connections). The peer receives the points and renders them with a dynamic renderer on a drawing layer in the background. So, it seems as if the ink "flows" out of the peer's pen tip. This means that the dynamically rendered content of the peer remains in the background (but in front of background objects/images), while the local content stays in the foreground. Therefore, a drawing user is not disturbed by his partner while sketching. As soon as the user interaction is finished, the stroke is generated and sent to the server as above. Stroke generation essentially means that a B-spline curve is fitted to a subset of the original coordinates of the interaction. If a client receives a remote stroke from the server, the according dynamically rendered content (of the remote peer) gets replaced smoothly by the stroke in the foreground. This replacement is almost invisible to the user, except for the fact that the stroke comes to the foreground. Hence, our software allows two or more users to work on visual content from different locations at the same time.

The developed CollaBoard whiteboard application offers a plane white area for drawing. The physical pens from the SMART Technologies pen tray [18] (see Figs. 3, 4) can be used to draw on the drawing layer and to choose a color. The eraser can be activated as well by using the tray. Images can be loaded onto the background and are automatically sent to the server and distributed to all other connected clients. At the bottom of the window, there are buttons for switching pages. Clicking the page forward or backward button results in a page change request to the server. The server stores the current page, switches to the new one and sends an update to all clients. Figure 4 shows the whiteboard application in use.

6 User study

A user study was conducted in order to analyze the performance of the CollaBoard system.

The user study task is designed to be similar to a real life scenario, in which two remote groups have to work together
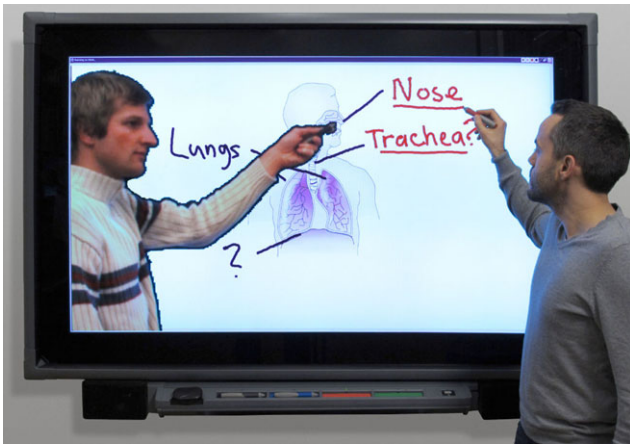
Fig. 4 The CollaBoard in use: Two remote peers are working on a common task using the whiteboard software

on a common problem. We believe that the new features of the CollaBoard offer a better and more natural way of collaboration among remote partners and, therefore, we compare it to other systems.

6.1 Apparatus and experimental conditions

The user study evaluates three different conditions. One condition is the CollaBoard system. The other one is a condition that represents standard video conference software. Finally, we also evaluated a control condition, in which the participants were working together at the same location and on the same whiteboard.

The hardware used for all conditions of the study was the CollaBoard screen with the tool tray and the interactive frame from SMART Technologies [18]. Users could use the tool tray to select one of four colors or an eraser; see Fig. 5. Only one tool can be used at a time. The CollaBoard whiteboard software was used for all conditions, but with different configurations. For the two remote conference conditions two rooms with identical hardware were prepared and for the audio link Skype was used [17].

Collocated condition (CO) In the collocated condition, both groups, which are usually working at remote locations, work together at the same CollaBoard in the same room. They share one whiteboard and, therefore, cannot sketch at the same time (i.e., no video overlay is used). The collocated condition simulates the optimal case, in which no remote collaboration is required and thus serves as a control condition for the user study.

Separate video condition (SV) In the separate video condition, a setup is simulated which is equal to traditional video conference software. Two groups are working in two different rooms with two CollaBoards. The whiteboard software is used, but with the real time strokes feature turned

off, i.e., if a user draws a stroke, the remote partner will see the stroke only after the local user removes the pen from the CollaBoard. The video of the remote user is presented on an extra 19″ screen (though no video overlay is used), which is placed on the left side of the CollaBoard. Video is transmitted with Skype. For the video link, a Logitech Pro9000 webcam sits on top of this extra screen, which records the upper body part of the CollaBoard user.

CollaBoard condition (CB) In the CollaBoard condition, all CollaBoard features are used as described in Sect. 5. Two groups are working in two different rooms with two CollaBoards. The video is recorded with the CollaBoard software and presented on top of the whiteboard content (people on content). The whiteboard software has the real time strokes feature turned on. The external 19″ screen and the webcam are not used, but the CollaBoard camera is active.

6.2 Participants

In total, 34 participants were recruited for the user study. The participants were split in two categories, CollaBoard users and local audience. CollaBoard users were actively working at the whiteboard and local audience members were joining the conference but were not allowed to work at the whiteboard (see below).

Fourteen subjects (12 male and 2 female, median age 22.5 years) were selected to be CollaBoard users. Twenty subjects (18 male, 2 female, median age 22 years) were selected to form the local audience.

For each study, two CollaBoard users had to work together on the study tasks. Each CollaBoard user had at least one person as local audience. Hence, we had fourteen groups to run seven studies. Four studies were done with students. Three studies were done with staff members. Local audience members were students. All students and staff members were from the mechanical engineering department. All subjects knew each other, but have not been working together on a digital whiteboard before. Staff members were familiar with digital whiteboards and basic conferencing tools (as in the SV condition). Students were given a general introduction and training phase of about two hours. No subject worked with the CollaBoard system before.

6.3 Task

The study task is designed to be similar to a real life scenario. The task should enforce the users to work collaboratively on a common problem and also provoke them to communicate a lot. Having a lot of interaction between the two users is crucial for evaluating the major improvements of the CollaBoard over classical conference systems. This is motivated by the media richness theory [2].
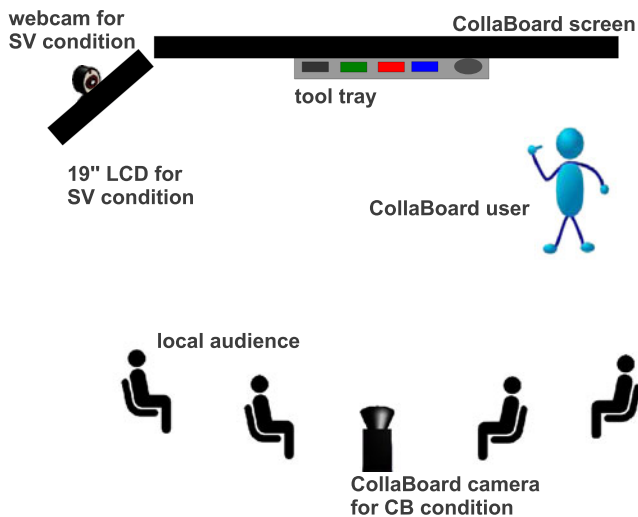
Fig. 5 Experimental setup for user study

In our study, we asked the participants to design a floor plan for a house. Both participants got the same task, but with different complementary requirements. One task for instance is to draw the floor plan of a new mansion. Group 1 has the requirement that the main entrance should be on the south side. Group 2 instead has the requirement that there is an entrance hall in the center of the house. Hence, the two remote partners have asymmetric information about the house they have to plan. Each participant has five requirements about the house which are different from his partner's requirements. They do not have sufficient information to design the house on their own. Both sets of requirements complement each other to a common solution. The uncertainty about the remote partner's requirements helps to enforce close collaboration.

The requirements are sorted from more general ones first to specific ones at the end. Therefore, the participants were asked to go through the given requirements stepwise.

Three different floor plan design tasks were created for the three different conditions.

6.4 Experimental setup

Two rooms with identical setups were prepared; see Fig. 5. Chairs were prepared for the local audience and placed to the left and right of the CollaBoard camera. For the SV condition, a 19″ display was placed to the left of the CollaBoard with a webcam on top. The CollaBoard user himself was able to move freely in the area between the screens and the local audience. A room microphone was placed next the CollaBoard tool tray and the loudspeakers were placed below the CollaBoard. The microphone could also pick up voices from the local audience.

For all three conditions, the two CollaBoard users were working on the same shared content. For the CO condition, both users were working in the same room on one CollaBoard.

6.5 Procedure

At the beginning of the study, two supervisors informed the participants about the task they were asked to solve. They were told that they have to solve three similar tasks on three different conditions. One condition is that two CollaBoard users work together locally and the other conditions are that they work together on the same problem in two different rooms with conference software. The local audience was split into two groups of similar size and assigned to a CollaBoard user.

We kept the order of the conditions the same for all studies, but we randomized the order of the three tasks according to a Latin square scheme. The ordering of the conditions was CO, SV, and finally CB. There were not enough participants to rotate the order of conditions as well.

Before each study on one condition, we quickly explained the features of the current condition as outlined in Sect. 6.1. Then we handed out the task sheets to the CollaBoard users and to the local audience. The local audience could participate in solving the task by giving hints to the CollaBoard user in the same room. But they were asked to remain seated throughout the study and could not join the CollaBoard user at the whiteboard. The participants had two minutes to read the task and then the CollaBoard users had to start solving the task.

The participants were told to solve the task quickly. The task was defined to be solved if all requirements for the floor plan were met by both groups. For the SV and CB condition, the participants were also told that they can work (sketch) simultaneously on a common solution.

In the CO condition, both groups stayed in the same room. The participants from both groups were not allowed to exchange their task sheets. After finishing the first task, the second group was asked to move to another room.

After all three tasks were solved under all conditions, the participants were asked to fill out a questionnaire with equal questions for the three setups. Figure 4 shows a typical situation at the CollaBoard.

7 Results

All groups managed to design correct floor plans and solved all tasks on every condition. No group had to be interrupted because they took too long. As suggested by the supervisors, all participants solved the task in a stepwise manner by starting with the first requirements first.

Figure 6 shows a resulting floor plan from one user study task. It was sketched by students during the user study in the CB condition.

Table 1 Average scores and standard deviation (SD) of the two user study questionnaires. The first questionnaire (question 1 to 13) was for the CollaBoard users and the second questionnaire (question A to E) was for the local audience. A Likert-type scale from 1 (strongly disagree) to 5 (strongly agree) is used

Question	CO		SV		CB	
	Mean	SD	Mean	SD	Mean	SD
COLLABOARD USERS (14 participants)						
1. I could easily make myself well understood	**4.36**	0.74	**3.64**	0.84	**3.79**	0.89
2. I could easily tell what my partner was pointing at	**4.43**	1.09	**2.57**	1.09	**3.29**	0.73
3. I could easily tell what my partner was referring to (verbally or by pointing gestures)	**4.43**	0.94	**3.00**	0.96	**3.21**	0.89
4. I found it easy to interrupt my partner when I saw him make an error	**4.29**	1.14	**3.29**	0.91	**3.43**	1.09
5. I could easily tell what my partner was looking at	**4.43**	1.09	**2.14**	0.86	**2.86**	0.95
6. My remote partner and I continuously communicated with each other	**4.64**	0.50	**3.93**	1.14	**4.00**	0.96
7. I was never confused	**4.21**	1.12	**3.21**	0.80	**2.86**	0.95
8. We never spoke at the same time	**3.36**	1.15	**3.21**	1.05	**3.07**	1.07
9. I often looked at my partner	**4.00**	1.18	**2.36**	1.08	**2.86**	1.35
10. I exactly knew when it was my turn to speak	**3.93**	1.00	**3.14**	0.95	**3.07**	1.00
11. I could easily tell if my partner was listening carefully to what I said	**4.21**	0.97	**2.79**	1.37	**3.00**	1.18
12. When I looked at my partner, I could always clearly see his or her face	**4.50**	1.16	**2.86**	1.41	**1.86**	0.95
13. I am happy with the elaborated floor plan	**4.36**	0.93	**3.50**	0.94	**3.79**	1.05
Mean over all questions from questionnaire	**4.24**	1.03	**3.05**	1.12	**3.16**	1.11
Mean over all questions from questionnaire without questions 7 & 12	**4.22**	1.02	**3.05**	1.12	**3.31**	1.06
LOCAL AUDIENCE (20 participants)						
A. I could easily tell what my remote partner was pointing at	**4.65**	0.75	**2.30**	0.86	**4.05**	0.76
B. I could easily tell what my remote partner was looking at	**4.40**	0.99	**2.10**	0.85	**3.40**	0.94
C. I was never confused	**4.20**	0.83	**3.25**	1.07	**3.60**	0.94
D. I often looked at my remote partner	**4.45**	0.69	**2.20**	1.06	**3.30**	1.13
E. I could easily tell what my remote partner was referring to (verbally or by pointing gestures)	**4.55**	0.83	**2.90**	0.85	**3.85**	0.75
Mean over all questions from local audience questionnaire	**4.45**	0.82	**2.55**	1.03	**3.64**	0.94
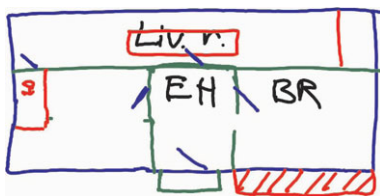
Fig. 6 Solution of one task sketched by students during the user study in the CB condition

Clearly, we expected that the CO condition outperforms the other conditions. We hoped that the CB condition performs better than the SV condition and that there is a significant difference between the SV and CB condition.

During the study, the supervisors also recorded the total task completion time for each condition. The mean completion time was 6.7 minutes for the CO condition, 7.1 minutes for the SV condition and 5.9 minutes for the CB condition. Because we did not rotate the order of the conditions there

was a training effect. Also, the standard deviation was quite high (around 2 minutes under each condition). Hence, we omit further analysis on the completion time.

7.1 Questionnaire results

After finishing all three tasks, the participants had to fill in a questionnaire for evaluating the usability. There was a questionnaire with 13 questions for the CollaBoard users and a questionnaire with 5 questions for the local audience members. Each question had to be answered three times, once for each condition. The scale of the questionnaire is a Likert-type scale from 1 ("strongly disagree") to 5 ("strongly agree"). The questions were partially taken from [6] and modified to fit the CollaBoard user study. Table 1 presents both questionnaires, the mean response per question and setup, and the standard deviation. The questionnaire was designed to allow for a summative analysis.

8 Discussion

As we expected, the CO condition outperforms the SV and the CB conditions. This holds for all questions on both questionnaires. But nevertheless, even the CO condition did not achieve an average score of 5 (strongly agree) on all questions. According to different statements of the participants, we assume that this is due to social issues between the two CollaBoard users. As mentioned above, the participants have not been working together at a digital whiteboard before. In several cases, especially in the CO condition, one user strongly dominated at the whiteboard, i.e., one user was drawing all the time while the other one just read out the requirements on his task sheet and did not participate actively in drawing the floor plan. Such a strong domination by one participant could not be observed during the SV or CB condition. It should be noted that this domination could also be influenced by the fact that only one user could sketch at a time in the CO condition.

The SV condition performs significantly worse than the CO condition for all questions in both questionnaires.

The results of the CollaBoard user questionnaire (see Table 1, mean over all questions) show that with a summative analysis, the CollaBoard does not perform significantly better than the SV condition. Although on average it is better, the variance in the data is too high to show a significant difference.

In contrast to this, the local audience questionnaire shows significant differences. The boxplot in Fig. 7(a) suggests that the CollaBoard helped the local audience a lot to follow the task. On all questions for the local audience, the CB condition was rated better than the SV condition. Especially question A shows that the people on content feature of the CollaBoard works as planned—the participants could follow gestures almost as easily as in the CO condition (see Fig. 7(b)).

In the CollaBoard user questionnaire, the CB condition is rated worse than the SV condition in four questions. As expected, especially for questions 7 and 12, the CollaBoard is rated significantly worse than the SV condition. But this is by design. As in the CB case, the video is on top of the content. The user working at the whiteboard usually does not see the face of his opponent directly (face to face), but from the side. He will see gestures very well instead. The remote user's face is clearly visible to everyone if he looks at his local audience. In the SV condition, the local user can see his opponent clearly but not in relation to the content and only if he decides to look at the webcam. Therefore, a user has to stop working to look at his partner in the SV setup.

Comparing question 12 to questions B and D in the local audience questionnaire implies that the local audience could actually see the remote peer's face. Hence, the CollaBoard user does probably not see the remote peer face to
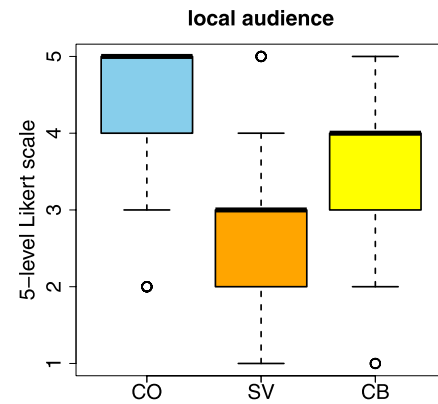
local audience

(a) Summative analysis of all questions of the local audience questionnaire

question A – local audience
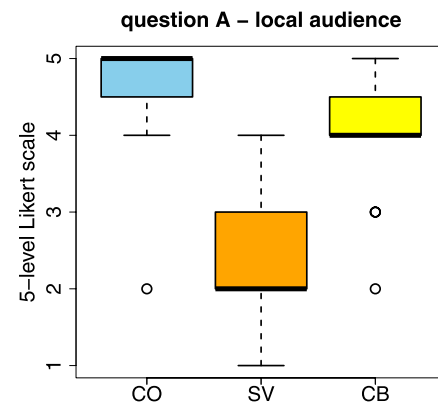
(b) Analysis of question A of the local audience questionnaire

Fig. 7 The boxplots for the local audience questionnaire show the median values. The lower and the upper end of the box are the 25th and the 75th percentile. The whiskers extend to 1.5 times the IQR. *Points* mark outliers

face most of the time, but the gaze in relation to the content is preserved.

Question 7 suggests the video on content feature confused many users. Most participants have been using Skype video conferencing at home before, but they all have not seen or used any video on content conferencing solution before. The confusion usually came from the fact that they were recorded from the back and were not aware of the fact that their body is inserted into their remote partner's content, even though they could see their partner on the screen. Therefore, some participants kept blocking the screen so that the remote partner could not work.

Hence, for a summative analysis of the CollaBoard user questionnaire, we ignore questions 7 and 12. In this case, the CollaBoard performs better than the SV condition (see Fig. 8).
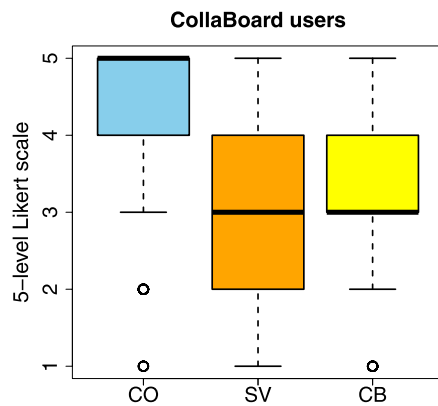
CollaBoard users

Fig. 8 Summative analysis of all questions of the CollaBoard user questionnaire except questions 7 and 12. The boxplot shows the median value. The lower and the upper end of the box are the 25th and the 75th percentile. The whiskers extend to 1.5 times the IQR. *Points* mark outliers

9 Summary and outlook

In this paper, we presented the CollaBoard—a novel whiteboard and conferencing system for remote collaboration. The CollaBoard is both, a video- and data-conferencing system. On the one hand, it provides real-time synchronization of user generated artifacts. On the other hand, it augments the shared workspace by overlaying the video image of the remote partner. For this, modern techniques such as real time foreground segmentation are used.

In a user study, the CollaBoard was compared to traditional systems. Results show that the CollaBoard provides a better user experience compared to traditional remote conferencing systems with whiteboards. We could also show that a local audience can more easily follow the task.

Further research is planned to solve the problem where a user working at the CollaBoard is blocking the screen. This happens for instance when a user is moving to close to the camera so that a large part of the field of view is blocked. Hence, the remote user cannot work at the whiteboard anymore. A possible solution to this problem is to render the video overlay semitransparent. A more advanced solution would be to continuously adapt the transparency of the video overlay based on different factors. For example, the video overlay could be made transparent if the local user is standing at the same location in front of the screen as the remote user.

References

1. Coldefy, F., Louis-dit Picard, S.: Remote gesture visualization for efficient distant collaboration using collocated shared interfaces. In: Proceedings of the Second IASTED International Conference on Human Computer Interaction, IASTED-HCI '07, pp. 37–42. ACTA Press, Calgary (2007)
2. Daft, R.L., Lengel, R.H.: Information richness: a new approach to managerial behaviour and organizational design. Res. Organ. Behav. **6**, 191–233 (1984)
3. Everitt, K.M., Klemmer, S.R., Lee, R., Landay, J.A.: Two worlds apart: bridging the gap between physical and virtual media for distributed design collaboration. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI'03, pp. 553–560 (2003). doi:10.1145/642611.642707
4. Friedland, G.: Adaptive audio and video processing for electronic chalkboard lectures. Lulu. Com. (2006)
5. Gaver, W.W., Sellen, A., Heath, C., Luff, P.: One is not enough: multiple views in a media space. In: Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems, CHI '93, pp. 335–341 (1993). doi:10.1145/169059.169268
6. Hauber, J., Regenbrecht, H., Billinghurst, M., Cockburn, A.: Spatiality in videoconferencing: trade-offs between efficiency and social presence. In: Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work, CSCW '06, pp. 413–422. ACM, New York (2006). doi:10.1145/1180875.1180937
7. Ishii, H., Arita, K.: Clearface: translucent multiuser interface for teamworkstation. In: Proceedings of the Second Conference on European Conference on Computer-Supported Cooperative Work, pp. 163–174 (1991)
8. Ishii, H., Kobayashi, M.: Clearboard: a seamless medium for shared drawing and conversation with eye contact. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '92, pp. 525–532 (1992). doi:10.1145/142750.142977
9. Ishii, H., Arita, K., Kobayashi, M.: Towards seamless collaboration media—from teamworkstation to clearboard. NTT Rev. **5**(1), 24–29 (1993)
10. Kirk, D., Stanton Fraser, D.: Comparing remote gesture technologies for supporting collaborative physical tasks. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '06, pp. 1191–1200 (2006). doi:10.1145/1124772.1124951
11. Kirk, D., Rodden, T., Fraser, D.: Turn it this way: grounding collaborative action with remote gestures. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '07, pp. 1039–1048 (2007). doi:10.1145/1240624.1240782
12. Kunz, A., Nescher, T., Kuchler, M.: Collaboard: a novel interactive electronic whiteboard for remote collaboration with people on content. In: 2010 International Conference on Cyberworlds (CW), pp. 430–437 (2010). doi:10.1109/CW.2010.17
13. Louwerse, M., Bangerter, A.: Focusing attention with deictic gestures and linguistic expressions. In: Proceedings of the 27th Annual Meeting of the Cognitive Science Society (2005)
14. Mester, R., Aach, T., Dümbgen, L.: Illumination-invariant change detection using a statistical colinearity criterion. In: Pattern Recognition. Lecture Notes in Computer Science, vol. 2191, pp. 170–177. Springer, Berlin (2001)
15. Microsoft Research ConferenceXP project (accessed 6 December 2010). URL http://research.microsoft.com/en-us/projects/conferencexp/
16. Migge, B., Kunz, A.: User model for predictive calibration control on interactive screens. In: 2010 International Conference on Cyberworlds (CW), pp. 32–37 (2010). doi:10.1109/CW.2010.18
17. Skype video conferencing software (accessed 6 December 2010). http://www.skype.com
18. SMART Technologies (accessed 6 December 2010). http://smarttech.com
19. Stotts, D., Smith, J.M., Jen, D.: The vis-a-vid transparent video facetop. In: Proc. of UIST 2003, pp. 57–58 (2003)

20. Tang, J.: Findings from observational studies of collaborative work. Int. J. Man-Mach. Stud. **34**(2), 143–160 (1991)
21. Tang, J.C., Minneman, S.: Videowhiteboard: video shadows to support remote collaboration. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Reaching Through Technology, CHI '91, pp. 315–322 (1991). doi:10.1145/108844.108932
22. Tang, J.C., Minneman, S.L.: Videodraw: a video interface for collaborative drawing. ACM Trans. Inf. Syst. **9**, 170–184 (1991). doi:10.1145/123078.128729
23. Tang, A., Neustaedter, C., Greenberg, S.: VideoArms: supporting remote embodiment in groupware. In: Video Proceedings of CSCW, vol. 4 (2004)
24. Tang, A., Neustaedter, C., Greenberg, S.: Videoarms: embodiments for mixed presence groupware. In: People and Computers XX—Engage, pp. 85–102 (2007)

Thomas Nescher is a researcher in the ICVR group (Innovation Center Virtual Reality) at ETH Zurich in Switzerland. He holds a M.Sc. degree in Computer Science with specialization in Visual Computing from ETH (2007). Thomas is currently a Ph.D. candidate student at ETH, examining optimization strategies for navigation in immersive virtual environments. His research interests cover Human Computer Interaction fields, ranging from remote collaboration to virtual reality applications and navigation in virtual environments.

Andreas Kunz was born in 1961, and studied Electrical Engineering in Darmstadt/Germany. After his diploma in 1989, he worked in industry for 4 years. In 1995, he became a research engineer and Ph.D. student at ETH Zurich/Switzerland in the Department of Mechanical Engineering. In October 1998, he finished his Ph.D. and established the research field "Virtual Reality" at ETH and founded the research group ICVR (Innovation Center Virtual Reality). In 2004, and became a private docent at ETH. Since July 2006, he is Adjunct Professor at BTH. Since 1995, he has been involved in students' education and since 1998 he has been giving lectures in the field of Virtual Reality. He also gives lectures abroad, in Switzerland as well as in other countries such as Germany, Sweden, USA, Romania, etc. Dr. Kunz published in IEEE Virtual Reality and PRESENCE, and reviews papers for several IEEE conferences.