

Merging Environments for Shared Spaces in Mixed Reality

Ben J. Congdon
Department of Computer Science
University College London
ben.congdon.11@ucl.ac.uk

Tuanfeng Wang
Department of Computer Science
University College London
tuanfeng.wang@cs.ucl.ac.uk

Anthony Steed
Department of Computer Science
University College London
a.steed@ucl.ac.uk

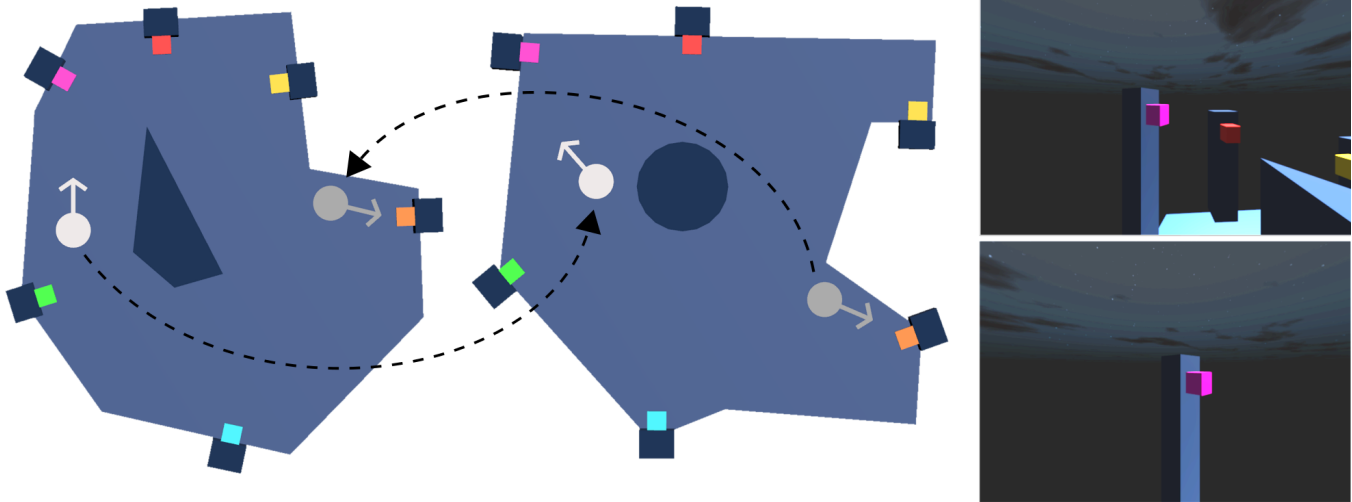


Figure 1: Overview of the technique. In order to maximize use of local physical space, the virtual world is constructed and presented differently to two users. Their movements are dynamically mapped into their collaborator's environment so as to create the impression that they are each sharing their own environment with the other. The left two images show floorplans for two rooms, each containing one local user and an avatar of a remote user. The remote user's position and orientation in their own room is used to place the avatar via a pre-generated forward mapping. The right two images show the viewpoint of the user in white in their own room (top) and the viewpoint of their mapped avatar (bottom).

ABSTRACT

In virtual reality a real walking interface limits the extent of a virtual environment to our local walkable space. As local spaces are specific to each user, sharing a virtual environment with others for collaborative work or games becomes complicated. It is not clear which user's walkable space to prefer, or whether that space will be navigable for both users.

This paper presents a technique which allows users to interact in virtual reality while each has a different walkable space. With this method mappings are created between pairs of environments. Remote users are then placed in the local environment as determined by the corresponding mapping.

A user study was conducted with 38 participants. Pairs of participants were invited to collaborate on a virtual reality puzzle-solving task while in two different virtual rooms. An avatar representing

the remote user was mapped into the local user's space. The results suggest that collaborative systems can be based on local representations that are actually quite different.

CCS CONCEPTS

• **Human-centered computing** → *Mixed / augmented reality; Virtual reality; Collaborative interaction;*

KEYWORDS

mixed reality, virtual reality, augmented reality, head-mounted display, computer graphics, virtual environments, remote collaboration, virtual co-location, planar map

ACM Reference Format:

Ben J. Congdon, Tuanfeng Wang, and Anthony Steed. 2018. Merging Environments for Shared Spaces in Mixed Reality. In *VRST 2018: 24th ACM Symposium on Virtual Reality Software and Technology (VRST '18), November 28-December 1, 2018, Tokyo, Japan*. ACM, New York, NY, USA, 8 pages. <https://doi.org/10.1145/3281505.3281544>

1 INTRODUCTION

One of the most pressing problems in virtual reality is that of locomotion around environments. With head-tracked systems, it is

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

VRST '18, November 28-December 1, 2018, Tokyo, Japan

© 2018 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-6086-9/18/11.

<https://doi.org/10.1145/3281505.3281544>

very attractive to allow movement over short distances by simply having the user walk a few steps. However, users have limited space available so the physical environment must be taken into consideration. Alternatives to real walking are in use today where movement is driven by teleportation, pointing, leaning, head direction, or by an external device such as a joystick, but the benefits of real walking on comfort, presence and problem solving are well documented [15, 22, 25, 27].

As real walking remains preferable there is a motivation to design environments around the physical features of the user's environment, by placing boundaries in the virtual world where there are obstacles in the physical environment. This might mean resizing the user's walkable area in the virtual world, perhaps a room, cockpit or game board. More sophisticated solutions might generate geometry for complicated walkable areas, or place virtual props to indicate physical obstacles. These approaches ensure that all elements of the environment can be reached by the user, preventing physical collisions. However, this scales poorly to multi-user VR. Users have different spaces available and it is not clear how best to combine these spaces. A naive approach would be to take the intersection of all user's spaces, but this will quickly lead to an unusably small space.

Augmented reality (AR) faces a similar issue. In AR the virtual environment is the user's physical environment with virtual additions. A significant part of the appeal of AR is that a user is not separated from the physical world, and can continue to interact with that environment. However, this poses a problem for collaboration, as again we have a different virtual environment per user that we must somehow combine. The typical solution is that a person in AR can only be joined by others in VR [4, 6, 13].

In this paper, we present a novel technique for virtual co-location and collaboration which allows each user to operate in their own virtual environment. The technique operates in real-time with little overhead. There are applications in virtual and augmented reality, though this paper uses virtual reality for demonstration. Remote users are placed in the local user's environment through pre-generated mappings between the spaces. As the user is not moved, no conflicts are introduced between the visual and vestibular system. However, the technique is dependent upon generating good mappings between environments. If environments are very different remote users may appear to move too quickly or to not quite interact with objects.

To evaluate how disruptive these effects might be we applied the technique to a pair of sample environments and ran a user study to determine whether users were aware of the mapping artefacts.

2 RELATED WORK

A great many methods for adapting a physical space to an immersive environment exist and are currently in use. A full survey is beyond the scope of this paper, but we do consider those which are commonly used and bear a resemblance to the technique proposed.

The simplest movement methods are those where movement is achieved through operation of a joystick, wand or gaze. Respectively, the user is moved in the direction they push a joystick, point or look. This allows a user to move independently of their available physical space. However, when compared with true locomotion

simulator sickness is increased and there is a negative impact on presence [15, 25]. In addition, the increased cognitive load has been shown to increase the time taken and the number of errors generated when conducting tasks in a virtual environment [22, 27].

Teleportation allows a user to move immediately to a selected location in the virtual environment. Though this has been shown not to cause significant simulator sickness [9], user disorientation is increased [3]. Environment recollection and navigation are also negatively affected when compared with real walking [15]. Walking-in-place is another technique which uses physical movement such as arm swinging as an analogy to true locomotion, and has been shown to be a significant improvement over joystick controls in presence and simulator sickness [18]. However, true locomotion has proven to be more desirable on both counts [25].

As real walking has so many positive characteristics in virtual reality, there is the motivation for techniques which use the limited physical space more intelligently. This allows a real walking interface to cover a much greater distance. Interrante et al. propose *Seven League Boots* [10], an approach which magnifies physical movements. Magnification is applied only in the user's movement direction so as not to exaggerate head sway. In practice, however, movement direction is hard to estimate, so the effect is disorientating while stationary or performing tasks [26].

Redirected walking takes this approach further, applying gain to translation but also to a user's rotation as they walk [14]. Thresholds have been found beneath which users do not perceive that gain is being applied [20]. Redirected walking has been proven to work well, but unless the user's path is known in advance a great deal of space is required. Steinickie et al. found that for users to believe they are walking in a straight line while really walking in a circular arc, the radius of that arc must be no less than 22 meters [19].

The technique used in this paper for adaptation to the virtual space is inspired by Sun et al., who propose computing a planar map between the virtual and physical floor plans [23]. The map is pre-calculated to place the virtual space within the available physical space and is surjective but not injective, potentially overlapping the virtual space on the physical many times. At run-time the map is consulted to render the virtual environment in the space of the physical environment.

The technique proposed by Sun et al. bears a resemblance to redirected walking, but rather than calculating redirections at runtime they are provided by the static map. This has the great advantage over redirected walking that the user will never leave the track space. However, unlike redirected walking, mapping can not be limited to perceptual thresholds. We can attempt to minimize distortion during the map generation stage, but as we do not know how fast the user will be moving we can not know how much redirection is being applied. For redirected walking, high gains are noticeable to users and lead to simulator sickness [20]. To avoid this problem, we generate two static forward mappings between each pair of spaces (A to B and B to A), and apply the map to user's avatars rather than the users themselves. As a result, users need not deal with any perceptual distortion or redirection, but still appear to be within the other user's space.

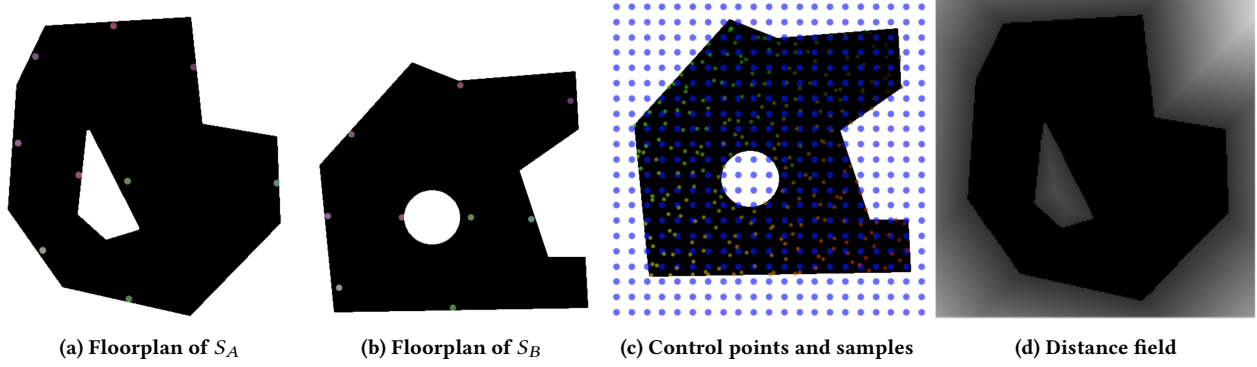


Figure 2: Static mapping optimization. Figures (a) and (b) contain the binary floor plans for room S_A and S_B . Correspondence points are colored for visualization. Figure (c) shows the sampled control points (in blue) and sampled positions used in Equation 3 (colored by 2D position). Figure (d) is the distance field described in Equation 4.

3 METHOD

Our goal is to allow two users (A and B) to interact with each other between two working spaces (e.g. room S_A and room S_B) with different layouts. To do this, we map A 's position in S_A to S_B , reconstruct their rotation and render their avatar in S_B . This process is repeated for B . We break this into two steps, separating position and rotation.

$$\mathbf{x}_{B,A} = f_{B \rightarrow A}(\mathbf{x}_{B,B}) \quad (1)$$

Where $\mathbf{x}_{B,A}$ is the 2D position of B in S_A and $f_{B \rightarrow A}$ is a pre-generated forward mapping from S_B to S_A . We assume the walkable areas in S_A and S_B are flat so ignore height for position mapping.

$$\mathbf{y}_{B,A} = g_{B \rightarrow A}(\mathbf{y}_{B,B}) \quad (2)$$

Where $\mathbf{y}_{B,A}$ represents the rotation of B in S_A and $g_{B \rightarrow A}$ is the function that reconstructs rotation across rooms.

We represent a 2D floorplan of a room in a binary image as shown in Figure 2a. Black pixels refer to free space where user can walk through while white pixels refer to obstacles, i.e. wall or furniture in the center of the room. In order to enable user interactions, we allow the users to add pre-defined corresponding locations between S_A and S_B , as shown in Figure 2a and 2b.

3.1 Positional Mapping

In this section we describe how to establish a forward mapping $f_{B \rightarrow A}$ from a 2D floorplan S_B to another 2D floorplan S_A . We expect the mapping $f_{B \rightarrow A}$ to (1) map any interior locations of S_B to a interior location of S_A ; (2) minimize the distance between corresponding positions; (3) minimize the angle and distance distortion; and (4) ensure local bijective for local smoothness. Inspired by a recent work [24], we represent our mapping $f_{B \rightarrow A}$ in a basis-function form and format our objectives and constrains into an optimization.

Forward mapping. We first uniformly sample N center points from S_B in 2D for our basis functions, $\mathbf{x}_i (i = 1, \dots, N)$, as shown in Figure 2c. Consider a 2D position \mathbf{x} in S_B , we construct our mapping $f_{B \rightarrow A}$ as:

$$f_{B \rightarrow A}(\mathbf{x}) = \sum_{i=1}^N w_i K_i(\mathbf{x}) + T(\mathbf{x}) \quad (3)$$

where w_i are the directions of basis functions K_i and $T(\cdot)$ is a 2D affine transformation. We chose a Gaussian function for our basis function where $K_i = \exp(-\frac{\|\mathbf{x} - \mathbf{x}_i\|^2}{2\sigma^2})$.

Objective function. First of all, we densely sample a subset of positions $\mathbf{u}_i (i = 1, \dots, M)$ from the free space of S_B (as shown in Figure 2c) to reduce our computational cost. Since we want our mapping $f_{B \rightarrow A}$ to be feasible as it will only map a free position (black pixel) from S_B to a free position in S_A . To efficiently achieve this, we convert this hard constraint to a soft one by casting it into an energy term. To do so, we first build a distance field D based on the binary floorplan of S_A , as shown in Figure 2d. In detail, we have:

$$D(\mathbf{x}) = \begin{cases} \min_{S_A(t)=0} (\|\mathbf{x} - \mathbf{t}\|), & S_A(\mathbf{x}) = 1 \\ 0, & S_A(\mathbf{x}) = 0 \end{cases} \quad (4)$$

Therefore, our term for obstacle free mapping is

$$E_{obs} = \sum_{i=1}^M D(f_{B \rightarrow A}(\mathbf{u}_i)) \quad (5)$$

To minimize the distance between two set of corresponding points, $\mathbf{C}_i^A \in S_A$ and $\mathbf{C}_i^B \in S_B$, where $i = 1, \dots, P$, we simply formulate this as a correspondence term

$$E_{corr} = \sum_{i=1}^P \|\mathbf{C}_i^A - f_{B \rightarrow A}(\mathbf{C}_i^B)\| \quad (6)$$

A locally isometric mapping requires its Jacobians J to satisfy $J^T J = 1$. To minimize the distance and angle distortion after applying $f_{B \rightarrow A}$, we convert this constraint into an energy term

$$E_{iso} = \|J^T J - 1\| \quad (7)$$

For local bijective, according to [16], the determinant of Jacobian of the inverse of $f_{B \rightarrow A}$ should be positive everywhere. We keep this as a constraint and our final optimization is

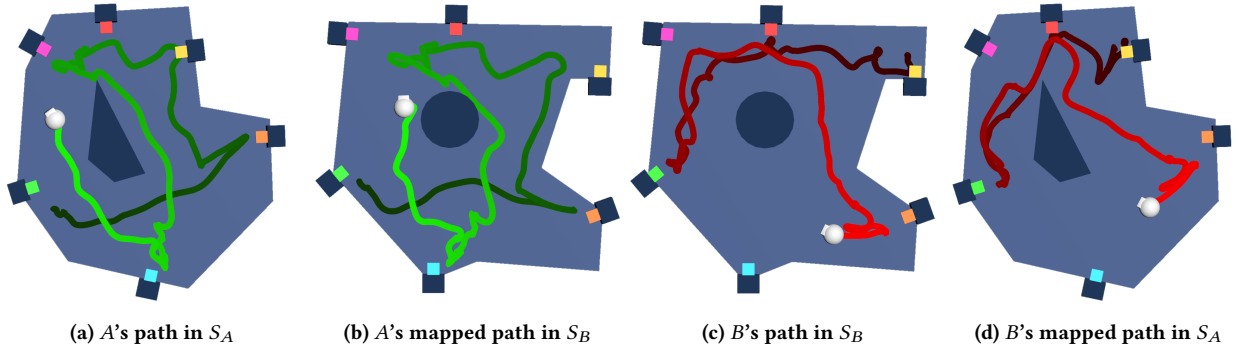


Figure 3: Demonstration of the effect of mapping on movement. Figures (a) and (c) show A and B 's movement in their own environments. Figures (b) and (d) show the mapped path of their avatars in the remote environment. Brighter lines mean more recent movement. We separate A and B 's paths into different images for clarity. In practice S_A would contain B 's avatar and A , while S_B would contain A 's avatar and B as seen in Figure 1.

$$E = E_{obs} + \lambda_{corr}E_{corr} + \lambda_{iso}E_{iso} \quad (8)$$

subject to the constraint

$$J_{-1}^T J_{-1} > 0 \quad (9)$$

When applied in real-time, the effect of the map on user position can be seen in Figure 3.

3.2 Reconstructing Rotation

After mapping B 's position into S_A B 's absolute rotation in S_B is no longer usable. It is likely that the relative direction to objects will not be preserved by the mapping process, so it may not be clear where B is looking. As described in Figure 4, we can counteract this by using B 's relative direction to objects in S_B to approximate B 's rotation in S_A .

$$g_{B \rightarrow A}(y_{B,B}) = y_{B,A} = (y_{forward,B,A}, y_{up,B,A}) \quad (10)$$

Where $y_{B,A}$ is a pair of 3D unit vectors (view-forward and view-up). To find these vectors we first select N corresponding points across rooms $\mathbf{p}_i (i = 1, \dots, N)$. These can be the correspondence points chosen for the mapping or the position of any object which is in both rooms. We chose to include the mapping correspondence points due to their significance in the task selected for the study (see Section 4.1). We also included the position of the remote user's head to allow for mutual gaze as a communication tool [1]. We calculate similarities $s_i (i = 1, \dots, N) \in [-1, 1]$ for each of the corresponding points based on angle between the points and the local user.

$$s_{i,B} = y_{forward,B,B} \cdot \frac{\mathbf{p}_{i,B} - \mathbf{x}_{B,B}}{\|\mathbf{p}_{i,B} - \mathbf{x}_{B,B}\|} \quad (11)$$

Where $\mathbf{p}_{i,B}$ is the position of point i in S_B , $s_{i,B}$ is B 's similarity to that point. We calculate the reconstructed forward vector for B in S_A as an average of the correspondence points in S_A weighted by similarities s .

$$y_{forward,B,A} = \sum_{i=1}^N w(s_i) \frac{\mathbf{p}_{i,A} - \mathbf{x}_{B,A}}{\|\mathbf{p}_{i,A} - \mathbf{x}_{B,A}\|} \quad (12)$$

Where w is the weighting function in $[0, 1]$. We found it helpful to use the non-linear weighting function $w(s_i) = (s_i * 0.5 + 0.5)^7$. This helped to give the impression that B was looking directly at objects in S_A . The particular degree chosen was based on informal testing. We found that lower degrees led to users appearing to look between objects, while higher degrees caused unrealistically fast head movement. To calculate the view up component of B 's orientation, we rotate to match $y_{forward,B,A}$, then further rotate by B 's head rotation on the local frontal plane. This captures the "roll" of B 's head, and is applied directly from tracking data with no mapping.

$$y_{up,B,A} = rotate(rotate(y_{up,B,A}, \theta_B), \theta_{B,roll}) \quad (13)$$

Where θ_B is the 3D rotation from the world forward vector to $y_{forward,B,A}$, $\theta_{B,roll}$ is B 's head rotation on the local frontal plane and $rotate$ is a function that applies 3D rotations.

3.3 Associated Object Positions

In this section we have discussed placing a user's head through mapping. However, some applications may require associated objects to move with a user such as their hands, feet, torso or held objects. In our tests for the accompanying study users were provided with virtual hands and bodies to help with communication [7]. Hand position and orientation was provided by tracking for a hand-held controller, while body position and orientation was generated from head and hand positions and orientations. We found that mapping objects individually did not produce good results as the mapping interfered with physical movements (e.g., gestures and locomotion) and made them appear unnatural. A superior approach is to preserve the relative position and orientation of associated objects after mapping.

$$M_{world \rightarrow x,A} = M_{B \rightarrow world,A} \times M_{x \rightarrow B} \quad (14)$$

Where x is an associated object belonging to B , and $M_{x \rightarrow B,A}$ is the homogeneous transformation matrix which describes position and orientation relative to x for positions and orientations relative to B 's calculated position in S_A . More simply, we find where x is relative to B in S_B , and place x there relative to B in S_A .

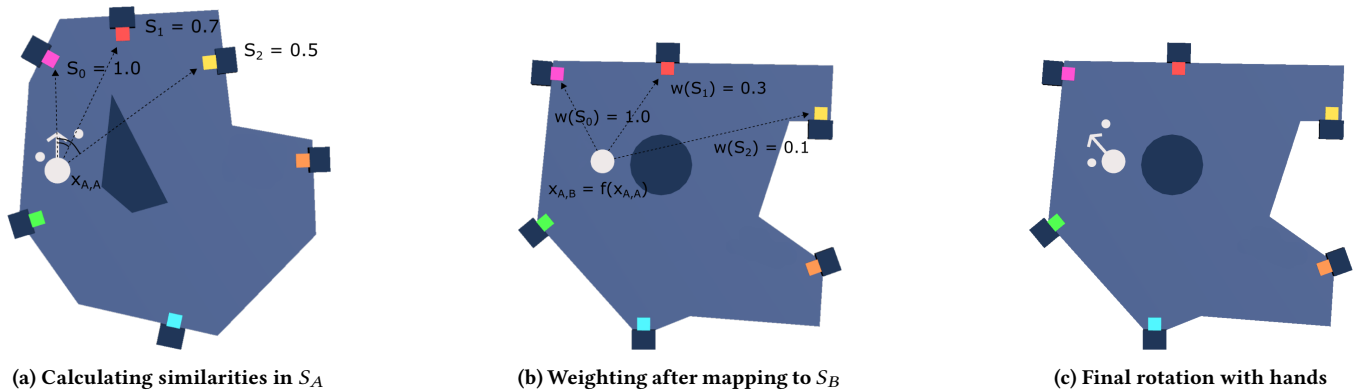


Figure 4: Calculating rotation and associated object positions. Similarities are calculated as a function of the angle between the user’s gaze direction and the direction to each correspondence point and to the other user’s face in S_A . After mapping the user’s position to S_B , a weighting function is applied to each similarity and the weighted direction vectors are summed to find the new gaze direction. The relative positions of the hands (the small white circles in the Figure) to the user are maintained.

4 USER STUDY

The technique proposed in this paper modifies the remote user’s physical movements and must reconstruct orientation information. Each user sees a mostly consistent view of the other’s actions. There are two potential down-sides. First the actions may not appear plausible in the other user’s space. In particular, users may appear to move too quickly or too slowly and move in curves rather than straight lines. Secondly, features of interest will be positioned differently across the spaces. Users might appear to not be looking or standing where they think they are looking or standing, and appear to be not quite touching or holding something they should be. The following user study was conducted to prove the concept and to determine if users would notice these effects in practice.

4.1 Study Design

Participants were invited to work together in pairs to complete a short puzzle in virtual reality. The time limit for the task was 5 minutes. Each participant was placed in their own virtual environment, which was designed to fit within the track space used for the study. The virtual environments for the two participants had a different floor plan and different points of interest (see Figure 2). Mappings were generated before the study as described in Section 3.1.

An avatar representing each participant’s head, hands and torso was created. Participants were able to see the full avatar of the remote user, but not their own head or body. The position of the remote participant was provided over the network and the remote participant’s avatar placed in the local user’s space as described in Section 3.1. Orientation of the remote participant was then reconstructed as described in Section 3.2. Participants were not informed mapping was taking place, or that the virtual environments were different.

Each room was surrounded with a set of buttons on small pillars. Each button was a different color. Either participant could push a button by moving their hand towards it. Once pushed, a button would be depressed for approximately 2 seconds before returning to its default position. Buttons were paired secretly before the

experiment began. If two paired buttons were pushed within 0.2 seconds of each other, a sound would play and the button pair would lock, not returning to their default position. If all buttons were locked the sequence in which they were pressed would be recorded, a chime would play and all buttons would return to neutral. For each new sequence, a new chime would play, up to the maximum of 6.

The design of the task was intended to (1) create a cognitive load, (2) encourage participants to move around and (3) draw attention to the movement of the remote participant. The objective was that the task represent a challenging but realistic scenario for the technique, where users are collaborating on a task with a significant spatial component.

4.2 Outcome Measures

Participants were asked to complete questionnaires before and after the study. The pre-questionnaire contained questions were on gender, age, occupation and how many times the participant had used virtual reality. We were interested to examine whether a greater familiarity with virtual reality might make the technique more noticeable to users so chose to investigate this as a between-subjects variable.

The first section of the post-questionnaire combined questions on co-presence [2] and embodiment [11] with the SUS presence questionnaire [17]. As distortion was only applied to user’s remote avatars and not to the users themselves, we did not require participants to answer questions on simulator sickness. The second section contained a set of questions specifically designed to assess whether the participant had noticed disruptive effects of the mapping. Participants were asked about a set of phenomena that did not occur during the study, with a statement describing “real” phenomena embedded amongst them. This is similar to the approach taken by Peck et al. [12] and Suma et al. [21].

The embedded question section of the questionnaire was preceded with the statement “This section is on unnatural phenomena you may have noticed in the virtual environment. Please rate the following statements. Note that these phenomena may or may not

have happened.". Participants were asked to rate each of the statements on a scale of 1 to 7, where 1 meant "did not notice or did not happen" and 7 meant "very obvious". One of the statements (below in italics) refers to a side effect of the technique that participants may have noticed where mapping was poor. The rest are "decoy" statements. The phenomena described in decoy statements were not present in the virtual environment. In the order presented to the participant, the statements were:

- The floor seemed to change colour
- The floor seemed to move beneath me
- The other person seemed to change shape
- *The other person seemed not to know where things were*
- I felt as if the world was changing shape
- I felt as if the world was rotated around me
- Objects in the world seemed to move without being touched
- Objects in the world seemed to change colour

Finally, participants were asked a series of open questions about the virtual experience. We were primarily interested in whether participants had found that the technique negatively affected their experience. To this end, participants were asked "what aspects of the virtual experience did you find disruptive or unpleasant?" and "what aspects of the virtual experience detracted from it as a whole?". We also included an open-ended question on unusual phenomena to see if participants had observed any artefacts of the technique in a way that had not been detected by the embedded statement.

4.3 Study Setup

Participants were equipped with the HTC Vive head-mounted display and hand controllers. Both provide tracking for position and orientation. The display has a resolution of 1080 x 1200 per eye, a refresh rate of 90hz, and a vertical field of view of 110 degrees. The HMDs were set to the approximate population mean inter-pupillary distance of 63mm [8]. Participants wore noise-canceling headphones. A voice-over-IP application was used with the built-in microphones in the Vive and set up so participants were able to converse naturally without pressing any buttons.

The room was divided into two 3.5m x 3.5m track spaces, with a 1m gap between track spaces to prevent collisions. The Vive uses external base stations for tracking, so a curtain was placed between the two track spaces to prevent interference. The virtual environments were developed in Unity3D and used the HLAPI for networking and the SteamVR plugin for VR. The systems were connected via ethernet over a local area network. Each Vive was powered by an NVIDIA GTX 1080 and maintained the maximum refresh rate of 90 frames per second throughout. Mappings were generated between the environments on an Intel Core i7 CPU with 8GB memory and took approximately an hour. The maps and environments used for the study are available online [5].

4.4 Study Protocol

The study took approximately 30 minutes per pair of participants. Participants were invited in teams of 2. On arrival, participants were given an information sheet describing the study. After an

opportunity to ask questions, the experimenter gave each participant a consent form to sign. After consent was obtained the experimenter asked participants to individually complete a short pre-questionnaire (see Section 4.2).

The experimenter showed participants to their track space and provided each with a head-mounted display, a pair of controllers and a pair of noise-canceling headphones. The experimenter loaded a demo environment to help participants familiarize themselves with the devices and with a real-walking movement interface. Participants were shown how to identify when they were approaching the edge of the track space. The experimenter then connected the participants over voice chat and tested audio levels.

When both participants were comfortable the experimenter loaded the study environment. The experimenter connected the two sessions, spawning the rooms and avatars for both participants. The experimenter started a timer and instructed participants that their goal was to "hear 6 chimes". After participants completed the task or 5 minutes elapsed, the experimenter ended the task by bringing participants back to the demo environment and instructing them to remove their headsets.

After participants removed their headsets they were asked to fill out the post-questionnaire (see Section 4.1). Finally, participants were debriefed and given the opportunity to ask questions and give feedback.

4.5 Participants

38 participants (23 female, 15 male) were recruited. The mean age was 26.02 with a standard deviation of 6.13. Participants were recruited from graduate study mailing lists and external advertisements and paid a small amount to cover travel expenses. Participants were required to be between the ages 18 and 65 and to be able to walk unassisted. This study was approved by the University College London Research Ethics Committee.

4.6 Results

Figure 5 shows the mean responses to the decoy questions and the embedded question across all participants. Means across all questions were contained within a relatively small range, from lowest ($M = 1.66$) to highest ($M = 2.26$), indicating no statement particularly stood out. The highest rated statement was "I felt as if the world was rotated around me" ($M = 2.26$, $SD = 1.46$). The lowest was the embedded statement, "The other person seemed not to know where things were" ($M = 1.66$, $SD = 1.15$).

As per the methodology used for the original SUS, we invite answers to questions on a 7 point scale. The score for each participant is the number of responses in the 5-7 range, which we then divided by the number of questions. We asked questions on presence ($M = .61$, $SD = .27$), co-presence ($M = .76$, $SD = .38$) and embodiment ($M = .65$, $SD = .30$). Of the 38 participants, 7 scored the maximum on presence, 26 scored the maximum on co-presence and 8 scored the maximum on embodiment.

Participants were broken into 3 groups based on the self-reported number of times each had used virtual reality ("never" with 14 participants, "1-2 times" with 16 participants, or "3+ times" with 8 participants). A one way univariate ANOVA was conducted to test for the effect of previous virtual reality experience on ratings for

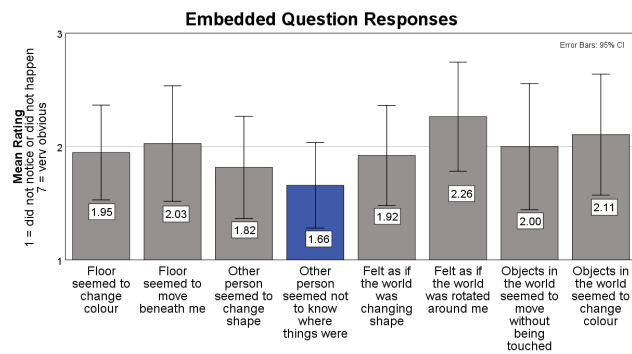


Figure 5: Mean participant responses to decoy questions (in grey) and the embedded question (in blue). Note that participants were asked to rate phenomena on a scale of 1 to 7, which has been capped to 3 here for clarity. The rating for the embedded question is particularly low, indicating that users believed the remote user knew where objects were in the local room.

the decoy questions and embedded question. No significant effect was found for the embedded question on experienced users ($M = 1.75$, $SD = 1.04$) over occasional users ($M = 1.5$, $SD = .89$) or those who had never used VR ($M = 1.79$, $SD = 1.48$), $F(2,36) = .254$, $p = .78$. No significant effect was found for VR experience on any of the decoy questions. No significant effect was found for VR experience on presence ($p = .07$) or co-presence ($p = .505$).

4.7 Discussion

The results appear promising as very few participants appeared to be aware of artefacts caused by the mapping. In response to the open ended questions 1 participant described an incident where "my partner was once floating above the obstacle in the middle". This may have been caused by an area in the map which optimized poorly. Possibly, though, it was caused by the remote participant walking through an obstacle in their own area without realizing, as there was no collision to prevent this in the study. Of the 38 participants 2 identified issues that were likely caused by the mapping. The first noted that "when the other player pressed a button, it looked like they were missing on my screen", while the second participant noticed their partner appeared to miss "when we tried to high-five".

It makes sense that participants would notice the latter two incidents as hand placement is a particularly difficult task for the technique. As described in Section 3.3, we preserve the position and orientation of hands across rooms to allow for gesture. However, this means the static map is not directly applied, so even at points where the two mappings agree (the pre-selected correspondence points, e.g., the buttons in the study) hand positions will appear to not quite match. If we were to solve this problem, virtual high-fiving still poses a challenge. The maps are static so hands cannot be included as correspondence points. This means the maps do not have the information necessary to make hands coincide. These distortions are likely to remain small even for very different rooms as they are limited by the distance between the user and their hands.

Co-presence scores were particularly high when compared with other metrics. Though we did not specifically ask for thoughts on co-presence in the open ended questions, responses to our final question "What are your general thoughts on the virtual experience?" received a number of positive comments such as "I enjoy the feeling that someone is together with me" and "it was great that I could do it with someone else".

The embedded statement was rated lower than all the decoy statements, which suggests participants were confident the remote user knew where objects were in their local rooms. We were surprised to find that in the results we collected VR experience had no significant effect on responses to the embedded statement. Our assumption was that new VR users would dismiss the artefacts as a limitation of VR technology. Though the task required co-operation and discussion, participants were not strictly required to look at one another, and this may have had a more significant impact on whether or not distortions were recognized.

5 CONCLUSION AND FUTURE WORK

In this paper, we described using mapping for space sharing in mixed reality. This is a novel technique for virtual co-location which allows each user to experience their own environment while appearing to interact with other users in their own spaces. The technique has applications in augmented and virtual reality. Our user study indicates the concept works in practice, with only 2 of 38 participants noticing mapping artefacts. The importance of mapping artefacts may also be arguable as users become more experienced, as all participants managed to solve some portion of the puzzle game. We conclude that mapping between spaces is a promising approach to remote co-location. However, while it seems that distortions went largely unnoticed in our study, it remains to be seen how well the technique will work on environments that are less similar. The technique also currently requires correspondence points to be manually identified, and an expensive optimization process to generate the static maps. This latter cost could be reduced by using pre-calculated partial maps between track spaces, reducing the final optimization time per room. Positional mapping and rotation reconstruction appeared to work well, but associated objects (and hands in particular) were the cause of many visible mapping artefacts and would benefit from a more sophisticated solution.

ACKNOWLEDGMENTS

The research was supported by the UCL Centre for Virtual Environments, Interaction & Visualisation funded by the UK's EPSRC under project number EP/G037159/1. The authors would like to thank Sebastian Friston and David Swapp for their suggestions and comments.

REFERENCES

- [1] Michael Argyle and Mark Cook. 1976. *Gaze and mutual gaze*. Cambridge U Press, Oxford, England. xi, 210–xi, 210 pages.
- [2] Gagatay Basdogan, Chih-Hao Ho, Mandayam A. Srinivasan, and Mel Slater. 2000. An Experimental Study on the Role of Touch in Shared Virtual Environments. *ACM Trans. Comput.-Hum. Interact.* 7, 4 (Dec. 2000), 443–460. <https://doi.org/10.1145/365058.365082>
- [3] Doug A Bowman, David Koller, and Larry F Hodges. 1997. Travel in Immersive Virtual Environments: An Evaluation Motion Control Techniques of Viewpoint. *Proc. IEEE Virtual Reality Annual Symposium (1997)*. <https://doi.org/10.1109/VRAIS.1997.583043>

- [4] M. L. Chenechal, T. Duval, V. Gouranton, J. Royan, and B. Arnaldi. 2016. Vishnu: virtual immersive support for HelpiNg users an interaction paradigm for collaborative remote guiding in mixed reality. In *2016 IEEE Third VR International Workshop on Collaborative Virtual Environments (3DCVE)*. 9–12. <https://doi.org/10.1109/3DCVE.2016.7563559>
- [5] Ben Congdon. 2018. Accompanying Data for "Merging Environments For Shared Spaces in Virtual Reality". (2018). <https://doi.org/10.6084/m9.figshare.7154702.v1>
- [6] Matthew Christopher Davis, Dang D. Can, Jonathan Pindrik, Brandon G. Rocque, and James M. Johnston. 2016. Virtual Interactive Presence in Global Surgical Education: International Collaboration Through Augmented Reality. *World Neurosurgery* 86 (2016), 103 – 111. <https://doi.org/10.1016/j.wneu.2015.08.053>
- [7] Trevor J. Dodds, Betty J. Mohler, and Heinrich H. Bühlhoff. 2011. Talk to the virtual hands: Self-animated avatars improve communication in head-mounted display virtual environments. *PLoS ONE* (2011). <https://doi.org/10.1371/journal.pone.0025759>
- [8] Neil A Dodgson. 2004. Variation and extrema of human interpupillary distance. In *Proc. SPIE 5291*. 36–46.
- [9] Sebastian Freitag, Dominik Rausch, and Torsten Kuhlen. 2014. Reorientation in virtual environments using interactive portals. In *2014 IEEE Symposium on 3D User Interfaces (3DUI)*. 119–122. <https://doi.org/10.1109/3DUI.2014.6798852>
- [10] Victoria Interrante, Brian Ries, and Lee Anderson. 2007. Seven league boots: A new metaphor for augmented locomotion through moderately large scale immersive virtual environments. In *IEEE Symposium on 3D User Interfaces 2007 - Proceedings, 3DUI 2007*. 167–170. <https://doi.org/10.1109/3DUI.2007.340791>
- [11] Elena Kokkinara, Konstantina Kilteni, Kristopher J. Blom, and Mel Slater. 2016. First Person Perspective of Seated Participants Over a Walking Virtual Body Leads to Illusory Agency Over the Walking. 6 (07 2016), 28879.
- [12] Tabitha C. Peck, Henry Fuchs, and Mary C. Whitton. 2009. Evaluation of reorientation techniques and distractors for walking in large virtual environments. *IEEE Transactions on Visualization and Computer Graphics* (2009). <https://doi.org/10.1109/TVCG.2008.191>
- [13] Thammathip Piumsomboon, Arindam Day, Barrett Ens, Youngho Lee, Gun Lee, and Mark Billingham. 2017. Exploring Enhancements for Remote Mixed Reality Collaboration. In *SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications (SA '17)*. ACM, New York, NY, USA, Article 16, 5 pages. <https://doi.org/10.1145/3132787.3139200>
- [14] Sharif Razzaque, Zachariah Kohn, and Mary C Whitton. 2001. Redirected Walking. *Proceedings of EUROGRAPHICS* (2001), 289–294.
- [15] Roy. A. Ruddle and S. Lessels. 2009. The benefits of using a walking interface to navigate virtual environments. *ACM Transactions on Computer-Human Interaction* 16, 1 (2009), 1–18. <https://doi.org/10.1145/1502800.1502805>
- [16] Christian Schüller, Ladislav Kavan, Daniele Panozzo, and Olga Sorkine-Hornung. 2013. Locally injective mappings. In *Proceedings of the Eleventh Eurographics/ACMSIGGRAPH Symposium on Geometry Processing*. Eurographics Association, 125–135.
- [17] Mel Slater, Martin Usoh, and Anthony Steed. 1994. Depth of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments* (1994). <https://doi.org/10.1162/pres.1994.3.2.130> arXiv:arXiv:1011.1669v3
- [18] Mel Slater, Martin Usoh, and Anthony Steed. 1995. Taking steps: the influence of a walking technique on presence in virtual reality. *ACM Transactions on Computer-Human Interaction* 2, 3 (1995), 201–219. <https://doi.org/10.1145/210079.210084> arXiv:arXiv:1011.1669v3
- [19] Frank Steinicke, Gerd Bruder, Jason Jerald, Harald Frenz, and Markus Lappe. 2010. Estimation of detection thresholds for redirected walking techniques. *IEEE Transactions on Visualization and Computer Graphics* 16, 1 (2010), 17–27. <https://doi.org/10.1109/TVCG.2009.62>
- [20] Frank Steinicke, Gerd Bruder, Jason Jerald, Harald Frenz, Markus Lappe, and Harald Frenz. 2008. Analyses of Human Sensitivity to Redirected Walking. *VrSt* 146, 3 (2008), 149–156. <https://doi.org/10.1145/1450579.1450611>
- [21] Evan A. Suma, Seth Clark, David Krum, Samantha Finkelstein, Mark Bolas, and Zachary Warte. 2011. Leveraging change blindness for redirection in virtual environments. In *Proceedings - IEEE Virtual Reality*. 159–166. <https://doi.org/10.1109/VR.2011.5759455>
- [22] Evan A. Suma, Samantha L. Finkelstein, Seth Clark, Paula Goolkasian, and Larry F. Hodges. 2010. Effects of travel technique and gender on a divided attention task in a virtual environment. In *3DUI 2010 - IEEE Symposium on 3D User Interfaces 2010, Proceedings*. 27–34. <https://doi.org/10.1109/3DUI.2010.5444726>
- [23] Qi Sun, Li-Yi Wei, and Arie Kaufman. 2016. Mapping virtual and physical reality. *ACM Transactions on Graphics* (2016). <https://doi.org/10.1145/2897824.2925883>
- [24] Qi Sun, Li-Yi Wei, and Arie Kaufman. 2016. Mapping Virtual and Physical Reality. *ACM Trans. Graph.* 35, 4, Article 64 (July 2016), 12 pages. <https://doi.org/10.1145/2897824.2925883>
- [25] Martin Usoh, Kevin Arthur, Mary Whitton C., Rui Bastos, Anthony Steed, Mel Slater, and Frederick Brooks P. 1999. Walking > walking-in-place > flying, in virtual environments. *SIGGRAPH '99 Proceedings of the 26th annual conference on Computer graphics and interactive techniques* (1999), 359–364. <https://doi.org/10.1145/311535.311589>
- [26] Betsy Williams. 2008. *Design and evaluation of methods for motor exploration in large virtual environments with head-mounted display technology*. Ph.D. Dissertation. Vanderbilt University.
- [27] Catherine A. Zambaka, Benjamin C. Lok, Sabarish V. Babu, Amy C. Ulinski, and Larry F. Hodges. 2005. Comparison of path visualizations and cognitive measures relative to travel technique in a virtual environment. *IEEE Transactions on Visualization and Computer Graphics* 11, 6 (2005), 694–705. <https://doi.org/10.1109/TVCG.2005.92>