Virtual Workspace Positioning Techniques during Teleportation for Co-located Collaboration in Virtual Reality using HMDs

Yiran Zhang, Thi Thuong Huyen Nguyen, Nicolas Ladevèze, Cédric Fleury, Patrick Bourdot

Virtual Workspace Positioning Techniques during Teleportation for Co-located Collaboration in Virtual Reality using HMDs

Yiran Zhang[*]
University Paris-Saclay, CNRS, LISN, VENISE team, Orsay, France
Huyen Nguyen[†]
University Paris-Saclay, LISN, VENISE team, Orsay, France
Nicolas Ladevèze[‡]
University Paris-Saclay, CNRS, LISN, VENISE team, Orsay, France
Cédric Fleury[§]
IMT Atlantique, Lab-STICC, UMR CNRS 6285, Brest, France
Patrick Bourdot[¶]
University Paris-Saclay, CNRS, LISN, VENISE team, Orsay, France

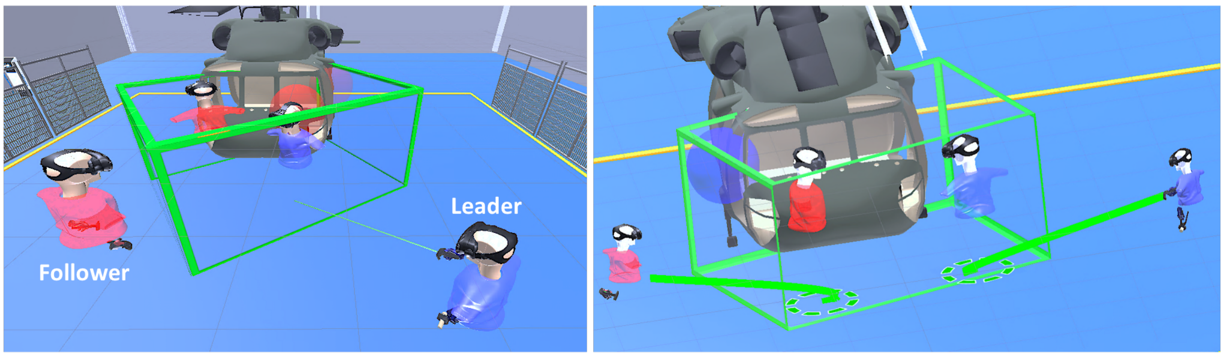

Figure 1: Two virtual workspace positioning techniques to help co-located users to recover their spatial consistency after a teleportation. The 3D volume framed in green represents the virtual representation in the VE of the user shared physical workspace. While positioning this virtual workspace, the users can predict their future position and orientation after the teleportation by observing their preview avatars (each user is represented by a distinct color). *Left*: **Leader-and-Follower** technique allows one of the users to fully manipulate the position and orientation of the virtual workspace. The other user can only communicates their own requirements for this manipulation. *Right*: **Co-manipulation** technique integrates the inputs from both of the users, allowing concurrent positioning.

Abstract

In many collaborative virtual reality applications, co-located users often have their relative position in the virtual environment matching the one in the real world. The resulting spatial consistency facilitates the co-manipulation of shared tangible props and enables the users to have direct physical contact with each other. However, these applications usually exclude their individual virtual navigation capability, such as teleportation, as it may break the spatial configuration between the real and virtual world. As a result, the users can only explore the virtual environment of approximately similar size and shape compared to their physical workspace. Moreover, their individual tasks with unlimited virtual navigation capability, which often take part in a continuous workflow of a complex collaborative scenario, have to be removed due to this constraint. This work aims to help overcome these limits by allowing users to recover spatial consistency after individual teleportation in order to re-establish their position in the current context of the collaborative task. We use a virtual representation of the user's shared physical workspace and develop two different techniques to position it in the virtual environment. The first technique allows one user to fully position the virtual workspace, and the second approach enables concurrent positioning by equally integrating the input from all the users. We compared these two techniques in a controlled experiment in a virtual assembly task. The results show that allowing two users to manipulate the workspace significantly reduced the time they spent negotiating the position of the future workspace. However, the inevitable conflicts in simultaneous co-manipulation were also a little confusing to them.

Index Terms: Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Virtual reality; Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Collaborative interaction

1 Introduction

Computer-supported collaborative work (CSCW) tools have been widely deployed in diverse fields such as industrial training, data exploration, product design, and entertainment, to name a few, to allow a group of users to communicate, interact with each other, and coordinate their activities to solve collaborative tasks. According to the collaborators' geographical location and whether the collaboration is performing simultaneously, the group interaction can be categorized according to a time/location matrix [13]. In this matrix, different group interactions can be distinguished as same-time (synchronous) and/or different-time (asynchronous) interactions, as well as same-location (co-located) and/or different-location (remote) interactions. This paper investigates co-located synchronous teamwork using head-mounted displays (HMDs), where multiple users share the same physical tracked space while immersing in a collaborative virtual environment (CVE).

In many collaborative virtual reality (VR) applications, the users'

[*]e-mail: yiran.zhang@universite-paris-saclay.fr
[†]e-mail: huyen.nguyen@universite-paris-saclay.fr
[‡]e-mail: nicolas.ladeveze@universite-paris-saclay.fr
[§]e-mail: cedric.fleury@imt-atlantique.fr
[¶]e-mail: patrick.bourdot@universite-paris-saclay.fr

relative position in the virtual environment (VE) matches their position in the real world. In this situation, the spatial consistency between the real (physical) and virtual environment enriches the VR experience by allowing, for example, a direct physical interaction between users [25] as well as the integration of shared tangible props in the scene [3,11,29]. However, this one-to-one mapping also limits the users' accessible area in the VE, and they can only explore the VE whose size is similar to the size of the physical workspace.

Teleportation is a widely used navigation technique that allows the users to explore a virtual space that is larger than their physical workspace while minimizing simulator sickness [15,19,36]. However, for the co-located users equipped with HMDs, after individual teleportation, the spatial relationship of their avatars in the virtual environment often differs from their counterpart in the real world. This position offset, also referred to as spatial desynchronization [23], makes the interaction that relies on the one-to-one mapping of the real and virtual workspace impossible. The lack of awareness of the position of other users in the real world increases the risk of collisions between them. In addition, they can only hear each other's voice from their position in the real world rather than their avatar's position. The dual-presence of the real and virtual audio stimuli generates perceptual conflicts and can greatly impact the user's task performance [10].

In this context, our work aims to overcome these limits by helping the users to recover the spatial consistency after individual teleportation. Such spatial consistency recovery techniques are useful in many scenarios, especially for complex workflows involving individual sub-tasks as well as collaborative sub-tasks (which specifically require spatial consistency). This is typically the case for complex virtual assembly simulations, where the users often have to perform a continuous virtual activity including individual and collaborative sub-tasks at different times, depending on the actual operation at hand. For example, during the assembly task, they can first navigate individually to different warehouses to obtain mechanical pieces. If the spatial consistency can be restored in the following collaborative phase when they come back to a shared space, they can then walk freely within this area and interact directly with each other without having any perceptual conflicts. Besides, shared tangible objects can also be integrated into the assembly task to coordinate the users' movement and provide them with additional passive haptic feedback.

In our approach, we deploy a virtual representation of the users' shared physical workspace in the VE. They can position such the virtual workspace in the VE while taking into consideration the requirements of their subsequent collaborative task. This step thus facilitates the recovery of the spatial consistency after teleporting inside it. We develop two techniques to allow the users to define the virtual workspace's position and orientation. In the scenario of two co-located users, the first technique enables one of the users to control the virtual workspace in the VE, while the other have to communicate their needs with verbal suggestions or other communication cues. The second technique integrates all of the users' inputs equally, thus enabling simultaneous positioning of the virtual workspace. Inspired by the virtual assembly task given as an example above, we envisioned a collaborative virtual riveting task to investigate the performance of these two techniques. The recovered spatial consistency allows the users to have direct physical contact to perform riveting, providing passive haptic feedback during the collaborative task. From the results of the controlled experiment, we derived some usability guidelines for such techniques.

The contributions of this work are:

1. The design of two interactive techniques that allow the users to recover a shared spatial consistency after individual navigation tasks to facilitate collaborative and tangible interaction between them. We intentionally developed two techniques that involve opposite types of collaboration in order to compare them: *Co-manipulation* is symmetric, while *Leader & Follower* is asymmetric.
2. Empirical results on participants' performance and preference when using these two techniques on a task alternating individual and collaborative sub-tasks.
3. An actual scenario demonstrating how recovering the spatial consistency can be useful for a collaborative VR task.

2 Related Work

Many collaborative VR application designs rely on a one-to-one mapping between the users' relative positions and rotations in the real and virtual environment. The spatial consistency provided by such mapping enables the possibility of introducing a tangible interface to the co-located users. Indeed, the blended real and virtual environment enriches the users' virtual experience and overcomes the lack of tactile feedback of VR, contributing to a higher sense of presence [21]. For example, a shared prop can be integrated in a virtual windshield [30] or a virtual car hood [3] assembly task to coordinate the users' co-manipulation. In addition, the spatial consistency between the real and virtual workspace allows the co-located users to interact directly with each other without going through an intermediary step, such as a handshake between them [25]. Finally, under the spatial consistency condition, the users' position in the real world is the same as in the VE, which helps to prevent possible collisions during real walking. Moreover, since the sound coming for the users matches with their virtual avatars' location, there is no perceptual conflict regarding the 3D spacial sound and thus the spatial information can be implicitly communicated as if it was in the real world.

One of the major drawbacks in the one-to-one mapping required by such applications is that it limits the size of the virtual environment to the same size of the users' physical workspace. It also constricts the use of virtual navigation since their individual navigation capabilities can break spatial consistency. One possible solution to avoid spatial desynchronization is to consider the co-located users as a group and allow them to virtually navigate as a single entity.

The physical workspace shared by the users can be embedded in the CVE by incorporating a virtual representation of the real environment into the virtual world. For example, 3DM [9] deploys a magic carpet to represent the tracking space. In addition, the user's physical workspace can be incorporated into the virtual environment by being imagined as a virtual vehicle [7] or a virtual cabin [14]. By manipulating such a virtual representation, the users can therefore navigate in the VE while preserving their spatial relationship. For example, C1x6 [22] allows co-located users to navigate inside a virtual vehicle as a group. The users can pilot the vehicle using a shared stationary 3D tracking sphere within the physical workspace. More recently, Multi-Ray jumping [35] allows co-located users to teleport as a group while maintaining their spatial offset during the navigation. When the navigator specifies a target teleportation position using a ray, the corresponding position of the passenger is computed and communicated using a second ray.

Group navigation during a continuous VR experience has its limits and is sometimes unnecessary. In cluttered or confined virtual environments, such as corridors, users often collide with or find themselves inside virtual objects in order to maintain spatial relationships among the group. To solve this problem, the previous study of Beck et al. [5] propose automatically moving users close to each other when they go through a narrow place and recovering the spatial consistency configuration after reaching a collision-free state. However, this approach causes short-term spatial desynchronization and induces users discomfort depending on their shifted offset to the open passage. Moreover, group navigation limits users' individual activities for loosely coupled collaboration stages, for example, individual object searching before the collaborative assembly task described in [10]. Therefore, it is crucial to preserve users' individual virtual navigation capabilities and help them recover spatial consistency in some areas of the VE when the need arises.

To meet these criteria, Min et al. [25] propose a recovery algorithm to adjust users' relative position and orientation in the virtual and physical world until they become aligned. The co-located users can use redirected walking technology independently to explore a VE larger than their physical workspace. When the spatial consistency is required to perform direct physical interaction in the VE, the users can trigger the recovery algorithm to achieve the recovered state. As the most natural method for traveling in a VE, walking can help the users better understand the size of VE by providing them with vestibular cues [8]. However, redirected walking technology usually requires a physical workspace larger than 6m × 6m [4], which is difficult to meet for many VR systems, especially the ones using HMDs. In addition, long physical walk may tire users after a certain time. Consequently, such solutions are not always suitable for large scale individual navigation in any VR systems.

In a single-user context, some previous works explore how to recover a spatial consistency between the real workspace and its virtual counterpart after large scale virtual navigation. They allow a user to teleport themself in a predefined virtual workspace maximizing their usable real space [39] or to choose the position and orientation of their future virtual workspace when they need to adapt the placement of their real space in the virtual environment [40]. However, these solutions are designed for a single user and cannot manage the spatial constraints of multiple users.

Inspired by these approaches, we extend them to the collaborative context by allowing users to customize their shared workspace position and orientation before teleportation while recovering a spatial consistency between them. We proposed in this paper two recovery techniques that enable individual virtual navigation and recover spatial consistency within user-defined areas when necessary for the subsequent collaborative interactions.

3 Spatial Consistency Recovery Techniques

In this section, we will detail the design and implementation of the two interactive techniques to recover spatial consistency in a co-located collaboration using HMDs. These recovery techniques incorporate a virtual representation of the physical workspace of co-located users in the VE. This representation is comparable to the one used for a single user in previous studies [39, 40], except that it handles multiple users. This virtual workspace representation includes a 3D volume with the same shape and size as the real workspace shared by the co-located users. In addition, we present the current group configuration by adding preview avatars [34, 35] in the 3D volume to directly show how each user will be positioned after teleportation. These avatars have different colors that correspond to the users' avatar colors in the VE.

By positioning the virtual workspace in the VE, users can define an area to recover spatial consistency regarding different scenarios and tasks. As collaborative interaction in the VE can be symmetric or asymmetric, we propose two different strategies to allow users to control the position and orientation of the virtual workspace: *Leader and Follower* and *Co-manipulation* techniques. Both of the techniques will be described for a pair of co-located collaborators but they can be extended to a bigger teamwork.

3.1 Leader and Follower Technique

Our first technique allows one of the users (named the *leader*) to position the virtual workspace, while the other user (named the *follower*) can only communicate verbally to the leader their requirements (see Figure 2). When approaching an area that requires spatial consistency for performing collaborative tasks, the leader can press the HTC Vive controller's touch-pad to trigger the control of the virtual workspace. It displays the virtual workspace for both users.

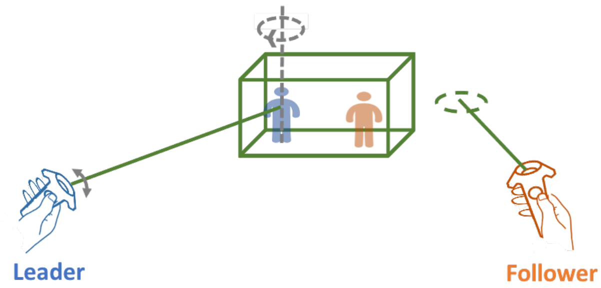

Figure 2: Leader and Follower technique: the leader controls the position and orientation of the virtual workspace and once it is configured, the follower can teleport directly into this space to recover the spatial consistency. The users can observe their future position within this workspace by looking at their preview avatars.

The intersection point between a virtual ray and the virtual ground determines the future position of the leader.

Based on this position and the actual spatial relationship between both users in the real world, the future position of the virtual workspace with the follower's position inside is computed. The leader can also rotate the virtual workspace around the vertical axis of their preview avatar by sliding their finger in a circle on the touch-pad with a one-to-one mapping. This design aims to help users better anticipate the virtual workspace configuration after teleportation by providing them with egocentric cues, as it rotates the virtual workspace around the user's future position. Our motivation is to leverage some generic study results which emphasise that egocentric cues can help users gather self-relevant information and estimate distances more accurately [24]. This choice is also encouraged by the result of a previous single-user virtual workspace positioning study, where we found that users preferred the egocentric design to the situation where the virtual workspace and the user's avatar are rotated as a whole unit [40].

The virtual workspace and the virtual ray controlled by the leader are always visible to the follower during the manipulation (see Figure 1 left). Consequently, even if the follower cannot take action, they can still communicate with their leader about the positioning of the virtual workspace. Once the future position of the virtual workspace is satisfactory to both of them, the leader can release the touch-pad to end the manipulation.

The leader is then teleported to this newly positioned workspace. The follower can use their ray to select the leader-defined virtual workspace and releases the touch-pad to teleport inside that workspace. The spatial consistency between users is thus restored. Instead of having the leader teleporting the two users at the same time, this process is split into two steps to avoid unwanted teleportation of the follower which can create frustration and disorientation.

3.2 Co-manipulation Technique

Unlike the first technique in which only one user (the *leader*) can position the virtual workspace, our second approach allows both users to simultaneously control the virtual workspace representation in the VE (see Figure 3). To achieve this goal, we considered different interaction techniques which have been proposed in the literature to enable multiple users to simultaneously manipulate a shared object. Indeed, collaborative object manipulation is one of the most important interaction tasks in CVE. One plausible solution is to average the translation and rotation of the user movements to obtain the final movement of the shared object [2, 16, 17, 29, 31]. In addition, the user input can be asymmetrically integrated by assigning different degree-of-freedom (DOF) control of the shared object to the users [3, 26].

In the context of teleportation, the users define a remote targeted

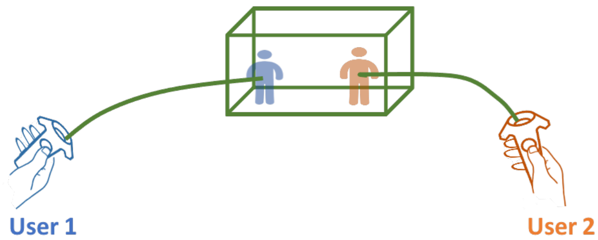

Figure 3: Co-manipulate technique from the front view: the two users will concurrently manipulate the workspace using a bending ray.

point using a pointing technique. To avoid introducing additional inputs, we propose a novel interaction technique that allows users to move and rotate the virtual workspace representation together based on the translation motion of targeted points on the virtual ground. In order to pull the virtual workspace towards users' desired configuration, we consider that the user-defined targeted positions and the users' preview avatars are connected by a mass-spring-damper system (see Figure 4). Similar physically-based approaches have been used to produce realistic virtual object grasping [6] or simulate collision during object manipulation [18].

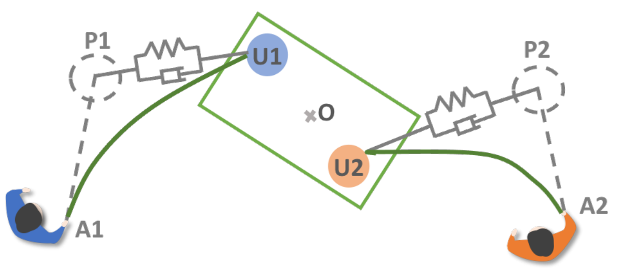

Figure 4: Co-manipulate technique from the top view: the position and orientation of the workspace are calculated from a physically-based approach using a mass-spring-damper system.

At each time step, the translational force coming from each user is first computed. If we assume that U_1 and U_2 are users' preview avatar positions inside the virtual workspace, P_1 and P_2 are the targeted positions defined by the users, then the force coming from *User 1* and *User 2* is computed as follows:

$$\vec{F_1} = k \cdot (P_1 - U_1) + b \cdot \vec{V_{p1}} \tag{1}$$

$$\vec{F_2} = k \cdot (P_2 - U_2) + b \cdot \vec{V_{p2}} \tag{2}$$

$$\vec{F} = \vec{F_1} + \vec{F_2} \tag{3}$$

where k and b are respectively the spring and damper coefficients, $\vec{V_{p1}}$ and $\vec{V_{p2}}$ are the velocity of the targeted point for *User 1* and *User 2*, respectively. The spring and damper coefficients were empirically set to 3.14N/m and 9.85N.s/m to maintain the critical damping of the system. Finally, we symmetrically integrate the user input by adding up the forces that come from both of them (Equation 3) and applying the total force to the point O, the center of gravity of the virtual workspace.

To allow the users to simultaneously control the virtual workspace rotation, we sum the torque coming from each user using the following formula:

$$\vec{T} = \vec{U_{1O}} \times \vec{F_1} + \vec{U_{2O}} \times \vec{F_2} \tag{4}$$

where the $\vec{U_{1O}}$ and $\vec{U_{2O}}$ are the vectors from the point U_1 and U_2 to the center of gravity of the virtual workspace, respectively.

Providing the users with appropriate feedback to show the current state of the shared object's position and orientation is critical in a collaborative manipulation task. Inspired by the Bent Pick Ray technique [27], we use a similar bending ray to continuously inform users about their mutual actions during the whole co-manipulation of the virtual workspace (see Figure 4). The curved ray starts at each user's virtual hand position (A_1 or A_2) and ends at the position of their preview avatar (U_1 or U_2). The user-defined targeted destinations (P_1 or P_2) serve as an additional control point to define a Bézier curve of the 2nd degree (i.e., parabolic curve segment). The deformation of the curved ray indicates the direction and intensity of the user's drag on the virtual workspace.

The users can press the HTC Vive controller's touch-pad to display the virtual ray. The system computes the distance between two targeted points defined by the users. The co-manipulation of the virtual workspace is triggered when this distance is smaller than a specific value (e.g., the length of the diagonal of the virtual workspace). This threshold is set to avoid the situation when one user wants to perform simple individual teleportation and accidentally activates the co-manipulation mode of the virtual workspace. As a first prototype, we implemented a simple approach to end the users' co-manipulation. When reaching an agreement on the configuration of the virtual workspace, either of them can end the co-manipulation by releasing the touch-pad. They will then be teleported into the newly defined workspace, and the spatial relationship between them will thus be recovered. However, by giving both users the ability to end the co-manipulation on their own, there is a risk that users can accidentally trigger this process due to misunderstandings in communication. In future studies, this design could be improved using some alternative solutions, such as locking the virtual workspace manipulation when one of the users is satisfied with the current configuration and waiting for the confirmation from the other.

Finally, if users want to stop the co-manipulation and return to the basic individual teleportation state, they can intentionally increase the distance between the two targeted points and exceed the predefined threshold. The threshold values used for starting and stopping the co-manipulation can be parameterized. For example, the stopping value can be greater than the starting one to allow users to manipulate the virtual workspace over a broader range. In addition, an additional threshold can be set between these two values. For example, when the distance between the two separate target points is about to reach this threshold, the curved rays will turn red, informing users in advance that the co-manipulation is about to end and that the virtual workspace is about to disappear.

4 User Study

We conducted a controlled experiment to compare the two spatial consistency recovery techniques presented in the previous section. We did not compare these two techniques with a baseline condition with which the users have to perform individual teleportation without any assistance to recover the spatial consistency. This is because, for such a baseline condition, it is nearly impossible for the users to recover their spatial consistency without removing their HMDs from time to time, verifying their spatial relationship in the real world, and applying it to the virtual world.

In this experiment, we set up a virtual riveting task in which the co-located participants were asked first to complete an individual task and then return to a designated area to achieve together the riveting of a helicopter shell. Participants had to position the virtual workspace representation according to the current riveting task location. The restored spatial consistency inside the newly defined workspace allows participants to interact directly with each other during the riveting process.

The experiment followed a within-subject design and assessed two TECHNIQUES, i.e., *Leader-and-Follower* and *Co-manipulation* described in Section 3 (see Figure 1). The experimental protocol has been approved by the ethics committee of the university.

4.1 Hypotheses

We assumed that *Co-manipulation* would help users to better position the workspace as their intention can be directly conveyed by the manipulation of the virtual workspace, and thus avoiding possible misunderstandings that can occur during the verbal communication in *Leader-and-Follower*. Moreover, we expected to find a power imbalance in the negotiation between the participants and differences in their respective contribution to the task in the *Leader-and-Follower* as it implies an asymmetric role assignment in the virtual workspace positioning process. Therefore, the following hypothesises are formulated:

H1 *Leader-and-Follower* will require more time to discuss and negotiate the future workspace position compared to *Co-manipulation*.

H2 *Co-manipulation* will induce better workspace positioning resulting in a better performance for the riveting task.

H3 *Leader-and-Follower* will be more challenging for the leader.

4.2 Participants

We recruited 24 participants, aged between 21 and 50 ($M = 27.27 \pm 5.51$), with normal or corrected-to-normal vision. 12 pairs were formed at the time of recruitment resulting in 8 male-male and 4 female-male groups. 18 out of 24 participants had previous VR experience and 10 of them rated their everyday usage of HMDs as very low.

4.3 Experiment setup

The VR setup includes two HTC Vive Pro Eye headsets [1]. The outside-in tracking supported by the HTC Vive Lighthouse Tracking System enables tracking of co-located users when one user is out of sight of another (for example, when one user is behind another), ensuring the safety of the user. The two workstations controlling the headsets are connected via a local network, and the tracking spaces of the two users are aligned to a common coordinate system by a calibration procedure [37]. User inputs were obtained through the two Vive handheld controllers that were used for each user. The experiment room supported a 3m $\times$ 4m tracking area. The virtual environment was rendered using Unity (released 2019.4.21) with a resolution of 1440 $\times$ 1600 pixels per eye at 90 Hz. The average time from sending a message from one headset until the reception was measured at 5ms.

4.4 Experiment task

Before the experiment, each participant of a pair was asked to walk to their respective starting points presented by dashed circles on the virtual floor. Each participant was equipped with two controllers, one for teleportation and another for the riveting task. The latter was presented in the VE as a hammer or a riveting pliers, depending on the participant's role. Each participant first performed an individual task to prepare the team riveting task. As illustrated in Figure 5, the participant with the hammer had to teleport close to the charging station to charge the hammer, while the other participant needed to teleport near a shelf to grab the rivets with the riveting pliers. To ensure the safety of the participants during the individual navigation phase towards the charging station and the shelf, warning signs appeared in their field of view with alarm sounds when they reached the limit of their physical working space, or stayed too close to each other and were about to collide [12, 23, 33]. After the participants finished their own individual tasks, they returned to the team riveting area indicated by the yellow frame on the ground. Then, participants had to position their virtual workspace to enclose three riveting positions required by the collaborative task.

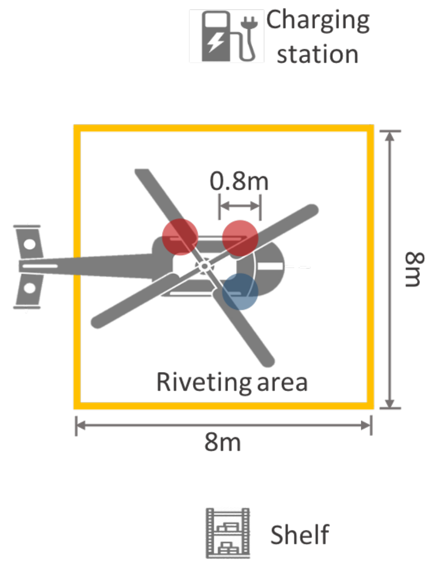

Figure 5: Top view of the VE implemented for the user study: a riveting area (including semi-transparent spheres presenting three predetermined riveting positions) as the shared working area, and two separate areas (including the charging station and the shelf) for individuals tasks.

Individual criteria for the future virtual workspace position were provided differently to the participants: one participant was informed of a part of the required riveting positions, while the other was informed of the remaining positions. Such design is based on the fact that each worker may have a different sequence workflow in an actual riveting process according to their expertise and preference. In addition, the negotiation is encouraged during the virtual workspace positioning stage to mitigate the imbalance of control in *Leader-and-Follower* condition.

The predetermined riveting positions were enclosed in a semi-transparent sphere (d = 0.8m) and displayed to the participants. There were three spheres in total in each operation and each sphere includes two riveting points. Each user was randomly assigned to see one or two of the three spheres either in red or blue. For example, in the case shown in Figure 5, one user can see the two spheres in red, while the other can see the remaining one in blue. To help the participants better determine the position and orientation of the virtual workspace, the color of the sphere is darkened when it is completely enclosed by the workspace.

After the participants configured the position and orientation of the virtual workspace and arrived at this newly positioned workspace with a recovered spatial consistency, the team riveting task began. The participant with the riveting pliers had to place the rivet in the drilled hole of the helicopter shelf, while the other completed the riveting by tapping the end of the rivet with the hammer. The rivet end and the hammer were respectively mapped to the upper area of the controllers held by the participants. As a result, they could feel the hit as they hammered the rivet, which provided them with passive haptic feedback for the collaborative riveting (see Figure 6). When the two riveting points inside one sphere were filled, the color of this sphere faded to gray, prompting the participants to walk to the next riveting position. During the team riveting, since the users' positions in the real world are the same as their avatars' in the VE, the warning signs were only triggered when the users reached the limit of their physical workspace.

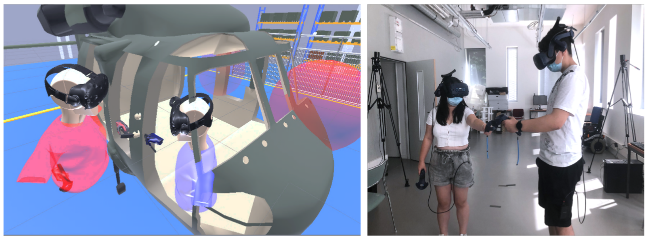

Figure 6: The collaborative riveting task requires that the two controllers of the users (which are displayed as a hammer or a riveting pliers depending on the role of the user) come into direct contact to perform the riveting. The users can feel the hit when hammering the rivet, which provides them passive haptic feedback for the collaborative task. *Left*: bird-eye view of the virtual environment. *Right*: view of the shared real environment.

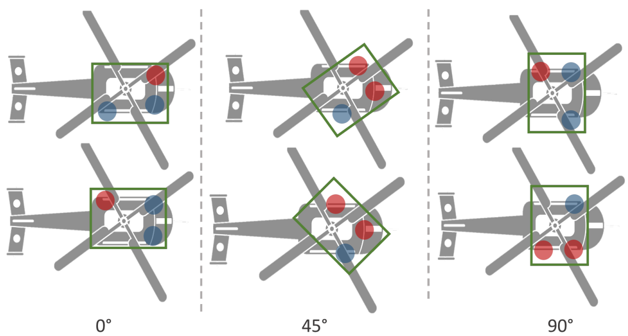

Figure 7: The six optimal virtual workspace positions for the six variations the riveting task. These variations included three types of rotations (0°, 45°, 90°) and different riveting positions.

Based on the size of the physical workspace and the riveting area, the co-manipulation starting and stopping values are set to 3m and 10m, respectively, and the stopping warning threshold was set to 8m.

4.5 Procedure

After arriving at the laboratory, each pair of participants received instructions on the task, signed an informed consent form, and filled out a demographic questionnaire. In each session, participants tried two conditions (techniques), a first one then the other as two sub-sessions, each sub-session including a set of trials. The order of presentation of the conditions was counterbalanced among the participants. At the beginning of each condition, the participants received a training trial. During this training, the experimenter was allowed to answer the participants' questions, if any. Then, the participants completed six trials in a randomized order with different targeted riveting positions, resulting in six optimal virtual workspace positions with three different rotations (0°, 45°, 90°), as illustrated in Figure 7. Participants filled out a questionnaire after each technique. At the end of the experiment, participants ranked the two techniques according to their preferences. The whole experiment lasted about 45 minutes, and the total VR exposure time was about 30 minutes on average.

4.6 Data collection

We registered 144 trials: 2 TECHNIQUES × 6 repetitions × 12 pairs. For each trial, we logged the following measures:

- *Task Completion Time (TCT)* is the total time spent by the participants completing one trial. Measurements began at the start of the individual task and ended when the six rivets were installed.
- *Workspace positioning time* is the time spent by the participants positioning the virtual workspace. The measurement started when both participants entered the riveting area and ended when they teleported inside the newly defined workspace (by releasing the touch-pad of their controller). For each of these measurements, we summed the values if more than one workspace positioning operation was required during the trial.
- *Riveting time* is the time spent by the participants completing the riveting task. The measurement began when the participants were first teleported into the shared workspace and continued until all six rivet placements were completed.
- *Number of positions* is the number of times the participants positioned the virtual workspace in one trial to complete the installation of the six rivets.
- *Number of warnings* is the number of the warnings triggered by the participants due to a collision with the workspace borders during the team riveting.

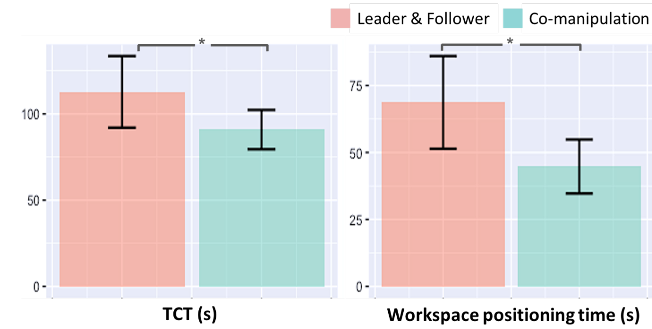

Figure 8: Mean *TCT* (left) and *workspace positioning time* (right) by technique. Error bars show 95% confidence intervals (CI)

We used the NASA-TLX questionnaire [20] to assess the cognitive task load. Participants were also asked to evaluate their leadership ("Who was the leader, you or your partner?"), contribution ("To what extent did you and your partner contribute to positioning the workspace?"), and talkativeness ("Who talked the most, you or your partner?"). Several previous studies have used similar questions to investigate leadership in collaborative tasks [32, 38]. Criteria were graded on a 21-point scale and later converted to a 100-point score.

4.7 Statistical results

We averaged the 6 repetitions of each technique to minimize the noise in the data. All statistical analyses were performed in R with a significance level of $\alpha = 0.05$ for all tests. Means (M) are reported with standard deviations.

For *TCT* (see Figure 8, left), we used Shapiro-Wilk test and QQ plots to analyze data normality. The data did not conform to normal distribution, so we applied a log transformation to it following the statistical recommendations [28] (p.316). The Kolmogorov's D-test then showed its goodness-of-fit to the log-normal distribution. We thus ran the analysis using the log-transform of *TCT*. The paired sample t-test revealed that the participants achieved the task significantly faster with *Co-manipulation* ($M = 90.90s \pm 17.95s$)) than with *Leader-and-Follower* ($M = 112.81s \pm 34.00s$, p = 0.037) with an effect size of 0.57.

Regarding *workspace positioning time* (see Figure 8, right), we followed the same analysis procedure as applied to *TCT* and observed a significant difference of TECHNIQUES in the paired samples t-test (p = 0.008) with a large effect size of 0.81. It was shown that the participants spent significantly longer time positioning the workspace with *Leader-and-Follower* ($M = 68.70s \pm 28.46s$) than with *Co-manipulation* ($M = 44.74s \pm 15.87s$)).

Concerning *riveting time* (see Figure 9, left), a paired sample t-test was used as the data was normally distributed. We did not find any significant difference between *Co-manipulation* ($M = 29.48s \pm 5.58s$)) and *Leader-and-Follower* ($M = 35.66s \pm 11.59s$, p = 0.075).

For *number of positions* (see Figure 9, right), we used a non-parametric test in conformity with the nature of count data. Wilcoxon signed-rank test revealed that the participants positioned the workspace significantly more often with *Leader-and-Follower* ($M = 1.22 \pm 0.32$) than with *Co-manipulation* ($M = 1.04 \pm 0.10$), p = 0.025) with an effect size of 0.74.

For *number of warnings* (see Figure 10, left), we used a non-parametric test for post-hoc analysis in conformity with this type of data. Wilcoxon signed-rank test showed that the participants detected significantly more the warnings with *Leader-and-Follower* ($M = 0.85 \pm 0.50$) than with *Co-manipulation* ($M = 0.58 \pm 0.46$), p = 0.034) with an effect size of 0.55.

Regarding the subjective questionnaire, we used non-parametric Wilcoxon signed-rank tests for post-hoc analysis. For NASA-TLX score, we did not find any significant difference on cognitive task load between *Leader-and-Follower* ($M = 43.09 \pm 12.32$) and *Co-manipulation* ($M = 42.47 \pm 11.75$), p = 0.99). In *Leader-and-Follower*, no significant differences in talkativeness were found among participants in the leader role ($M = 42.92 \pm 16.71$) and follower role ($M = 48.33 \pm 18.50$), p = 0.506). Leaders ($M = 50.00 \pm 11.48$) were also found to have no significant differences in the level of contribution as followers ($M = 52.50 \pm 5.84$, p = 0.792). However, we detected an imbalance of leadership in *Leader-and-Follower* condition, with higher value for leaders ($M = 40.00 \pm 14.61$) compared to followers ($M = 57.92 \pm 17.64$), p = 0.015) with a large effect size of 1.11. Finally, 14 out of 24 participants preferred *Co-manipulation* over *Leader-and-Follower*. However, this result was not statistically significant according to the Binomial test (p = 0.541). Six participants out of 10 who preferred *Leader-and-Follower* took a leadership role in the experiment.

4.8 Discussion

The results provide evidence that *Co-manipulation* was more efficient than the *Leader-and-Follower* for workspace positioning. In *Co-manipulation*, the participants spent significantly less time on completing the task. In particular, the time used for negotiation and positioning the workspace was significantly decreased when both participants were able to manipulate the workspace, compared to the *Leader-and-Follower*. It therefore supports **H1**. This can be explained by the fact that each participant could adjust the position of the workspace according to the individual criteria provided to them. Besides, participants' desired positions and intentions could be communicated implicitly through the manipulation of the workspace. In the *Leader-and-Follower*, communicating the desired positions could be difficult, and we observed three different approaches that the participants used to achieve this. The participants could either (i) teleport themselves close to the required riveting point, (ii) use a ray to point to the target position, or (iii) use a virtual object (e.g., helicopter or avatar) as a reference to verbally describe its position. However, the use of additional teleportation or pointing, as well as the potential misunderstandings that could arise from the verbal communication, resulted in the need to use more time to exchange the information on riveting positions between the participants.

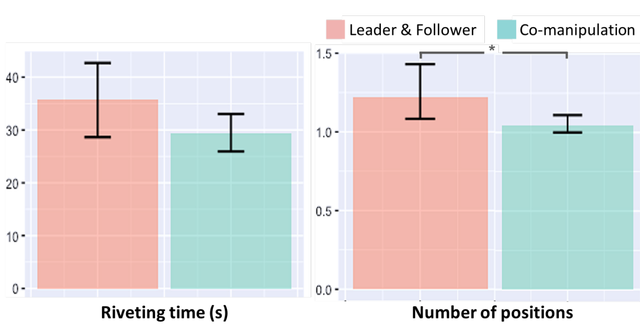

Figure 9: Mean *riveting time* (left) and *number of positions* (right) by technique. Error bars show 95% CI.

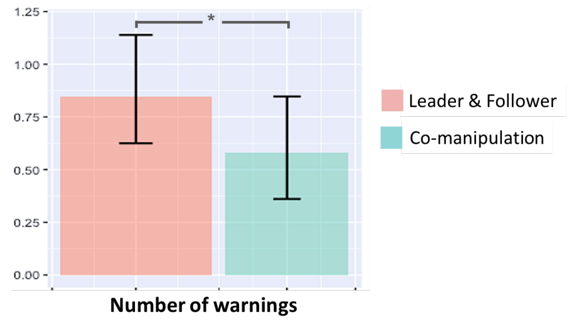

Figure 10: Mean *number of warnings* by technique. Error bars show 95% CI.

Contrary to our expectations, we were not able to find any significant difference between *Co-manipulation* and *Leader-and-Follower* in terms of *riveting time*, which does not support **H2**. However, we found that the participants triggered significantly more warnings and executed a significantly larger number of workspace positioning operations in *Leader-and-Follower* during the riveting process. Indeed, the accuracy of the workspace positioning affects the performance of the subsequent riveting tasks. If the user-defined workspace does not include all the required rivet positions, warnings will be triggered when users attempt to access a rivet outside the workspace. The virtual workspace then needs to be re-positioned to complete the task. Therefore, some minor accuracy differences may still exist between these two techniques.

Participants agreed that it was the leader who directed the workspace positioning in *Leader-and-Follower*. However, no significant difference in levels of verbal activity and contribution was found, which rejects **H3**. This can be explained by the fact that the individual position criteria were provided to the follower, which required them to actively join the workspace positioning task. We also did not detect any significant difference between *Leader-and-Follower* and *Co-manipulation* for cognitive task load. Although the *Co-manipulation* can be more efficient, a few participants also felt "out of control" (P4) in such a condition as the simultaneous manipulation inevitably produced "conflicts" (P11) during the virtual workspace positioning. This may also explain why a large number of participants, especially those in the leader role, preferred to use the *Leader-and-Follower* technique.

According to the participants' self-evaluation of their VR experience, the 12 pairs consisted of three novice-novice groups, four expert-novice groups and five expert-expert groups. In the *Co-manipulation* condition, the expert-expert groups ($M = 80.39s \pm 11.81s$) outperformed the expert-novice groups ($M = 98.92s \pm$

24.79s) and the novice-novice groups ($M = 97.94s \pm 9.77s$) with less average task completion time. However, the same result was not found in the *Leader-and-Follower* condition, as the expert-expert groups ($M = 121.88s \pm 45.68s$) needed more time to complete the task than the two other groups ($M = 106.80s \pm 29.34s$ for expert-novice groups and $M = 105.71s \pm 7.56s$ for novice-novice groups). It is difficult to draw formal conclusions based on such a limited number of pairs by groups, and further studies are needed. However, several explanations are still worth discussing. One is that the *Leader-and-Follower* technique could rely more on the approach the users use to achieve spatial information exchange than on their level of VR experience. Another is that when both users are experts, it is possible that the follower may not accept their status and therefore may challenge the leader's decisions more frequently, thus increasing the execution time (i.e., the follower has their own understanding of workspace management due to their VR expertise and would like to act as a leader).

5 Conclusion

This paper proposes and evaluates two interactive techniques that allow two co-located users to recover spatial consistency between their real and virtual workspace after individual teleportation. These recovery techniques use a virtual representation of the users' shared physical workspace in the VE. The *Leader-and-Follower* technique allows only one user to position the virtual workspace, while the *Co-manipulation* enables both of them to manipulate the virtual workspace at the same time. The recovered spatial consistency allows the users to access surrounding virtual objects through physical movements, enables direct physical interaction between them, and avoids the perceptual conflicts of dual presence of the users and their avatars in the virtual scene.

We conducted a controlled experiment to evaluate the performance of the two techniques in a collaborative virtual riveting task. The results showed that the positioning of the virtual workspace can be significantly faster with *Co-manipulation* than with *Leader-and-Follower*. In *Co-manipulation*, users' intents can be directly communicated by manipulating their future virtual workspace, shortening the time it takes to communicate its correct position and allowing an efficient subsequent collaborative task. In addition, significantly more attempts to reposition the virtual workspace and more warnings given when the users collided with the workspace borders during the collaborative task were measured in *Leader-and-Follower*. However, despite a better performance of the *Co-manipulation*, it also introduces conflicts in the way to position the virtual workspace. Moreover, it is sometimes difficult for users to understand their impact on controlling the final position and orientation of the workspace during the co-manipulation.

Nevertheless, *Leader-and-Follower* may be a relevant technique to reach spatial consistency when one of the users is well aware of all the requirements of the subsequent collaborative tasks. For example, in training or education application, the trainer may be responsible for placing the workspace to include all the required training contents for the following tasks. Moreover, this technique can also be applied for asynchronous collaborative interaction. In a collaborative system that uses tangible interaction, a user can define a workspace and leave a tangible prop in it. When other users arrive at the scene, they can continue to use that same prop within the workspace configured by the previous user. However, these two techniques should be further improved when a physical prop is used in the workspace. When recovering the spatial consistency between the users' real workspace and the virtual workspaces, the physical object must be paired with its virtual counterpart to allow the users to touch it when reaching. The physical object thus adds more constraints to the possible positioning of virtual workspace in the VE, even imposing a unique positioning solution.

Both of these techniques can be extended to collaborative tasks involving more than two users. However, the increasing number of users brings more conflicts and difficulties to concurrent manipulation. Therefore, further investigation is necessary to determine which technique is more appropriate for different numbers of users and in various collaboration scenarios. Moreover, different spring-damper values can be tested in the co-manipulation condition, giving users unbalanced control over the positioning of the virtual workspace and generating an alternative in-between the co-manipulation and the leader-follower approach. It could be thus interesting to investigate the impact of this asymmetric strategy on users during the collaborative virtual workspace positioning in some future studies.

Finally, these spatial consistency recovery techniques can also be used in some AR training applications. When using these AR setups, the user can see and hear other users in the real world from a different position and orientation than their avatars after individual virtual navigation, which can be confusing. Our techniques can solve this spatial desynchronization issue and correct the position and orientation mapping of the users for the collaboration phase.

Acknowledgments

This work was supported by French government funding managed by the National Research Agency under the Investments for the Future program (PIA) with the grant ANR-21-ESRE-0030 (CONTINUUM project).

References

[1] HTC Vive Pro Eye. `https://business.vive.com/fr/product/vive-pro-eye-office/`. Accessed: 2022-01-05.

[2] L. Aguerreche, T. Duval, and A. Lécuyer. Short paper: 3-hand manipulation of virtual objects. In *JVRC 2009*, pp. 4–p, 2009.

[3] L. Aguerreche, T. Duval, and A. Lécuyer. Comparison of three interactive techniques for collaborative manipulation of objects in virtual reality. In *CGI 2010 (Computer Graphics International)*, 2010.

[4] M. Azmandian, T. Grechkin, M. T. Bolas, and E. A. Suma. Physical space requirements for redirected walking: How size and shape affect performance. In *ICAT-EGVE*, pp. 93–100, 2015.

[5] S. Beck, A. Kunert, A. Kulik, and B. Froehlich. Immersive group-to-group telepresence. *IEEE Transactions on Visualization and Computer Graphics*, 19(4):616–625, 2013.

[6] C. W. Borst and A. P. Indugula. Realistic virtual grasping. In *IEEE Proceedings. VR 2005. Virtual Reality, 2005.*, pp. 91–98. IEEE, 2005.

[7] P. Bourdot and D. Touraine. Polyvalent display framework to control virtual navigations by 6dof tracking. In *Virtual Reality, 2002. Proceedings. IEEE*, pp. 277–278. IEEE, 2002.

[8] D. Bowman, E. Kruijff, J. J. LaViola Jr, and I. P. Poupyrev. *3D User interfaces: theory and practice, CourseSmart eTextbook*. Addison-Wesley, 2004.

[9] J. Butterworth, A. Davidson, S. Hench, and M. T. Olano. 3dm: A three dimensional modeler using a head-mounted display. In *Proceedings of the 1992 Symposium on Interactive 3D Graphics*, I3D '92, pp. 135—-138. Association for Computing Machinery, 1992.

[10] W. Chen, N. Ladeveze, C. Clavel, D. Mestre, and P. Bourdot. User cohabitation in multi-stereoscopic immersive virtual environment for individual navigation tasks. In *Virtual Reality (VR), 2015 IEEE*, pp. 47–54. IEEE, 2015.

[11] L.-P. Cheng, S. Marwecki, and P. Baudisch. Mutual human actuation. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, pp. 797–805, 2017.

[12] G. Cirio, M. Marchal, T. Regia-Corte, and A. Lécuyer. The magic barrier tape: a novel metaphor for infinite navigation in virtual worlds with a restricted walking workspace. In *Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology*, pp. 155–162. ACM, 2009.

[13] C. A. Ellis, S. J. Gibbs, and G. Rein. Groupware: some issues and experiences. *Communications of the ACM*, 34(1):39–58, 1991.

[14] C. Fleury, A. Chauffaut, T. Duval, V. Gouranton, and B. Arnaldi. A generic model for embedding users' physical workspaces into multi-

scale collaborative virtual environments. In *ICAT 2010 (20th International Conference on Artificial Reality and Telexistence)*, 2010.
[15] J. Frommel, S. Sonntag, and M. Weber. Effects of controller-based locomotion on player experience in a virtual reality exploration game. In *Proceedings of the 12th international conference on the foundations of digital games*, pp. 1–6, 2017.
[16] A. S. García, J. P. Molina, P. González, D. Martínez, and J. Martínez. An experimental study of collaborative interaction tasks supported by awareness and multimodal feedback. In *Proceedings of the 8th International Conference on Virtual Reality Continuum and its Applications in Industry*, pp. 77–82. ACM, 2009.
[17] A. S. García, J. P. Molina, D. Martínez, and P. González. Enhancing collaborative manipulation through the use of feedback and awareness in CVEs. In *Proceedings of the 7th ACM SIGGRAPH international Conference on Virtual-Reality Continuum and Its Applications in industry*, p. 32. ACM, 2008.
[18] G. Gonzalez-Badillo, H. I. Medellin-Castillo, and T. Lim. Development of a haptic virtual reality system for assembly planning and evaluation. *Procedia Technology*, 7:265–272, 2013.
[19] M. J. Habgood, D. Moore, D. Wilson, and S. Alapont. Rapid, continuous movement between nodes as an accessible virtual reality locomotion technique. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 371–378. IEEE, 2018.
[20] S. G. Hart and L. E. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In P. A. Hancock and N. Meshkati, eds., *Human Mental Workload*, vol. 52 of *Advances in Psychology*, pp. 139 – 183. North-Holland, 1988.
[21] H. G. Hoffman. Physically touching virtual objects using tactile augmentation enhances the realism of virtual environments. In *Proceedings. IEEE 1998 Virtual Reality Annual International Symposium (Cat. No. 98CB36180)*, pp. 59–63. IEEE, 1998.
[22] 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 Transactions on Graphics (TOG)*, 30(6):1–12, 2011.
[23] J. Lacoche, N. Pallamin, T. Boggini, and J. Royan. Collaborators awareness for user cohabitation in co-located collaborative virtual environments. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, p. 15. ACM, 2017.
[24] J. M. Loomis, J. A. Da Silva, N. Fujita, and S. S. Fukusima. Visual space perception and visually directed action. *Journal of Experimental Psychology: Human Perception and Performance*, 18(4):906, 1992.
[25] D.-H. Min, D.-Y. Lee, Y.-H. Cho, and I.-K. Lee. Shaking hands in virtual space: Recovery in redirected walking for direct interaction between two users. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 164–173. IEEE, 2020.
[26] M. S. Pinho, D. A. Bowman, and C. M. Freitas. Cooperative object manipulation in immersive virtual environments: framework and techniques. In *Proceedings of the ACM symposium on Virtual reality software and technology*, pp. 171–178. ACM, 2002.
[27] K. Riege, T. Holtkamper, G. Wesche, and B. Frohlich. The bent pick ray: An extended pointing technique for multi-user interaction. In *3D User Interfaces (3DUI'06)*, pp. 62–65. IEEE, 2006.
[28] J. Robertson and M. Kaptein. *Modern Statistical Methods for HCI*. Springer, 2016.
[29] R. A. Ruddle, J. C. Savage, and D. M. Jones. Symmetric and asymmetric action integration during cooperative object manipulation in virtual environments. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 9(4):285–308, 2002.
[30] H. Salzmann, J. Jacobs, and B. Froehlich. Collaborative interaction in co-located two-user scenarios. In *Proceedings of the 15th Joint virtual reality Eurographics conference on Virtual Environments*, pp. 85–92. Eurographics Association, 2009.
[31] H. Salzmann, M. Moehring, and B. Froehlich. Virtual vs. real-world pointing in two-user scenarios. In *Virtual Reality Conference, 2009. VR 2009. IEEE*, pp. 127–130. IEEE, 2009.
[32] A. Steed, M. Slater, A. Sadagic, A. Bullock, and J. Tromp. Leadership and collaboration in shared virtual environments. In *Proceedings IEEE Virtual Reality (Cat. No. 99CB36316)*, pp. 112–115. IEEE, 1999.
[33] M. Wang, S. L. Lyckvi, C. Chen, P. Dahlstedt, and F. Chen. Using advisory 3d sound cues to improve drivers' performance and situation awareness. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 2814–2825, 2017.
[34] T. Weissker and B. Froehlich. Group navigation for guided tours in distributed virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, 27(5):2524–2534, 2021.
[35] T. Weissker, A. Kulik, and B. Froehlich. Multi-ray jumping: comprehensible group navigation for collocated users in immersive virtual reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 136–144. IEEE, 2019.
[36] T. Weissker, A. Kunert, B. Fröhlich, and A. Kulik. Spatial updating and simulator sickness during steering and jumping in immersive virtual environments. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 97–104. IEEE, 2018.
[37] T. Weissker, P. Tornow, and B. Froehlich. Tracking multiple collocated htc vive setups in a common coordinate system. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 593–594. IEEE, 2020.
[38] J. Wideström, A.-S. Axelsson, R. Schroeder, A. Nilsson, I. Heldal, and Å. Abelin. The collaborative cube puzzle: A comparison of virtual and real environments. In *Proceedings of the third international conference on Collaborative virtual environments*, pp. 165–171, 2000.
[39] Y. Zhang, N. Ladeveze, C. Fleury, and P. Bourdot. Switch techniques to recover spatial consistency between virtual and real world for navigation with teleportation. In *International Conference on Virtual Reality and Augmented Reality*, pp. 3–23. Springer, 2019.
[40] Y. Zhang, N. Ladevèze, H. Nguyen, C. Fleury, and P. Bourdot. Virtual navigation considering user workspace: Automatic and manual positioning before teleportation. In *26th ACM Symposium on Virtual Reality Software and Technology*, pp. 1–11, 2020.