Simplifying Collaboration in Co-Located Virtual Environments Using the Active-Passive Approach

Jose Garcia Estrada [*]   Jan P. Springer [†]   Helen Wright [‡]

University of Hull

ABSTRACT

The design and implementation of co-located immersive virtual environments with equal interaction possibilities for all participants is a complex topic. The main problem, on a fundamental technical level, is the difficulty of providing perspective-correct images for each participant. There is consensus that the lack of a correct perspective view will negatively affect interaction fidelity and therefore also collaboration. Several research approaches focus on providing a correct perspective view to all participants to enable co-located work. However, these approaches are usually either based on custom hardware solutions that limit the number of users with a correct perspective view or software solutions striving to eliminate or mitigate restrictions with custom image-generation approaches.

In this paper we investigate an often overlooked approach to enable collaboration for multiple users in an immersive virtual environment designed for a single user. The approach provides one (active) user with a perspective-correct view while other (passive) users receive visual cues that are not perspective-correct. We used this active-passive approach to investigate the limitations posed by assigning the viewpoint to only one user. The findings of our study, though inconclusive, revealed two curiosities. First, our results suggest that the location of target geometry is an important factor to consider for designing interaction, expanding on prior work that has studied only the relation between user positions. Secondly, there seems to be only a low cost involved in accepting the limitation of providing perspective-correct images to a single user, when comparing with a baseline, during a coordinated work approach. These findings advance our understanding of collaboration in co-located virtual environments and suggest an approach to simplify co-located collaboration.

Keywords: Immersive Virtual Environments, Co-Located Work, Collaboration, Interaction Design

Index Terms: H.5.1 [Multimedia Information Systems]: Artificial, augmented, and virtual realities—; H.5.3 [Group and Organization Interfaces];

1 INTRODUCTION

Support for collaboration has been the object of interest very early on in the development of immersive virtual environments (VE) and displays [5]. Collaboration in VEs is often focused on cases where users do not share a physical location [2], such as distributed VEs. The case where several users collaborate within a physical location is largely under-investigated, in comparison with distributed VEs. However, large projection-based display systems are now in widespread use and allow for easy support of side-by-side interaction similar to meetings in the real world. Unfortunately, most VEs

---

[*] `mailto:j.f.garcia-estrada@2010.hull.ac.uk`
[†] `mailto:j.springer@hull.ac.uk`
[‡] `mailto:h.wright@hull.ac.uk`

used in these displays are designed for a single user, i. e. they usually only provide a single perspective-correct view.

Support for perspective-correct views for multiple users in immersive VEs is essentially a bandwidth problem. For each user two images need to be computed, one for the left and one for the right eye. Usually, these images are generated at a rate of 60 Hz per eye or 120 Hz for the whole frame. For two users this rate needs to be doubled, for three users tripled and so on. This means two-user stereo requires 240 Hz, three-user stereo 360 Hz *frame rate* assuming a single output device. However, current display technology (e. g., digital displays, light-emitting projectors) is rather restricted in how much bandwidth can be processed with a single output device. Usually, such devices are capable of providing a display-refresh rate at 60 Hz and, more recently for consumer devices, 120 Hz to support a pair of stereo images at 60 Hz within one video frame. Past attempts to lift this restriction involved separating left/right images for a single user with polarization filters and multiple users with time-multiplexed shuttering [3]. An alternative is to synchronize the color wheels of multiple DLP projectors to show time-multiplexed images per eye per user and distribute the primary color channels over multiple devices [8]. Because the mirror elements in a DLP can be switched at frequencies in the order of 10,000 Hz or higher, potentially a large amount of left/right-eye image pairs could be processed [13]. However, this is still very experimental technology that would also require substantial changes to the image-generation process to provide the necessary amount of images in the (very short) time frame required.

An alternative approach for supporting several co-located users working on a shared display is using a standard VE, assigning a correct perspective to one user, and accepting that other users will be limited to a single perspective-correct view. Such an approach, identified as active-passive [4] or leader-follower [10], has been previously researched but not quantified for direct manipulation tasks nor its usability and user acceptance determined. A question is how can we quantify that limitation. An answer to this question will provide insight into the implications of adopting an active-passive approach and inform the design process of VEs.

We report on a study that investigates collaboration in co-located work. We accept the limitation posed by assigning a correct perspective view to only one user and study the effects of interaction performance. The study uses a spatial-ability task involving coordinated selection and manipulation of simple target geometry. The findings of our study quantify the overhead on task performance during coordinated work when adopting the active-passive approach. These findings advance our understanding of collaboration in co-located virtual environments and suggest an approach to simplify co-located collaboration. The results of our study suggest that the location of target geometry is an important factor to consider for designing interaction in VEs, expanding existing approaches that emphasize only the relation between user positions.

2 RELATED WORK

Providing a correct perspective for all co-located users can be considered as an obvious requirement for collaboration in a co-located VE. The difficulty to implement such a solution, however, has led to the exploration of other approaches. Projection of multiple images on a shared display, combined with mechanical or electronic

hardware, constitutes the main approach to provide correct perspective. Time-multiplexing and polarization approaches introduced in [1] were extended and improved in [6], finally resulting in a six-user co-located VE [8]. Another approach projected several images on the same display and used physical barriers to separate correct perspective [7].

Simon [12] uses a single image, produced by averaging the viewpoints of several co-located users, for multi-user interaction using a curved screen. This approach, however, averages the viewpoints so no user benefits from a correct perspective but settles for an averaged one. Other approaches focus on rendering images based on where the user is looking or the portion of the scene visible to a particular viewer. Marbach [9] renders the scene using each user's head position and gaze direction to compute this viewer's perspective. The resulting images are blended when users look at different portions of the screen or averaged when they look at an overlapping area. This approach combines co-location using an averaged view with a separation approach when users look at different parts of the screen. A similar approach, tested with two users, renders images to different sections of a large curved screen using separate perspectives for each user [11]. For two users, this produces the impression almost of two individual displays surfaces using a mediated viewpoint for overlapping areas.

Despite considerable existing work on collaboration in VEs, few researchers have characterized and quantified the limitations posed by co-locating multiple users in a single-user VE. Using the position of one user to compute a perspective-correct view in the VE that user will always be the center of projection. Additional co-located users will receive the cues intended for that primary user [10]. Assigning the viewpoint to one user whilst co-located users receive cues intended for that user is known as leader-follower approach [10] or active-passive approach [4].

A study of depth judgment in VEs found differences in performance based on the position of the passive user in relation to the location of the active user [10]. Visual cues were presented for the active user on an L-shaped display. The passive user verbally agreed to changes in location with the active user during the task. The study found significant differences in performance based on the location of the users in the VE and in relation to the front display surface. An interesting outcome of the study is in establishing that the active user can assist the passive user to improve performance. The authors explained their results as the effect on the position of the passive user by the position of the active user. However, the verbal communication component of their task makes it difficult to assess a similar effect on tasks involving direct interaction.

Positioning the passive user according to the location of the active user to facilitate co-located work was proposed based on empirical data [10] as well as through theoretical considerations [4] as a design guideline for co-located work using the active-passive approach. However, they only address co-located work in the context of verbal communication but not for tasks based on direct manipulation of spatial entities. We are going to shed some light on this by describing a user study and its results in the rest of this paper.

3 EXPERIMENT

We conducted a controlled experiment to study the performance of subjects carrying out a spatial ability task requiring coordinated work. The task design used the active-passive approach and introduced a simple navigation tool to help the active user to assist the passive user in mitigating non-optimal perspective-correct views (figure 1). We assume that the active user focuses on their task and is not easily aware of the actions of the passive user. The navigation tool was designed to facilitate action coordination between the active and the passive user by making explicit to the active user what objects in the scene are within the view frustum. The target geometry in the scene is replicated in the navigation tool. The target geometry's replica is highlighted when that geometry is outside the frustum. This visual feedback was intended to alert the active
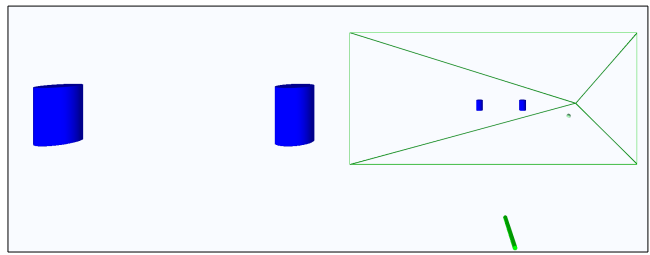

Figure 1: Screen shot of the VE used in our experiment. The target geometry in the VE, i. e. the two cylinders on the left, is replicated in scale in the navigation tool on the right, which is an inset in the final rendered image. It is displayed at the tip of the wand used for picking the target geometry. The navigation tool also indicates the active user's viewpoint with respect to the physical screen setup by drawing the edges of the active user's view frustum.

user that the passive user is unable to see their corresponding target geometry.

We made a pragmatic decision to carry out the experiment with two co-located users. In a preliminary experimental setup we used a pair of users in a bottom-up approach for investigating the active-passive approach. However, this exhibited difficulties in differentiating individual contributions to the common task as well as the statistical complexity of using two different individuals, which would result in measuring the combination of the contributions by the individuals. We solved this problem by using a single individual for each session paired with an automated participant played by the computer. We call this automated participant a *virtual subject*. The virtual subject was modeled based on observations of real subjects during preliminary work and performed the role of the passive user as well as the active user. As a passive user it simply calculates and carries out its interaction with the visible target geometry. In active user mode the virtual subject uses recorded trajectories designed to assist the passive user and calculates its interaction with the target geometry on the fly. The virtual subject was designed to perform in a uniform fashion, providing all participants with a companion at a consistent skill and expertise level.

3.1 Hypotheses

Viewpoint assignment to only one user, i. e. the active user, introduces the possibility that the other user, i. e. the passive user who lacks a correct perspective, will be unable to see the target geometry. In that case the passive user must wait until that geometry is visible again to interact with it.

We hypothesize that the provision of a navigation tool for the active user allows a significant reduction in waiting time for the passive user.

The main dependent variable in our experiment is the time the passive user has to wait until the target geometry is visible again to continue with the task. This variable is referred to as waiting time T_{wait}. The second dependent variable is the overall time required to complete the task. This variable is labeled as task completion time T_{total}. The independent variable is the availability of the navigation tool (C_{tool}) or the lack of it ($C_{no-tool}$).

3.2 Participants

Fourteen unpaid subjects, age 20 to 56 (average 26.85), participated in the study. The subject group consisted of nine males (64.29%) and five females (35.71%). Participants filled out a pre-test questionnaire. All subjects reported normal to corrected-to-normal vision. 43% of the subjects reported previous exposure to 3d-computer games or immersive VEs.

3.3 Apparatus

The experiment was carried out on a large projection-based display system. The display consisted of a rear-projected screen measur-

ing 5.32×2.44 m using two Christie Mirage 2000 DLP projectors producing an overall image resolution of 2240×1024 pixels. The two projectors were edge-blended and created a seamless single image on the screen. A Vicon MX tracking system was used to track the subjects and interaction tool. Stereoscopic viewing was implemented using tracked active shutter glasses. Participants used a tracked hand-held pointing device for selection and direct manipulation. A tripod, located equidistantly to the real subject's starting position, was used to mark the virtual subject's starting position. The human subject was always located to the right of the virtual subject in the tracking volume and in front of the screen. Once the task started the human subject was allowed to move freely and interact with the target geometry.

3.4 Task

Subjects start the trial from a fixed initial position 1.2 m from the origin in the VE (as well as in the physical display system). The task is designed to elicit interaction and coordination as well as be simple and easily executable by subjects from any background. Subjects carry out the task in coordination with the virtual subject. Coordinated actions were required in order to progress with the task. The task starts with the target geometry displayed in front of the subjects in negative parallax for the active user. Apart from using color as visual cues, shadows were used to provide an additional depth cue when users displaced the geometry (figure 2a). Subjects are required to approach the target geometry for selection (figure 2b). The target geometry is highlighted when selected, providing visual feedback to the user (figure 2c). When the target geometry is simultaneously selected by both the active and passive user (figure 2d) the target geometry becomes movable. The target geometry can now be dragged to a drop-off location indicated by a white circle (figure 2e). This requires the subject to drag the object and hold it over the drop-off location. When the target geometry is correctly placed by both users and simultaneously kept there the trial is completed (figure 2f). The actions of the virtual subject are produced based on pre-recorded movements and the virtual subject does not react to the human subject's actions. It only acts upon triggered events e. g., visibility of the target geometry. Each trial starts using one of four different initial positions (P1, P2, P3, and P4), counterbalanced throughout the experiment. The four positions were designed to place the passive user's geometry at the boundary of the rendering frustum (figure 3).

3.5 Experimental design

The active-passive approach in combination with the navigation tool resulted in a 2×2 within-subjects factorial design. The availability of the navigation tool or the lack of it was the within-subjects factor, which included two different experimental conditions: subject has control of a correct perspective, i. e. active user, or lacks control of it, i. e. passive user. Control of the viewpoint was randomly assigned to and balanced between the subject's positions per trial. Subjects received no indication whether they were the active or the passive user. The experiment tested for differences in time based on the ability to keep the target geometry visible with or without the aid of the navigation tool.

Figure 3 shows how the target geometry was arranged at the beginning of the trial. Each trial began by assigning a position for the target geometry to the active user from a set of predetermined positions, where both the human and the virtual subject had their own set of four positions (figure 3a and figure 3b). The target-geometry position assigned to the active user was used to calculate where to place the target geometry for the passive user. The outcome of the calculation was a coordinate on the X axis that placed the passive user's target geometry at the edge of the viewing frustum.

3.6 Procedure

The procedure consisted of three steps. First, participants were briefed on the experiment, provided with consent forms, and completed a screening questionnaire. The interpupillary distance was

measured for each subject and used to configure the system before each experimental session. After filling out the questionnaire the task was explained to subjects using a mock-up of the object geometry used in the experiment. The second step introduced the subjects to the virtual environment. Subjects were taught how to use the wand and made aware of the differences between having objects in the scene rendered from their perspective or from the virtual subject's perspective. Subjects carried out training consisting of 8 trials. Training was the same for all subjects regardless of their level of familiarity with VEs. The subjects were instructed to be as efficient as possible in performing the task. The participants performed 80 trials altogether being randomly assigned one of the four test conditions every time, i. e. active vs. passive user and navigation tool or not. Finally, the subjects were asked to complete a post-test questionnaire.

3.7 Results

Quantitative data collection focused on two measurements. First, we measured the time the passive user waited for assistance from the active user and was thus not able to progress with the task (T_{wait}). Second, we recorded the overall time to complete the task (T_{total}). All data was collected in log files generated by the experiment software. The collected quantitative data contained several outliers, identified using visual inspection of box plots and confirmed using Tukey's hinges. Outliers beyond three times the interquartile range from the rest of the scores were removed before analysis. Out of 1120 measurements 80 data points were removed as outliers accounting for 7.14% of all data. After outlier removal the remaining measurements were averaged to generate performance metrics for each participant. Outlier removal did not result in removing any of the conditions or eliminating any participants.

The data was separated into four groups with respect to the predetermined starting positions of the target geometry (P1, P2, P3, and P4, as described in section 3.5 and shown in figure 3). Shapiro-Wilk test for normality and normal Q-Q plots showed that the data was not normally distributed. Therefore, statistical analysis was performed using non-parametric tests for all four selection positions, specifically Friedman and Wilcoxon signed-rank tests were used. Waiting time T_{wait} as well as completion time T_{total} were compared for the C_{tool} and $C_{no-tool}$ conditions, respectively. In the following we will provide the results of the statistical analysis broken down for user condition per starting position of the target geometry.

Human Subject as Active User and Position P1 The initial position of the target geometry was located 0.1 m to the left and 0.9 m in front of the subject's starting position. T_{wait} for the C_{tool} condition varies between 4.63% and 15.54% of T_{total} while T_{wait} for the $C_{no-tool}$ condition varies between 3.7% and 19.5% of T_{total}. The Wilcoxon test applied to the T_{wait} variable shows no significant difference between the C_{tool} and $C_{no-tool}$ condition ($z = -0.659$, $p = 0.510$). T_{wait} median value for the C_{tool} condition was $m = 2.52$ and $m = 3.31$ for the $C_{no-tool}$ condition. The Wilcoxon test applied to the T_{total} variable shows no significant difference for the C_{tool} condition compared with $C_{no-tool}$ condition ($z = -.722$, $p = 0.470$). The T_{total} median for the C_{tool} condition was $m = 20.84$ and $m = 20.84$ for the $C_{no-tool}$ condition. A diagram overlaying both variables per subject is shown in figure 4.

Human Subject as Active User and Position P2 The initial position of the target geometry was located 0.2 m to the left and 1 m in front of the subject's starting position. T_{wait} for the C_{tool} condition varies between 3.92% and 25.89% of T_{total} while T_{wait} for the $C_{no-tool}$ condition varies between 4.38% and 21.97% of T_{total}. The Wilcoxon test applied to the T_{wait} variable shows significant decrease for nine participants in the $C_{no-tool}$ condition while five participants performed better in the C_{tool} condition ($z = -2.731$, $p = 0.006$). T_{wait} median value for the C_{tool} condition was $m = 3.04$ and $m = 2.09$ for the $C_{no-tool}$ condition. The Wilcoxon test applied to the T_{total} variable shows no significant difference for the C_{tool} condition compared with $C_{no-tool}$ condition ($z = -1.036$, $p = $

(a) Target geometry at start of the trial.

(b) Approach to target geometry.

(c) Individual selection.

(d) Simultaneous selection of target geometry.

(e) Approach to the target area.

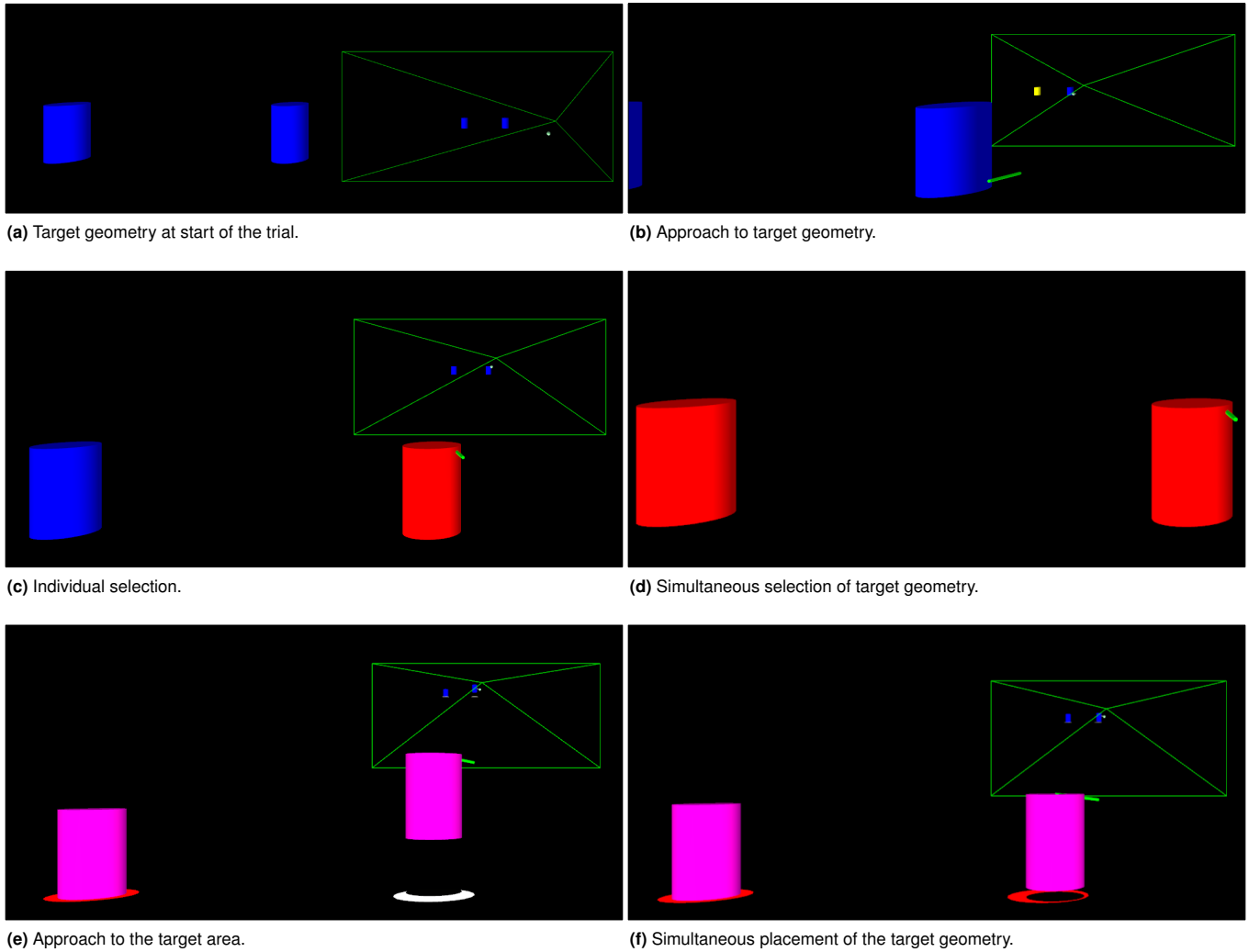**(f)** Simultaneous placement of the target geometry.

Figure 2: Sequence of screen shots of the task performed in our experiment. (a) – (b) At the start of the trial the target geometry is placed in front of the subject in negative parallax and needs to be approached for selection. (c) – (d) The subject selects the target geometry in coordination with the virtual subject. (e) – (f) During the docking phase the subject uses drag and drop to place the target geometry on the indicated area.

(a) Geometry positions for the human subject as the active user.

(b) Geometry positions for the human subject as the passive user.
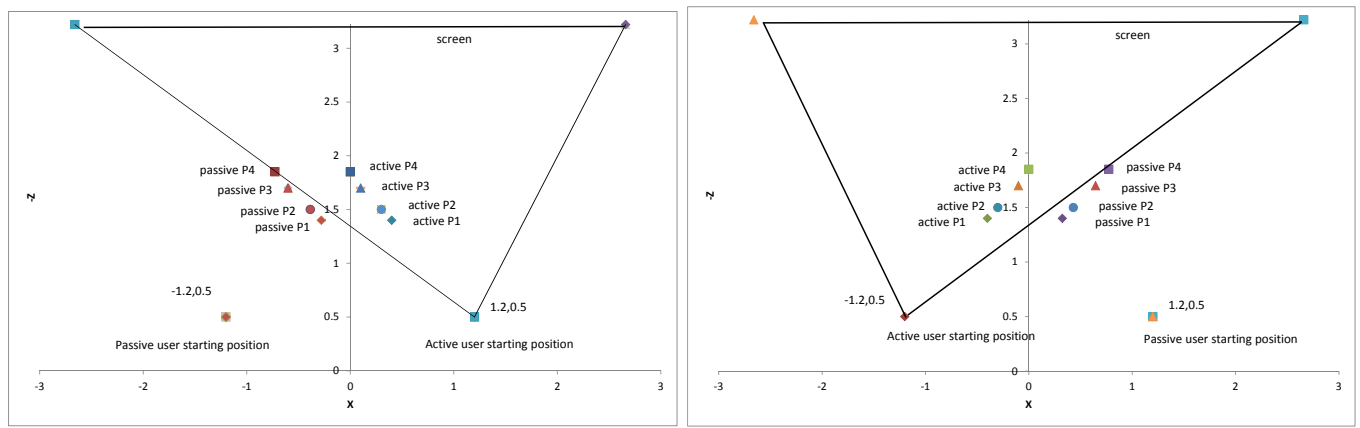
Figure 3: Arrangement of the positions for target geometry to be selected for (a) the human subject as the active user and (b) the human subject as the passive user. Note that for the virtual subject (a) and (b) are reversed.
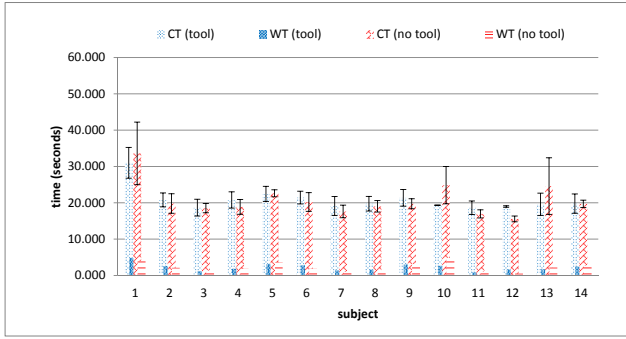
Figure 4: Human subject as active user and target geometry in position P1. For C_{tool} condition T_{wait} values represent 4.63% of T_{total} for subject 11 to 15.54% of T_{total} for subject 1. For $C_{no-tool}$ condition T_{wait} values represent 3.7% of T_{total} for subject 12 to 19.5% of T_{total} for subject 10. Note that T_{wait} is overlaid onto T_{total} rather than stacked. Error bars are standard deviation.
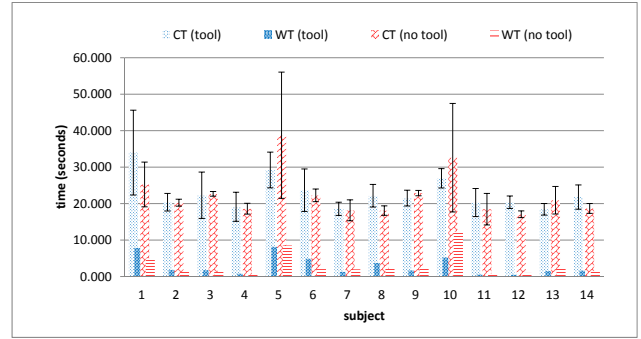

Figure 6: Human subject as active user and target geometry in position P3. For C_{tool} condition T_{wait} values represent 2.8% of T_{total} for subject 12 to 28.0% of T_{total} for subject 5 for the C_{tool} condition. For C_{tool} condition T_{wait} values represent 3.2% of T_{total} for subject 4 to 37.8% of T_{total} for subject 10 for the $C_{no-tool}$ condition. Note that T_{wait} is overlaid onto T_{total} rather than stacked. Error bars are standard deviation.
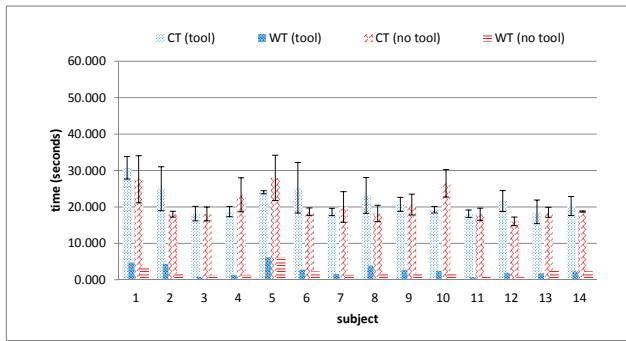

Figure 5: Human subject as active user and target geometry in position P2. For C_{tool} condition T_{wait} values represent 3.92% of T_{total} for subject 11 to 25.89% of T_{total} for subject 5 for the C_{tool} condition. For C_{tool} condition T_{wait} values represent 4.38% of T_{total} for subject 12 to 21.97% of T_{total} for subject 5 for the $C_{no-tool}$ condition. Note that T_{wait} is overlaid onto T_{total} rather than stacked. Error bars are standard deviation.
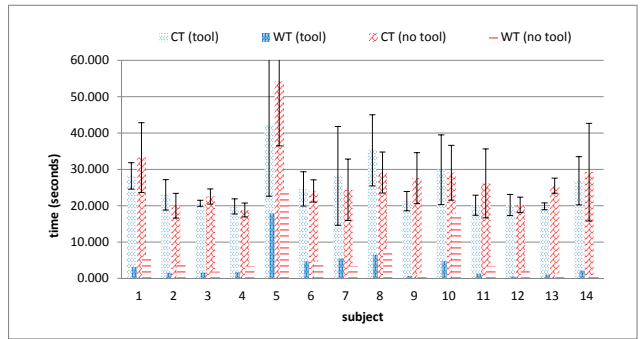

Figure 7: Human subject as active user and target geometry in position P4. For C_{tool} condition T_{wait} values represent 2.5% of T_{total} for subject 12 to 42.0% of T_{total} for subject 5 for the C_{tool} condition. For C_{tool} condition T_{wait} values represent 2.67% of T_{total} for subject 14 to 61.20% of T_{total} for subject 10 for the $C_{no-tool}$ condition. Note that T_{wait} is overlaid onto T_{total} rather than stacked. Error bars are standard deviation.

0.300). The T_{total} median for the C_{tool} condition was $m = 21.59$ and $m = 20.74$ for the $C_{no-tool}$ condition. A diagram overlaying both variables per subject is shown in figure 5.

Human Subject as Active User and Position P3 The initial position of the target geometry was located 0.4 m to the left and 1.2 m in front of the subject's starting position. T_{wait} for the C_{tool} condition varies between 2.8% and 28.0% of T_{total} while T_{wait} for the $C_{no-tool}$ condition varies between 3.2% and 37.8% of T_{total}. The Wilcoxon test applied to the T_{wait} variable shows no significant difference between the C_{tool} and $C_{no-tool}$ condition ($z = -0.094$, $p = 0.925$). T_{wait} median value for the C_{tool} condition was $m = 3.14$ and $m = 3.87$ for the $C_{no-tool}$ condition. The Wilcoxon test applied to the T_{total} variable shows no significant difference for the C_{tool} condition compared with $C_{no-tool}$ condition ($z = -.722$, $p = 0.470$). The T_{total} median for the C_{tool} condition was $m = 22.78$ and $m = 22.48$ for the $C_{no-tool}$ condition. A diagram overlaying both variables for these results per subject is shown in figure 6.

Human Subject as Active User and Position P4 The initial position of the target geometry was located 0.5 m to the left and 1.4 m in front of the subject's starting position. T_{wait} for the C_{tool} condition varies between 2.5% and 42.0% of T_{total} while T_{wait} for the $C_{no-tool}$ condition varies between 2.67% and 61.20% of T_{total}. The Wilcoxon test applied to the T_{wait} variable shows significant median decrease for eleven participants in the C_{tool} condition while three participants performed better in the $C_{no-tool}$ condition

($z = -2.668$, $p = 0.008$). T_{wait} median value for the C_{tool} condition was $m = 4.04$ and $m = 6.50$ for the $C_{no-tool}$ condition. The Wilcoxon test applied to the T_{total} variable shows no significant difference for the C_{tool} condition compared with $C_{no-tool}$ condition ($z = -1.036$, $p = 0.300$). The T_{total} median for the C_{tool} condition was $m = 25.73$ and $m = 27.44$ for the $C_{no-tool}$ condition. A diagram overlaying both variables for these results per subject is shown in figure 7.

Virtual Subject as Active User and Position P1 The initial position of the target geometry was located 0.1 m to the right and 0.9 m in front of the virtual subject's starting position. T_{wait} for the C_{tool} condition varies between 13.3% and 23.6% of T_{total} while T_{wait} for the $C_{no-tool}$ condition varies between 13.5% and 19.6% of T_{total}. The Wilcoxon test applied to the T_{wait} variable shows no significant difference between the C_{tool} and $C_{no-tool}$ condition ($z = -1.338$, $p = 0.124$). T_{wait} median value for the C_{tool} condition was $m = 3.82$ and $m = 3.85$ for the $C_{no-tool}$ condition. The Wilcoxon test applied to the T_{total} variable shows no significant difference for the C_{tool} condition compared with $C_{no-tool}$ condition ($z = -1.66$, $p = 0.096$). The T_{total} median for the C_{tool} condition was $m = 22.53$ and $m = 23.8$ for the $C_{no-tool}$ condition. A diagram overlaying both variables results per subject is shown in figure 8.

Virtual Subject as Active User and Position P2 The initial position of the target geometry was located 0.2 m to the right and 1 m in front of the virtual subject's starting position. T_{wait} for the C_{tool}
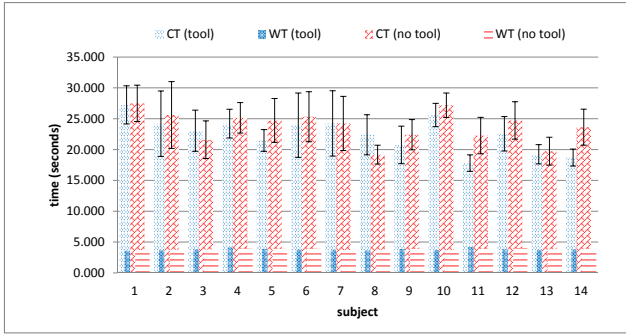
Figure 8: Virtual subject as active user and target geometry in position P1. For C_{tool} condition T_{wait} values represent 13.3% of T_{total} for subject 1 to 23.6% of T_{total} for subject 11 for the C_{tool} condition. For C_{tool} condition T_{wait} values represent 13.5% of T_{total} for subject 1 to 19.6% of T_{total} for subject 13 for the $C_{no-tool}$ condition. Note that T_{wait} is overlaid onto T_{total} rather than stacked. Error bars are standard deviation.

Figure 10: Virtual subject as active user and target geometry in position P3. For C_{tool} condition T_{wait} values represent 11% of T_{total} for subject 1 to 17.8% of T_{total} for subject 11 for the C_{tool} condition. For C_{tool} condition T_{wait} values represent 11.3% of T_{total} for subject 1 to 18.7% of T_{total} for subject 3 for the $C_{no-tool}$ condition. Note that T_{wait} is overlaid onto T_{total} rather than stacked. Error bars are standard deviation.
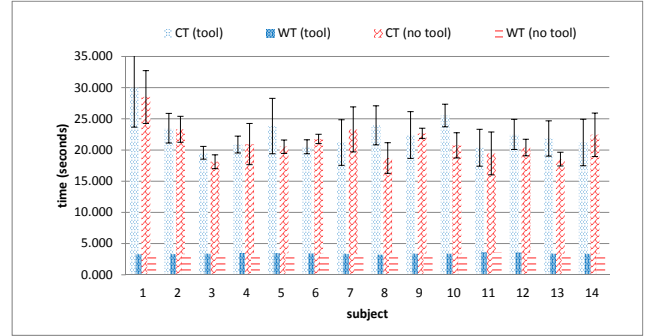

Figure 9: Virtual subject as active user and target geometry in position P2. For C_{tool} condition T_{wait} values represent 11.7% of T_{total} for subject 1 to 20.2% of T_{total} for subject 11 for the C_{tool} condition. For C_{tool} condition T_{wait} values represent 12.6% of T_{total} for subject 1 to 22.8% of T_{total} for subject 11 for the $C_{no-tool}$ condition. Note that T_{wait} is overlaid onto T_{total} rather than stacked. Error bars are standard deviation.
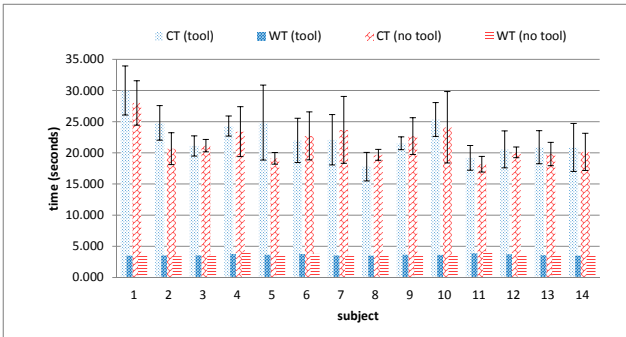

Figure 11: Virtual subject as active user and target geometry in position P4. For C_{tool} condition T_{wait} values represent 10.2% of T_{total} for subject 1 to 15.1% of T_{total} for subject 9 for the C_{tool} condition. For C_{tool} condition T_{wait} values represent 10.8% of T_{total} for subject 1 to 14.6% of T_{total} for subject 6 for the $C_{no-tool}$ condition. Note that T_{wait} is overlaid onto T_{total} rather than stacked. Error bars are standard deviation.
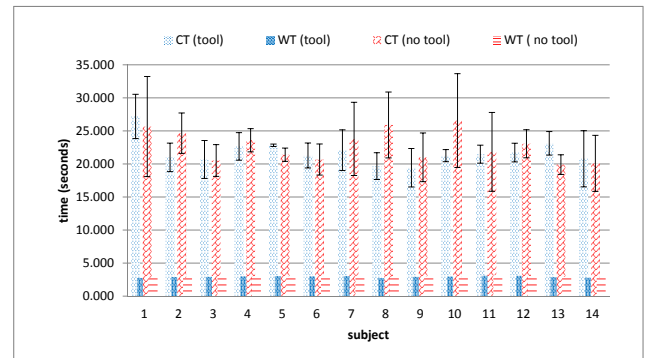
condition varies between 11.7% and 20.2% of T_{total} while T_{wait} for the $C_{no-tool}$ condition varies between 12.6% and 22.8% of T_{total}. The Wilcoxon test applied to the T_{wait} variable shows no significant difference between the C_{tool} and $C_{no-tool}$ condition ($z = -0.659$, $p = 0.510$). T_{wait} median value for the C_{tool} condition was $m = 3.65$ and $m = 3.66$ for the $C_{no-tool}$ condition. The Wilcoxon test applied to the T_{total} variable shows no significant difference for the C_{tool} condition compared with $C_{no-tool}$ condition ($z = -0.09$, $p = 0.925$). The T_{total} median for the C_{tool} condition was $m = 22.52$ and $m = 21.67$ for the $C_{no-tool}$ condition. A diagram overlaying both variables results per subject is shown in figure 9.

Virtual Subject as Active User and Position P3 The initial position of the target geometry was located 0.4 m to the right and 1.2 m in front of the virtual subject's starting position. T_{wait} for the C_{tool} condition varies between 11% and 17.8% of T_{total} while T_{wait} for the $C_{no-tool}$ condition varies between 11.3% and 18.7% of T_{total}. The Wilcoxon test applied to the T_{wait} variable shows no significant difference between the C_{tool} and $C_{no-tool}$ condition ($z = -0.71$, $p = 0.48$). T_{wait} median value for the C_{tool} condition was $m = 3.38$ and $m = 3.36$ for the $C_{no-tool}$ condition. The Wilcoxon test applied to the T_{total} variable shows no significant difference for the C_{tool} condition compared with $C_{no-tool}$ condition ($z = -1.04$, $p = 0.3$). The T_{total} median for the C_{tool} condition was $m = 22.67$ and $m = 21.40$ for the $C_{no-tool}$ condition. A diagram overlaying both variables results per subject is shown in figure 10.

Virtual Subject as Active User and Position P4 The initial position of the target geometry was located 0.5 m to the right and 1.4 m in front of the virtual subject's starting position. T_{wait} for the C_{tool} condition varies between 10.2% and 15.1% of T_{total} while T_{wait} for the $C_{no-tool}$ condition varies between 10.8% and 14.6% of T_{total}. The Wilcoxon test applied to the T_{wait} variable shows no significant difference between the C_{tool} and $C_{no-tool}$ condition ($z = -0.91$, $p = 0.363$). T_{wait} median value for the C_{tool} condition was $m = 2.94$ and $m = 2.95$ for the $C_{no-tool}$ condition. The Wilcoxon test applied to the T_{total} variable shows no significant difference for the C_{tool} condition compared with $C_{no-tool}$ condition ($z = -1.66$, $p = 0.096$). The T_{total} median for the C_{tool} condition was $m = 21.8$ and $m = 22.76$ for the $C_{no-tool}$ condition. A diagram overlaying both variables results per subject is shown in figure 11.

Preliminary Summary In summary, the variable T_{wait} showed significant difference for the $C_{no-tool}$ condition for position P2 and for the C_{tool} condition for position P4 when the human subject was leading the collaborative interaction. There was no significant difference for the T_{total} variable for any of the tested positions. As expected there was no significant difference between any of the variables when the virtual subject was leading the collaborative interaction. An overview of waiting time as a proportion of completion time per starting position of the target geometry with respect to the navigation-tool condition is given in table 1.

	Tool				No Tool			
	P1	P2	P3	P4	P1	P2	P3	P4
min %	4.6	3.9	2.8	2.5	3.7	4.4	3.2	2.7
max %	15.5	25.9	28.0	42.0	19.5	22.0	37.8	61.2
avg %	10.6	11.8	11.9	12.4	10.5	9.6	11.9	19.0

Table 1: Waiting time as a proportion of completion time per starting position of the target geometry with respect to the navigation-tool condition for the human subject.

4 DISCUSSION

In contrast to solutions based on time-multiplexing and polarization [1, 6, 8, 7] our approach computes only one perspective image at a time. Furthermore, rather than use an average perspective to accommodate several users [12, 9], or divide the screen into individual displays [11], we decided to use the active-passive [4] or leader-follower [10] approach. Participants carried out the task with a virtual subject to guarantee that all of them performed under the same conditions. In an attempt to mitigate the lack of correct perspective for the passive user, we provided a navigation tool to the active user which showed when the passive user was unable to proceed with the task due to hidden geometry. We measured the length of time the target geometry was not visible (T_{wait}), and the overall time taken to complete the task (T_{total}). We expected a significant difference between the C_{tool} and the $C_{no\text{-}tool}$ condition on account of the assistance the tool affords the active user when helping the passive user. In fact, there were no significant differences in performance between the C_{tool} and $C_{no\text{-}tool}$ conditions. However, our study did reveal two curiosities, firstly with respect to the waiting time (T_{wait}) and secondly the effect on task performance, as indicated by the overall task-completion time (T_{total}), of the participants' positions and scene geometry.

The ideal condition for interaction in a co-located VE is when all users have a correct perspective. Using a virtual subject allowed us to observe this condition when the human subject is the active user because the virtual subject does not need a perspective-correct view at all. This can be thought of as a baseline. Compared to this baseline table 1 shows that, on average, T_{wait} is 20% or less of the total time it took the human subjects to complete the task. This does not seem a great cost to accept for a very simple implementation supporting co-located collaborative work in a single-user display setup. In contrast, Simon [12] reports 45% to 60% differences in selection task performance between a condition with an averaged viewpoint and a non-head-tracked condition. However, a direct comparison of ours and Simon's results is problematic since they used a fixed viewpoint when the user was not tracked and tested object selection and manipulation separately.

In contrast with the study described in [10], our users did not have to agree before changing their locations during the task. This results in free movement similar to how they would perform in practical settings. Nonetheless we can compare some aspects of our study with theirs, specifically their finding of a significant difference in performance when the two subjects were not standing side-by-side. In our study, for example in position P3, we did not observe significant differences in performance even though participants' positions are closer to a front-back arrangement (one user standing in front of the other) than a side-by-side arrangement. In addition, the steady increase in the maximum waiting time in table 1 for target-geometry positions P1 to P4 could be indicative of another effect not investigated by [10], namely scene composition. This matter warrants further investigation.

5 CONCLUSIONS

Several research approaches have attempted to solve or mitigate the difficulty in providing a correct perspective for multiple users in a co-located VE. Those approaches propose solutions with different degrees of success and complexity. However, that complexity can be reduced by supporting multiple users in a single-user VE. In this paper we investigated the implications of having several users in a single-user VE through the use of the active-passive approach.

The results of the study regarding the initial hypothesis, i. e. that the provision of a navigation tool for the active user allows a significant reduction in waiting time for the passive user, are inconclusive. However, our results show that a very simple implementation of a co-located collaborative setup using a single-user VE does not result in great additional cost. Our results also add to the work of [10], which only consider distance between users, by tentatively pointing at an additional effect with respect to target-geometry location. This seems to be a promising direction for future investigation.

REFERENCES

[1] M. Agrawala, A. C. Beers, I. McDowall, B. Fröhlich, M. Bolas, and P. Hanrahan. The Two-User Responsive Workbench: Support for Collaboration Through Individual Views of a Shared Space. In *Proceedings of ACM SIGGRAPH 97*, Computer Graphics Proceedings, Annual Conference Series, pages 327–332. ACM, 1997.

[2] V. Bayon, G. Griffiths, and J. R. Wilson. Multiple Decoupled Interaction: An Interaction Design Approach for Groupware Interaction in Co-Located Virtual Environments. *International Journal of Human-Computer Studies*, 64(3):192–206, 2006.

[3] R. Blach, M. Bues, J. Hochstrate, J. Springer, and B. Fröhlich. Experiences with Multi-Viewer Stereo Displays Based on LC-Shutters and Polarization. In *IEEE VR 2005 Workshop: Emerging Display Technologies*, pages 15–17. IEEE, 2005.

[4] S. Coburn, L. Rebenitsch, and C. Owen. Passive Viewpoints in a Collaborative Immersive Environment. In *Virtual Augmented and Mixed Reality. Designing and Developing Augmented and Virtual Environments*, pages 3–12. Springer, 2013.

[5] C. Cruz-Neira, D. J. Sandin, T. A. DeFanti, R. V. Kenyon, and J. C. Hart. The CAVE: Audio Visual Experience Automatic Virtual Environment. *Commun. ACM*, 35(6):64–72, 1992.

[6] B. Fröhlich, R. Blach, O. Stefani, J. Hochstrate, J. Hoffmann, K. Klüger, and M. Bues. Implementing Multi-Viewer Stereo Displays. In *Proceedings of the 13th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG 2005)*, pages 139–146. Eurographics, 2005.

[7] Y. Kitamura, T. Konishi, S. Yamamoto, and F. Kishino. Interactive Stereoscopic Display for Three or More Users. In *Proceedings of ACM SIGGRAPH 2001*, Computer Graphics Proceedings, Annual Conference Series, pages 231–240. ACM, 2001.

[8] A. Kulik, A. Kunert, S. Beck, R. Reichel, R. Blach, A. Zink, and B. Froehlich. C1x6: A Stereoscopic Six-User Display for Co-Located Collaboration in Shared Virtual Environments. *ACM Trans. Graph.*, 30(6):#188, 2011.

[9] J. Marbach. Image Blending and View Clustering for Multi-Viewer Immersive Projection Environments. In *Proceedings IEEE Virtual Reality 2009 Conference*, pages 51–54. IEEE, 2009.

[10] B. Pollock, M. Burton, J. W. Kelly, S. Gilbert, and E. Winer. The Right View from the Wrong Location: Depth Perception in Stereoscopic Multi-User Virtual Environments. *IEEE Trans. Vis. Comput. Graphics*, 18(4):581–588, 2012.

[11] J. P. Schulze, D. Acevedo, J. Mangan, A. Prudhomme, P. Nguyen, and P. Weber. Democratizing Rendering for Multiple Viewers in Surround VR Systems. In *Proceedings of 2012 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 77–80. IEEE, 2012.

[12] A. Simon. First-Person Experience and Usability of Co-Located Interaction in a Projection-Based Virtual Environment. In *VRST 05: Proceedings of the 2005 ACM Symposium on Virtual Reality Software and Technology*, pages 23–30. ACM, 2005.

[13] F. Zheng, T. Whitted, A. Lastra, P. Lincoln, A. State, A. Maimone, and H. Fuchs. Minimizing Latency for Augmented Reality Displays: Frames Considered Harmful. In *Proceedings of 2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 195–200. IEEE, 2012.