Joint Virtual Reality Conference of EGVE - ICAT - EuroVR (2009)
M. Hirose, D. Schmalstieg, C. A. Wingrave, and K. Nishimura (Editors)

Collaborative Interaction in Co-Located Two-User Scenarios

H. Salzmann[1] J. Jacobs[2] B. Froehlich[3]

[1]Carmeq GmbH, Member of Volkswagen Group, Germany
[2]Volkswagen AG, Germany
[3] Bauhaus-Universität Weimar, Germany

Abstract

We investigated the utility of co-located collaborative interaction metaphors for assembly tasks in the automotive industry. In a first expert review, we compared the usability of a regular single-user stereoscopic system to a two-user system. Our experiments revealed that the two-user system greatly facilitates basic collaborative interactions, when users are standing next to each other. However, it is unsuitable for scenarios wherein users are facing each other and more sophisticated collaborative interactions are required. Our second study focused on these types of collaborative assembly tasks using head mounted displays instead of our projection-based setup. In a virtual assembly task, two workers had to mount the windshield of a car by using two different interaction methods. The first method employed tangible props and the other method relied solely on virtual interaction. Our evaluation shows that the multi-user prop-based interaction results in significantly higher accuracy and was clearly preferred by our users. These results are an important step towards the acceptance of such virtual techniques for reliable ergonomic evaluations of virtual assembly tasks.

Categories and Subject Descriptors (according to ACM CCS): Computer Graphics [I.3.1]: Hardware Architecture—Three-dimensional displays; Computer Graphics [I.3.6]: Methodology and Techniques—Ergonomics; Computer Graphics [I.3.6]: Methodology and Techniques—Interaction techniques; Computer Graphics [I.3.7]: Three-Dimensional Graphics and Realism—Virtual reality

1. Introduction

Immersive virtual environments are widely employed in the automotive industry. In the area of assembly simulations they are used for two different aspects: testing and optimization of the assembly processes; and the assessment of the ergonomics of an assembly sequence. In both situations the display of a full scale model and the involvement of multiple people are crucial for the validity and significance of these evaluations. It is often argued that nowadays the largest part of the work at the assembly lines is done by robots, but there are still many tasks, especially in the maintenance sector, which are done by two workers working closely together.

We implemented different collaborative assembly scenarios supporting side-by-side as well as face-to-face interactions between two co-located users. All our scenarios have in common that they require direct interaction within the reach of the users' arms. For side-by-side interactions of two workers, such as standing in front of a car and mounting engine parts, the workers employ different tools or they assume different roles. We also implemented face-to-face interactions wherein two workers are collaboratively mounting a windshield. Our goal was to assess the suitability of different co-located immersive virtual reality solutions and appropriate interaction forms for these types of tasks.

All regularly used stereoscopic projection systems in our labs support only one tracked user, which limits their usability for tasks involving two active users. To improve the situation, we set up a prototypical projection-based two-user system providing two users with individual stereoscopic images. We evaluated this setup in an expert review for collaborative side-by-side interactions and compared it to the commonly used single-user projection setup. The results indicate that the two-user system greatly facilitates basic collaborative interactions and it may prove to be a standard evaluation platform for side-by-side tasks.

Our experts were also interested in looking at other tasks; more specifically, tasks wherein users have to work face-to-face. These tasks can only be supported by an HMD-based two-user system since any projection setup would result in problematic occlusions if an object would have to be displayed between the two users. Thus, we set up an HMD-based study comparing two methods of collaborative interaction for a windshield assembly task performed by two workers. The first method used a tangible two-user prop representing the windshield and the other relied solely on virtual interaction. Our findings indicate that in this HMD-based face-to-face assembly task, untrained users rapidly became proficient with the task and the prop-based method resulted in enough realism to validate ergonomic issues in a virtual environment. The haptic link and the constraints provided by the windshield prop where major factors in improving the coordination between two users. Due to missing force-feedback, the virtual method is much less suitable for ergonomic evaluations since the task duration times were much higher than in real life, and the two-user windshield mounting task could not be performed with enough precision. In general, co-located collaboration in automotive assembly tasks involving face-to-face combined with side-by-side interactions seems to be best supported by HMD-based multi-user setups augmented with multi-user props. Due to better ergonomics, multi-user projection displays are well suited if the task requires only side-by-side interactions.

2. Related Work

The main research topics touched by our work are interaction with three-dimensional user interfaces augmented by real-world objects and multi-user scenarios in collaborative virtual environments (CVEs).

Hinckley et al. [HPGK94] show that the use of familiar real-world objects as a tangible interface may facilitate interaction in the virtual world. They used a tracked doll head or rubber ball to represent a tomography dataset of the human head. These passive interface props help users to understand the functionality of the represented objects or tools. Novice users directly interact with props and their corresponding virtual representation without requiring much training. The success of Nintendo's Wii shows that the use of tracked handles connected to a virtual object is a strong interaction metaphor that may also partially compensate for missing force feedback. In our work we use tracked handles, called flysticks, to represent the handle of virtual assembly tools. Our windshield prop is an aluminum pole with a handle at each end. This two-user prop provides passive force-feedback and a direct physical link between the two users. It also represents the physical dimensions of the virtual object.

Margery et al. [MAP99] defined three levels of cooperation. In our scenarios, two users coexist in a CVE and can perceive each other and communicate with each other (Level 1). Each user can individually modify the scene (Level 2) and for our symmetric interaction scenarios, the users can simultaneously manipulate the same object (Level 3). Ruddle et al. [RSJ02] describe some important aspects that should be taken into account when designing symmetric interactions. It is necessary to define rules that manage inputs from the participating users (e.g. transferring the forces from each user's hands into object movement). Pure virtual object manipulation without haptic feedback requires dealing with physically impossible situations such as hand-object penetrations.

Collaborative virtual environments (CVEs) support the interaction of multiple remote or co-located users in a shared virtual environment. Our system shares similarities with other co-located setups like the Studierstube project [SFH*02], the PIT [APT*98] and the Two-User Responsive Workbench [ABF*97]. Two or more users are co-located in the same room and experience the same virtual scene. Each of these systems generates an individual stereoscopic view for each user, enabling collaboration on a virtual scenario.

3. Side-by-Side Collaboration

The intention of this study is to investigate the advantage of a stereoscopic two-user system over a conventional single-user solution for common automotive scenarios. Our first scenario is an authentic assembly planning task that was simulated some time ago in the development process of a new car. We invited the original six participants, real experts in the field, to be our interviewees. The focus was on the investigation of the accessibility of certain car parts by tools. The second scenario involves a typical video-based training of a new assembly method by a hands-on training sequence in our multi-user system. We presented this extended training scenario to five experienced automotive coaches in our interviews. What both scenarios have in common is that they require the execution of certain sequential interaction tasks as well as asymmetric interaction. Sequential means that they have to do similar actions using a single tool. For example, both have to fix screws, but there is only one screwdriver available. Asymmetric actions are those wherein one user moves the object and the other has to fix some screws. This way they are assuming different roles in the assembly task.

The two scenarios support similar interaction techniques. Users were able to point at different parts with their real fingers, which was enabled by the perspectively correct views in the multi-viewer projection system. In our previous work [SMF09] we found out that pointing with the real hand is an intuitive interaction metaphor which is highly appreciated in our collaborative virtual scenarios. In the current experiment, we supported two flystick input devices (optically tracked handles with buttons) that were used in two ways. One option was to have one flystick for the scene movement and the other controlled a virtual tool attached to it. This enables one user to navigate the scene while the other can check mounting points with the virtual tool in his hand. This task division initiates communication between users. The user with the tool has to ask for viewpoint changes if he cannot properly see or reach the part he is supposed to manipulate. Alternatively we assigned a different tool to each flystick. Users often switch their roles by simply exchanging input devices. Figure 1 shows the handover of a tool by passing the input device to the other user.

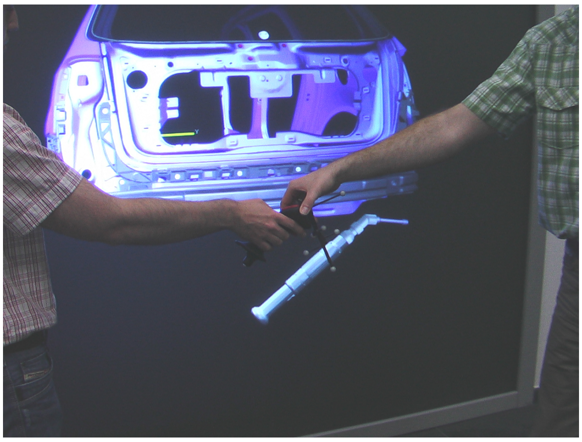

Figure 1: *The users are able to hand over a virtual tool that is mapped onto a tracked flystick while investigating several screw mountings.*

Our two-viewer display system uses liquid crystal (LC) shutter elements mounted in front of LC-projectors to separate the individual users [FBS*05] and provides a framerate of 60Hz per eye per user. Polarization is used to separate the left and right eye view of each user. The size of the display is 3 meters by 2.25 meters, the display's resolution is 1400x1050 pixels and the projectors' brightness is 3500 lumens. The shutters in front of the projectors and wired shutter glasses are controlled by a custom micro-controller circuit. Each projector pair for a single user is driven by a separate computer equipped with a dual head graphics card. Both systems are driven by our inhouse software, which keeps the states of both applications in sync.

3.1. Task Description

3.1.1. Assembly Planning Scenario

The first scenario is inspired by an existing protocol of a typical assembly simulation. The task was aimed to inspect different areas in the front of a car:

- The users initiate a discussion by pointing at certain parts with the real hand and naming the parts.
- One user checks screw mountings on the left side of the car's hood. The other on the right side. They have to perform a handover of the used tool.
- One user has to move the front part of the car such that an otherwise occluded part becomes visible. The other user then has to fix some screws on this part using a tool.
- The users have to navigate to see the windshield wiper and check for collisions on the surrounding car parts.

3.1.2. Training Scenario

The second scenario is a simulation of a training application. For certain assembly tasks, animated 3D-sequences are well accepted methods used to illustrate and train the mounting of car parts. We showed such a tutorial of the mounting of a part at the rear end of the car on a separate projection screen next to the two-user display (Figure 2). To improve upon the standard process, we allowed users to interactively perform the just seen task with our stereoscopic two-user system. The scenario can be divided into these steps:

- The users initiate a discussion by pointing at certain parts with the real hand and naming the parts.
- One user has to fix screws of the back bumper with a tool. The other has to position the car in the desired orientation.
- The users have to check several mounting points if they are reachable and fixable with the automatic screwdriver.
- The users are able to exchange the flysticks at any time.

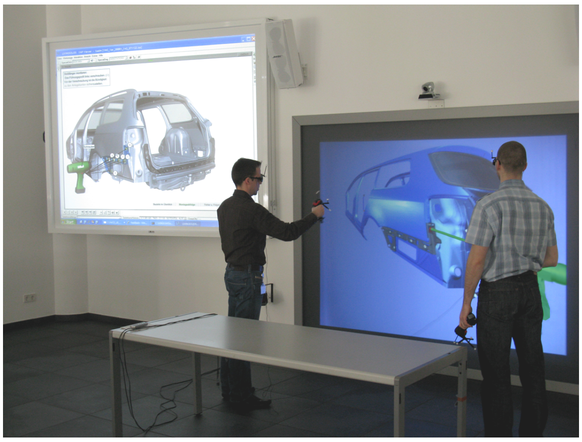

Figure 2: *The training scenario allows users to first watch a video of the assembly sequence on a regular monoscopic display on the left and then perform this sequence in the stereoscopic two-user system in single-user as well as in two-user mode.*

3.2. Evaluation

For an initial evaluation of the usability of our two-user display, for both scenarios we performed an expert review involving six automotive experts. With this method as described in [PRS02], usually 75 percent of the usability problems can be found by asking only five experts. We chose a semi-structured group interview, allowing new questions to be brought up which result from the discussion with the interviewees. The interview is guided by general objectives that should be explored and it does not have to stick to a fixed set of questions. Each interview took about 20 minutes.

In the beginning we instructed the participants that their task is to compare the new two-user setup to the usual single-user setup. A group of interviewees had to go through a given agenda. For each scenario, users first performed their tasks in single-user mode, wherein only one user was tracked and the second user saw the images computed for the first user's viewpoint. Then both users were tracked and provided with individual stereoscopic images (two-user mode). Finally, users were exposed to single-user mode again. Thus each pair of users performed a scenario three times, which took approximately 25 minutes. The first pass allowed the participants to become familiarized with the scenario task. In the following passes, they were then able to compare both configurations.

3.3. Results

Independent of the scenario, the experts stated that it is very hard or even impossible for the single-user slave (non-tracked user) to interact in any way since the single-user master's (tracked user) head movement affects his view. Some of them had the fear of becoming sick considering that the movement was not controlled by them and occurred without their complete awareness. They also realized that our artificial single-user mode has the same shortcomings as the stereoscopic projection system they were used to and that the distortion gets worse the more they are away from the tracked user. We also found out that in earlier immersive presentations, novice users from other departments were not told that they had a distorted view and so they accepted the distortions as a given disadvantage of virtual reality technology overall.

The asymmetric task division, wherein one user controlled navigation and the other a tool, was not well received. In particular, training experts said that it would be a better option to give one experienced user of virtual reality technology (e.g. the teacher or instructor) the ability to control navigation and menu functions to avoid a chaotic presentation in a potentially larger group of users. In a real assembly situation there is no need to navigate the scene because the observers are usually moving around the car. In our setup we needed this function to compensate for the relatively small viewing volume which prevents the users from walking around the car as they would in reality. The possibility to hand over a tool or to exchange tools with different functions between users was seen as an improvement compared to single-user mode. For the single-user slave, the tools had the same distortion as the parts of the scene, which made a correct grasping and operation impossible. Another effect was that distorted parts and tools were clipped at the boundaries of the viewing frustum.

The expert-interview has shown that in a setup consisting of only a single projection screen, it is not essential to have avatars representing the users or their hands. In most cases the users are standing relatively close together and do not look at each other. In addition shutter glasses allow the users to see their real bodies and hands such that gestures and postures can be well communicated. In both scenarios we did not use a virtual hand during the interviews since we expected it is not that important in a projection-based setup where flysticks are used as props for graspable tools. All of the experts confirmed that they did not miss their virtual hand. One of them said that a virtual hand is only an additional object that occludes the more important virtual tools and car parts.

4. Face-to-Face Collaboration

At the end of our first study the experts pointed out that it would be interesting to have two or more users looking from different sides at the car while they are collaboratively working. Our projection-based system enables two users to look at a virtual model through the same window defined by the projection screen. This works well for scenarios where both users are standing next to each other in front of a virtual model such as the front or back of a car. Many collaborative assembly tasks require that two workers are partly facing each other to successfully perform the task. Our first study showed that it is hardly possible to simulate those tasks using projection-based setups even if they support multiple tracked users and multiple projection screens like a CAVE^{TM}. If the users would attempt to face each other, the projection of the virtual environment would always be occluded by their real bodies – except for cases wherein a table-top setup is sufficient. However, this is not the case for most automotive 1:1 scale simulations. For enabling face-to-face scenarios, we decided to implement a HMD-based two-user setup. We represented each user in the virtual world by a basic avatar consisting of a virtual head and two virtual hands. Each user was able to see his own hands and the hands and head of the other user. The two users' movements were restricted to a 2.80m x 2.50m optical tracking-frame with a 2.20m height. We used nVisorSX HMDs providing a resolution of 1280x1024 and an update rate of 60Hz while covering a diagonal field of view of 60 degrees. For the hands we used tracked ART-gloves with active LED-markers. The graphics update rate was 60Hz for all our scenarios.

Our second study investigates assembly scenarios wherein two users are collaboratively working on the same object. In reality such a concurrent interaction of two people becomes necessary when relatively large and heavy objects need to be moved. Assembly tasks like this can be found in the production process of a car on the assembly line or more often during maintenance tasks in garages. Even during early development stages of a new car, it is necessary to perform simulations of the later production process considering it should be possible to efficiently assemble and disassemble a car. Such simulations are evaluated by assembly planners and ergonomists. Together they carefully design the assembly line processes following guidelines of ergonomics and effectiveness. Based on standard ergonomic simulations, which do not differ very much between different cars, we chose a suitable real world scenario for collaborative assembly tasks. Such tasks are described in detail by process animations in CAD systems, which we used as guidelines to design our virtual scenarios. One assembly process usually consists of several steps, but for our tasks we extracted only the collaboratively performed process steps.

4.1. Windshield Assembly Task

Windshield assembly is typically done by two workers using only one of their hands. They pick up the windshield from a rack next to the car by grasping it on each side using handles. The worker standing right grasps it with his left hand and the worker standing left uses his right hand. Then they carefully move the windshield over a distance of 2 meters to the front of the car and place it in its end-position. The intended task completion time (TCT) suggested by the CAD process animation is 20 seconds.

Figure 3: *Snapshot from the video showing the process animation of the windshield assembly task.*

4.2. Interaction Techniques

Various approaches have been suggested to implement simultaneous manipulation of a single object by two or more virtual hands or users [CFH97] [FTB*00] [RSJ02] [PBF02] [DLT06]. In contrast to this previous work, we decided to compare an unadulterated virtual method to a method based on a tangible two-user prop which provides a physical link between two users.

The prop-based method uses a simple aluminum pole with two attached handle bars to represent the windshield. Our windshield-prop weighs only about 3.5 kg, while the real windshield weighs about 14 kg, but at least the same centroid was provided. The pole prop also enables an additional communication channel between both users by transmitting input forces via the pole.

The virtual method required that two users had to manipulate the windshield without the passive haptic feedback of the windshield prop. Only visual feedback was provided by the movement of the virtual windshield. The windshield model was attached to the user's hand once the hand was in close proximity to the holding point. The windshield was automatically released once the target position was reached. We used simple averaging of translations and rotations of the two hands to move the virtual windshield.

4.3. Evaluation

The stability of the hand positions over the duration of the task is an indicator of the reliability of the virtual evaluation of ergonomic aspects for the described assembly task. Thus, we were interested in how much the positions of the hands in the virtual condition varied with respect to the virtual windshield's holding points. The prop-based approach should result in quite fixed hand positions. Visual dominance is the cause for an effect we could observe in earlier vr-sessions: if a virtual object connected to the user's hand stops moving, the user also stops moving his hand even if there is no haptic feedback forcing him to stop. In our case, we expected that the participants would try to keep their hands inside the holding spheres of the virtual handles on the object. That should lead to a stable distance between the user's hands holding the windshield over the duration of the task. We speculated that visual dominance and verbal communication between users could compensate for missing haptic feedback.

Twenty subjects aged 25 to 45 (18 male, 2 female) volunteered to participate in our study. All of them had correct stereo vision. They had different backgrounds, including assembly planners, ergonomists and students with different virtual reality experience. Our post questionnaire aimed at demographic issues, VR-experiences and the task itself. We also included items addressing collaboration [BHG01], involvement [HB03], awareness [GMH01], co-presence [SSA*01], usability and preference. The questions were to be answered on a scale of 1 to 5 (1 = to a very small extent and 5 = to a very large extent). As statistical measures, we used one-way ANOVA to compare mean differences using a 5% confidence interval. At the end of the questionnaire we provided space for general comments and remarks.

The participants completed the task in pairs of two using the virtual method as well as the prop method. At the beginning they were told to watch a video showing the task in order to get familiar with the movements they had to perform and how to manipulate the object (see Figure 3). They were instructed that it was important that they mimicked the task as precisely as possible. They were also able to complete one test trial in advance before the evaluation started. Each pair

Figure 4: *Comparison of prop-based and virtual manipulation technique. Holding points were marked by small spheres.*

performed ten trials of each task and with each manipulation method. For each trial we recorded the positions of the hands via the gloves and the center of the windshield during the manipulation as well as the time needed to complete the task.

After putting on the HMDs and the gloves, the participants where asked to stand in front of the car next to the rack with the windshield and to orient themselves in the virtual environment. They then picked up the windshield by its handles and began to move it to their target position. They had to walk about 2 meters while balancing the windshield by adjusting the positions and orientations of their hands. During the preparation of the scenario we identified a deviation of two degrees and one centimeter from the object center as a suitable tolerance that still allows to find the target position in the virtual condition without frustration and with acceptable task completion times. The tolerance for the prop-based method was the same. As soon as the object was within the tolerance area near the target position, the object was colored dark-green. Then the participants had to hold this position for three seconds in order to prevent random matches, and the object's color changed to light-green and then froze into place. At this point all measurements were stopped and the task was completed. Collisions between windshield and car were not considered.

4.4. Results

Table 1 shows the presence measures for our collaborative HMD-scenario and the two different interaction methods. Collaboration and awareness were significantly better for the prop-based method, and involvement was still marginally better as well. Co-presence did not prove significant, which seems reasonable considering that the visual representations of both users were identical for both tasks. At this point we also want to mention that the participants reacted with great pleasure when seeing each other through their HMDs; some of them waved at each other or even tried to shake hands.

As Table 2 indicates, the usability of the prop-based interaction was rated significantly higher than the usability of the purely virtual interaction. Most of the participants (93.75%) preferred the prop interaction over the virtual interaction method (6.25%).

Table 1: *Presence measures for the assembly scenario corresponding to the used technique.*

item	prop	virtual	ANOVA
collaboration	4.15 ($\sigma = 0.88$)	3.41 ($\sigma = 1.28$)	F(1,38)=4.05 p=0.009
involvement	4.12 ($\sigma = 0.80$)	3.60 ($\sigma = 1.23$)	F(1,38)=4.06 p=0.052
awareness	3.93 ($\sigma = 1.16$)	3.20 ($\sigma = 1.23$)	F(1,38)=3.98 p=0.010
co-presence	4.50 ($\sigma = 0.73$)	3.91 ($\sigma = 1.24$)	F(1,38)=4.30 p=0.166

Table 2: *Mean differences regarding usability and preference for the two interaction techniques.*

item	prop	virtual	ANOVA
usability	4.03 ($\sigma = 0.90$)	2.57 ($\sigma = 1.23$)	F(1,38)=3.91 p=0.001
preference	93.75%	6.25%	

The distance between the centers of the handles on the windshield prop was 93 cm. Table 3 shows that while using the prop-based method, the distance of the user's hands on the handles varied by 2.26 cm (1.13 cm per user), which may be due to slight hand adjustments without dropping the prop. Using the virtual method, this distance varied by 26.34 cm (13.17 cm per user). The average task completion time (TCT) was 17.16 sec for prop-based and 42.08 sec for the virtual condition. A closer look at individual TCTs revealed that some of the participants were repeatedly able to reach times comparable to the prop manipulation. As previously mentioned, the intended TCT given by the CAD process animation is 20 seconds; therefore, so the prop method's TCT matches very well. Some participants were even faster because they started to get competitive and did not have to care about damages to the car. The main differences in TCT and hand positions were found in the fine adjustment phase close to the target position of the windshield. We assume that the participants needed a lot more time without the prop due to the missing force feedback, which could not be sufficiently compensated for by verbal communication and visual feedback. In general we observed that the participants talked more to each other when using the virtual method to instruct their counterpart.

Table 3: *ADDH is a measure for the average deviation from the default distance of user's hands for each technique and corresponding task completion times.*

	prop	virtual
ADDH	2.26 cm ($\sigma = 0.86$)	26.34 cm ($\sigma = 5.70$)
avg TCT	17.16 sec ($\sigma = 6.14$)	42.08 sec ($\sigma = 24.86$)

The participants, all of whom knew each other, synchronized their movements by talking to one another. Mostly at the beginning they counted "1, 2, 3" or similar commands to start the interaction. At the end during the fine adjustment of the windshield, which took the most time, they exchanged commands like "up, down, forward" to instruct their partner. They even realized if their partner was moving too fast between start and target position and asked him to slow down. This type of communication adds to the findings of Otto et al. [OR03] and Roberts et al. [RWO*04], who observed that in non co-located collaborative tasks, verbal communication is a key feature. Even though our setup is co-located, there is not much of a difference when compared to non co-located setups due to the use of HMDs, with direct voice communication and the windshield prop being the exceptions.

5. Discussion

Our initial experiment comparing a single-user projection system to a two-user setup showed that individually correct stereo perspectives greatly facilitate side-by-side interactions of two users such as pointing or using different tools. The effect of clipped objects at the boundaries of the viewing frustum can be partly compensated by one user controlling the scene movement, but the interaction space is clearly limited by the single projection screen. The interviewed experts confirmed that it is unnecessary or even confusing to have virtual hand representations since users can see the hand gestures of their counterpart and their own through the glasses. Virtual hand representations trigger the users to switch their focus between the real and the virtual hand, which disturbs the stereo perception. These findings are in contrast to our HMD-task wherein virtual representations of hands and heads are essential since the view on real body parts is blocked by the HMD.

The HMD-based assembly task was further discussed in a post evaluation. We showed pictures taken from the HMD-trials to ergonomists and discussed to what extent they thought that ergonomic issues could be evaluated with such virtual methods. They believed that taking snapshots of users during the interaction could be very helpful to evaluate the body postures at certain timesteps. Even the 13 cm deviation of the hand position per user during the unadulterated virtual interaction with the windshield could be tolerable. However, the long task completion time of the purely virtual scenario is ergonomically critical and would distort the results of the ergonomic evaluations. The task duration is a particularly important factor for the ergonomics of a task. It is investigated how long workers have to rest in static poses carrying certain weights and how often they have to repeat the complete task over one working day. We did not simulate correct weights in our virtual scenario, but when looking at different postures, our ergonomists are able to judge how much weight a worker should carry and for how long he should remain in a particular pose.

Usually a complete assembly process includes several steps including side-by-side as well as face-to-face collaboration. So it happens that two workers are placing a big part into the desired position and then one of them has to fix it with a tool such as a screwdriver while the other one holds the part in position. While the basic usability of the projection-based setup is much better than for the HMD setup, only the HMD setup can be used for almost all task sequences. Nevertheless, a two-user CAVE™ would already work for a much larger set of scenarios than the single projection wall used in this study, even though true face-to-face situations can be only simulated in a multi-user HMD system.

6. Conclusions and Future Work

We presented the evaluation of collaborative interaction techniques in a projection-based and a HMD-based stereoscopic two-viewer system for automotive application scenarios. We evaluated the techniques in virtualized real world assembly tasks involving two co-located collaborating users. Our studies confirmed that having a perspectively correct view of the virtual world is a basic requirement and a premise for collaborative work with virtual assembly simulations. Prop-based two-user interaction provides an interactive link between the two users and improves task performance and collaboration. Additionally the subjective ratings (see Table 2) confirmed that the participants preferred the prop-based over the pure virtual method. For day-to-day use with automotive experts, we recommend building a simple construction set for building props of the commonly manipulated objects. The passive haptic feedback and the direct connection between both participants make it an indispensable part of a variety of multi-user assembly tasks in virtual environments.

Our results show that it should be carefully considered which type of display technology is suitable for a given task. Projection-based multi-user setups are mostly suited for tasks executed in a side-by-side fashion due to mutual occlusion of the displayed graphics otherwise. The only way to perfectly support face-to-face situations with projection-based systems is by using a truly distributed setup consisting of two separate projection systems as in the study introduced by Heldal et al. [HSS*05]. HMD-based setups can be used for side-by-side and face-to-face situations, but in general they come with other shortcomings such as the discomfort of HMDs, limited FOV and an increased potential for motion sickness. Surprisingly in our HMD-scenario, the participants performed very well and were moving relatively fast. Even their sensation of presence and co-presence was quite high. Nevertheless wireless HMDs would be a great benefit to simulate more complicated assembly tasks wherein users have to cross over each other. HMD-based multi-user setups in combination with multi-user prop-based interaction turned out to be the most general and also a quite effective approach for co-located collaboration in automotive assembly scenarios.

Once users have realized the advantage of a two-user system over a single user system, they immediately suggested involving even more tracked users with individual perspectives into the virtual assembly scenario. Each user needs to be provided with an individual stereoscopic view as well as with a task-specific view. Thus further development of stereoscopic display technology supporting three or more users is important. However, these technical developments need to be accompanied by the corresponding development of effective and task-specific collaborative interaction techniques for multiple users. Multi-user props may therefore be the key to successful simulation of assembly tasks in virtual environments.

Acknowledgments

We would like to thank the participants of our studies for their time and comments, Manuel Gleisberg for his help with setting up the two-user projection system and the side-by-side evaluation, the team of the Volkswagen VR-Lab for valuable discussions and continuous support, and the reviewers for their detailed and constructive comments.

References

[ABF*97] AGRAWALA M., BEERS A. C., FROEHLICH B., HANRAHAN P., MCDOWALL I., BOLAS M.: The Two-User Responsive Workbench: Support for Collaboration Through Individual Views of a Shared Space. In *Proceedings of Computer Graphics 1997 (SIGGRAPH '97)* (1997), pp. 327–332. 2

[APT*98] ARTHUR K., PRESTON T., TAYLOR R. M., BROOKS F. P., WHITTON M. C., WRIGHT W. V.: *Designing and Building the PIT: a Head-Tracked Stereo Workspace for Two Users*. Tech. rep., University of North Carolina at Chapel Hill, 1998. 2

[BHG01] BIOCCA F., HARMS C., GREGG J.: The Networked Minds measure of social presence: Pilot test of the factor structure and concurrent validity. In *4th annual International Workshop on Presence* (2001), pp. 9–11. 5

[CFH97] CUTLER L. D., FRÖHLICH B., HANRAHAN P.: Two-handed direct manipulation on the responsive workbench. In *SI3D '97: Proceedings of the 1997 symposium on Interactive 3D graphics* (New York, NY, USA, 1997), ACM, pp. 107–114. 5

[DLT06] DUVAL T., LECUYER A., THOMAS S.: SkeweR: a 3D Interaction Technique for 2-User Collaborative Manipulation of Objects in Virtual Environments. In *IEEE Symposium on 3D User Interface 2006* (2006), pp. 71–74. 5

[FBS*05] FROEHLICH B., BLACH R., STEFANI O., HOCHSTRATE J., HOFFMANN J., KLUEGER K., BUES M.: Implementing Multi-Viewer Stereo Displays. In *WSCG (Full Papers)* (2005), pp. 139–146. 3

[FTB*00] FROEHLICH B., TRAMBEREND H., BEERS A., AGRAWALA M., BARAFF D.: Physically-based Manipulation on the Responsive Workbench. In *Proceedings of IEEE Virtual Reality 2000* (2000), pp. 5–11. 5

[GMH01] GERHARD M., MOORE D., HOBBS D.: Continuous Presence in Collaborative Virtual Environments: Towards the Evaluation of a Hybrid Avatar-Agent Model for User Representation. In *International Conference on Intelligent Virtual Agents* (2001), pp. 137–155. 5

[HB03] HOFMANN J., BUBB H.: Presence in industrial virtual environment applications - susceptibility and measurement reliability. *Being There: Concepts, Effects and Measurement of User Presence in Synthetic Environments* (2003), 237–246. 5

[HPGK94] HINCKLEY K., PAUSCH R., GOBLE J. C., KASSELL N. F.: Passive real-world interface props for neurosurgical visualization. In *Proceedings of ACM CHI'94 Conference on Human Factors in Computing Systems* (1994), pp. 452–458. 2

[HSS*05] HELDAL I., SCHROEDER R., STEED A., AXELSSON A.-S., SPANTE M., WIDESTRÖM J.: Immersiveness and Symmetry in Copresent Scenarios. In *IEEE Virtual Reality 2005* (2005), pp. 171–178. 7

[MAP99] MARGERY D. M., ARNALDI B., PLOUZEAU N.: A general framework for cooperative manipulation in virtual environments. In *Proceedings of EGVE 99: Eurographics Workshop on Virtual Environments* (1999), Springer, pp. 169–178. 2

[OR03] OTTO O., ROBERTS D.: Importance of Communication Influences on a Highly Collaborative Task. In *Proceedings of the 7th IEEE International Symposium on Distributed Simulation and Real Time Applications* (2003), pp. 195–201. 7

[PBF02] PINHO M. S., BOWMAN D. A., FREITAS C. M.: Cooperative Object Manipualtion in Immersive Virtual Environments. In *VRST 2002* (2002). 5

[PRS02] PREECE J. J., ROGERS Y., SHARP H.: *Interaction Design: beyond human computer interaction*. Wiley & Sons, 2002. 3

[RSJ02] RUDDLE R. A., SAVAGE J. C., JONES D. M.: Symmetric and Asymmetric Action Integration During Cooperative Object Manipulation in Virtual Environments. *ACM Trans. on Computer-Human Interaction 9*, 4 (2002), 285–308. 2, 5

[RWO*04] ROBERTS D., WOLFF R., OTTO O., KRANZLMUELLER D., ANTHES C., STEED A.: Supporting Social Human Communication between Distributed Walk-in Displays. In *Proceedings of VRST'04* (2004), pp. 81–88. 7

[SFH*02] SCHMALSTIEG D., FUHRMANN A., HESINA G., SZALAVARI Z., ENCARNACAO L. M., GERVAUTZ M., PURGATHOFER W.: The Studierstube Augmented Reality Project. In *Presence Teleoperators and Virtual Environments* (2002), vol. 11, pp. 33–54. 2

[SMF09] SALZMANN H., MOEHRING M., FROEHLICH B.: Virtual vs. Real-World Pointing in Two-User Scenarios. In *IEEE Virtual Reality Conference* (2009), pp. 127–130. 2

[SSA*01] SCHROEDER R., STEED A., AXELSSON A.-S., HELDAL I., ABELIN A., WIDESTRÖM J., NILSSON A., SLATER M.: Collaborating in networked immersive spaces: as good as being there together? *Computers and Graphics 25* (2001), 781–788. 5