

Magnoramas: Magnifying Dioramas for Precise Annotations in Asymmetric 3D Teleconsultation

Kevin Yu*
Research Group MITI,
Hospital Recht der Isar
TUM, Munich, Germany

Dirk Wilhelm[¶]
Research Group MITI,
Hospital Recht der Isar
TUM, Munich, Germany

Alexander Winkler[†]
Chair of Computer Aided
Medical Procedures
TUM, Munich, Germany

Ulrich Eck[¶]
Chair of Computer Aided
Medical Procedures
TUM, Munich, Germany

Frieder Pankratz[‡]
Institute for Emergency
Medicine
LMU, Munich, Germany

Daniel Roth^{**}
Chair of Computer Aided
Medical Procedures
TUM, Munich, Germany

Marc Lazarovici[§]
Institute for Emergency
Medicine
LMU, Munich, Germany

Nassir Navab^{††}
Chair of Computer Aided
Medical Procedures
TUM, Munich, Germany

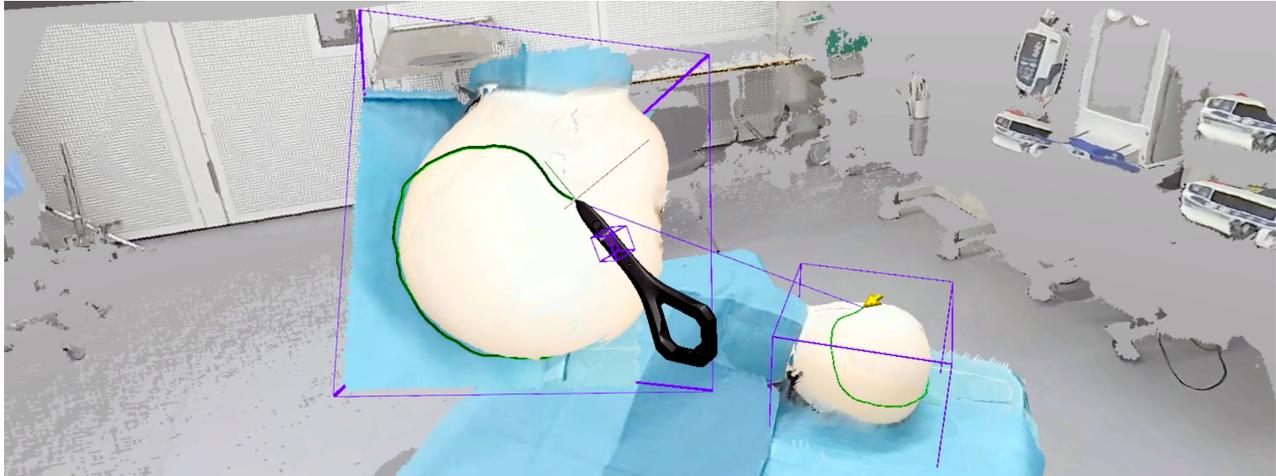


Figure 1: **Our proposed method: Magnorama.** Magnoramas allow the flexible extraction, transformation, and annotation of a region of interest (right) inside the real-time captured point cloud. A Magnorama can be interactively positioned, rotated, and scaled by the user. Increasing the size of the Magnorama provides the user with a magnified view of the region of interest. By that, it supernaturally augments the precision of annotations while remaining in the scene context.

ABSTRACT

When users create hand-drawn annotations in Virtual Reality they often reach their physical limits in terms of precision, especially if the region to be annotated is small. One intuitive solution employs magnification beyond natural scale. However, scaling the whole environment results in wrong assumptions about the coherence between physical and virtual space. In this paper, we introduce *Magnoramas*, a novel interaction method for selecting and extracting a region of interest that the user can subsequently scale and transform inside the virtual space. Our technique enhances the user's capabilities to perform supernaturally precise virtual annotations on virtual objects. We explored our technique in a user study within a

simplified clinical scenario of a teleconsultation-supported craniectomy procedure that requires accurate annotations on a human head. Teleconsultation was performed asymmetrically between a remote expert in Virtual Reality that collaborated with a local user through Augmented Reality. The remote expert operates inside a reconstructed environment, captured from RGB-D sensors at the local site, and is embodied by an avatar to establish co-presence. The results show that Magnoramas significantly improve the precision of annotations while preserving usability and perceived presence measures compared to the baseline method. By hiding the 3D reconstruction while keeping the Magnorama, users can intentionally choose to lower their perceived social presence and focus on their tasks.

Keywords: Interaction techniques, medical information system, virtual reality.

Index Terms: [3D user interaction]: Human factors and ergonomics—Teleoperation and telepresence;

1 INTRODUCTION

Immersive virtual environments hold great potential to support collaborative and assistive tasks, such as joint exploration [1] or collaborative medical procedures [2]. They can provide avatar embodiment [3, 4] and augmented forms of interaction in ways that would not be possible in the physical world or traditional media [5, 6, 7, 8].

*e-mail: kevin.yu@tum.de

[†]e-mail: alexander.winkler@tum.de

[‡]e-mail: frieder.pankratz@med.uni-muenchen.de

[§]e-mail: marc.lazarovici@med.lmu.de

[¶]e-mail: dirk.wilhelm@tum.de

[¶]e-mail: ulrich.eck@tum.de

^{**}e-mail: daniel.roth@tum.de

^{††}e-mail: nassir.navab@tum.de

Such collaborative environments can consist of purely virtual environments (e.g., [9, 10, 11, 12]), augmented environments (e.g., [1, 13]) or asymmetric combinations that merge virtual as well as augmented reality aspects (e.g., [14, 15, 16, 8, 17, 18, 19], see [20, 21] for further systematic reviews). One of the applications of the latter class is telepresence and, more specifically, teleconsultation [2, 18, 22], in which two or more users, physically apart from each other, can interact and guide another through a specific procedure.

To provide a mixed reality or asymmetric teleconsultation, as in the case of medical emergencies, it is necessary to provide bidirectional communication, visualization or replication of the situation, and context [22, 2], for example, by 3D reconstruction [23, 24]. Despite recent progress, remote collaboration in virtual- or mixed reality scenarios still faces several challenges that consider the coherence of shared environments (and the relation to the physical space), such as sharing awareness [25] or avoiding collisions [26].

Moreover, it can rationally be assumed that interactions, especially drawing in 3D reconstructions, are error-prone either due to the technical artifacts of noise and reconstruction errors or, more importantly, the lower accuracy when compared to drawing with a physical pen and paper, which arises from the lack of physical support [27] and the fact that drawing in 3D has higher manual effort or cognitive and sensorimotor demands [28].

In medical teleconsultation, however, precision in interaction and guidance can be critical to ensure a patient’s survival. Surgeons, paramedics, and first responders are likely to encounter injuries in which immediate treatment is of paramount importance. Still, they might not be trained to or not possess enough experience to perform certain interventions. Even trauma surgeons may lack specialized skills for specific procedures. One of these emergency surgery procedures is craniectomy, where the patient’s skull needs to be opened to promptly release pressure from a swelling of the brain. Teleconsultation may be used for marking the steps to perform the craniectomy but requires exceptionally accurate annotations as guidance, which would directly relate to interventional incisions.

Little work has explored how to support such high-precision interactions while ensuring important communicative functions for successful collaboration. In this regard, common tools for consultation include virtual avatars that can point and gesture at real-world objects or draw mid-air annotations, which both users can see. When users need to draw a line accurately, an intuitive approach is moving closer. However, moving closer may still limit the precision capabilities due to factors such as jitters of the drawing device resulting from tracking errors. These reduced sensorimotor capabilities will become apparent in mid-air drawing tasks. Yet, adapting common methods from purely virtual applications for sketching (e.g., VRSketch) or drawing (e.g., TiltBrush), such as scaling the whole environment depicted as a real-time point cloud, would most likely result in sickness effects, mislead the perception of size and spatial relation between objects and context, and may hinder necessary communicative interactions such as mutual gaze, joint attention [29] and shared awareness [25].

1.1 Contribution

To address this problem, we propose Magnoramas (see Figure 1, which can be described as interactive dioramas for selectively magnifying regions of interest of a real-time 3D reconstruction for remote teleconsultation. We describe our approach and compare our method to i) a baseline and ii) to a variant of our method where users can only see the Magnorama, but the context is masked. Our method outperforms the baseline in terms of precision while having similar usability and task load ratings, thus providing initial evidence for the applicability. Our findings show that removing the scene context (and hence the partner’s avatar) reduces social

presence. This novel interaction method and its evaluation provide valuable insights, demonstrate high potential, and guide the design of future telepresence systems.

2 RELATED WORK

We divide the related work into three major categories that present the related context and previous work for our approach: (i) Virtual Reality (VR) interaction with a World-In-Miniature (WiM), (ii) drawing precision in VR, and (iii) co-interaction between multiple parties during teleconsultation.

2.1 Interaction with a World-In-Miniature

The well-known work on WiMs by Stoakley et al. [30] follows a related concept and utilizes a tracked physical clipboard. In VR, the entire room is down-scaled and attached to the clipboard inside the virtual environment (VE). Users could move furniture in the miniaturized version and observe the original furniture moving inside the actual room. The authors recognize the potential of enlarging the WiM for more fine-grain control of manipulation in exchange for range. However, to the best of our knowledge, they do not follow up on this idea and neglect the potential of detail selection and improved precision.

In the follow-up works, the metaphor of WiM is primarily researched for interaction [31], spatial locomotion and navigation [32, 33, 34, 35, 36]. Wingrave et al. [37] added scaling and scrolling functionality to the WiM and investigated the use of WiMs for spatial navigation. They, however, used scaling only to shrink the environment. They found that users rarely re-scale the WiM and often leave it at a comfortable size.

Pierce et al. [38] introduce an interaction method alluding to Voodoo-dolls, which applies the idea of indirect object manipulation that is present as well in a WiM. In this method, users can create copies of virtual objects. Any interaction performed on the copy is simultaneously performed on the original object.

In contrast, our method allows such interactions on any virtual subspace and its content, therefore, is not limited to particular virtual objects. Additionally, no magnification was used for their method.

While WiMs and Magnoramas have common characteristics – such as the duplicated view on virtual space and indirect manipulation – the core aspect is that the scaling factor is inverted.

2.2 Precision of Drawing and Annotating

In the following section, we provide an overview of related work regarding (the improvement of) freehand drawing and annotation inside a three dimensional (3D) VR environment. In this regard, we consider drawing to be a context free action and annotating to be an object-centered application of drawing or object placements/attachments. One of the re-occurring difficulties during unconstrained drawing in a 3D environment is the inclusion of the third dimension. One common pitfall is the misjudgment of depth such that drawn line strokes may appear closer or farther than intended by the user [39]. Additionally, drawing a straight line poses a challenge since no haptic feedback nor cognitive, nor sensorimotor aids are provided, unlike drawing on a physical surface [40, 27, 28]. Multiple related works investigated the assistance in freehand 3D manipulations or drawing with a tracked pen, either by including purely visual non-constraining guides [41, 42], constraining guides [43], or haptic modalities [40, 44].

Barrera et al. [45] investigate the relationship between spatial ability and the user’s proficiency to redraw a given shape in VR. They found that helping the user identify the correct viewpoint and starting point of the next stroke positively affects line precision. Additionally, they conclude that dynamic viewpoints and feedback on head-movements via a compass or a map can improve the sense of depth.

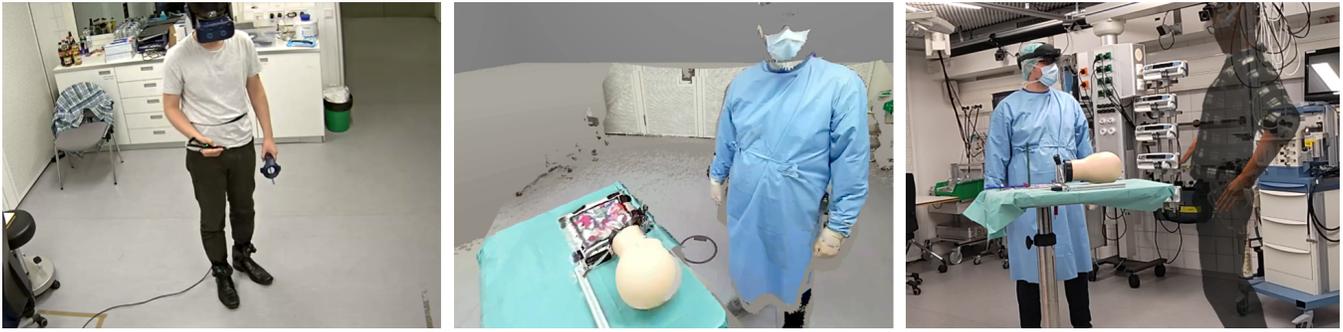


Figure 2: **A scenario using the asymmetric telepresence system.** From left to right: The remote expert wearing the VR tracking setup for animating his avatar and allowing annotations (left). A view on the local user from the first-person perspective of the remote expert in VR (center). The HoloLens 2 is not visible in the point cloud due to the high reflection coefficient of the transparent display. 3rd person AR view on the avatar of the remote expert and the local user in the shared environment (right).

Since drawings in these works and other VR applications are anchored statically in the virtual space, users changed their body position to gain a different perspective. A Magnorama is a cut-out of the drawing region. Users can transform it with their hands and quickly change their point of view to gain a better understanding of the geometry as well as to object details. Simultaneously, the final drawing results will not be changed in position and maintain the spatial correctness.

2.3 Teleconsultation and Collaborative Interaction

Collaborative virtual environment approaches can be distinguished between avatar-mediated systems (e.g., [46, 47, 6, 48]), 3D-reconstruction-based telepresence approaches (e.g., [13, 49, 50, 51, 24, 1]), and mixed/asymmetric approaches (e.g., [17, 8]). These provide the basic context for an application use-case. Research in object manipulation, shared drawings, or annotations for remote guidance is central to teleconsultation systems. The next paragraphs discuss methods in a shared teleconsultation system using annotations.

Oda et al. [16] present a method for VR/Augmented Reality (AR) teleconsultation to guide the consultee in placing a real physical object onto a second real physical object. They introduce the concept of Virtual Replicas, which is an instantiated virtual version of the physical object to be moved. By defining points on the virtual replica, the consultant can create links connecting the physical object to the replica. Unlike our method, virtual replica require knowledge and 3D model of the annotated object and does not provide methods on increasing the precision while defining the annotations. Oda and colleagues further [52] use a cutout from a real-time captured point cloud in a collaborative AR collaboration system for a more precise pointing gesture of distant objects. Kolkmeier et al. [53] use an RGB-D camera to capture the 3D environment of the consultee in real-time and visualize it for the consultant inside a VR head-mounted display (HMD). Their presented work incorporates a real-time captured point cloud and an abstracted user representation (head and hands) of the consultant drawing annotations. Weibel et al. [2] present ARTEMIS, an asymmetric telepresence system using VR and AR HMDs. Drawing annotations is possible in this system as well but with no additional solution for increased precision.

These works indicate the need for precise annotations in teleconsultation systems. However, none of the the systems helps the users to draw annotations that are more accurate than they could achieve with their natural dexterity, which presents a gap in research.

2.4 Hypotheses

Our review shed light on three major areas of related work. Our research goal was to provide a method that would successfully assist the presented use-case or related requirements. Reviewing the literature on drawing precision and projecting the findings on our proposed method, we assumed that **H1: The magnification of details in the form of Magnoramas increases freehand drawing precision** since Magnoramas aim to improve information detail but also act as “lever” for motor precision. Further, since the interaction method is novel and less natural than more coherent interaction, we also assumed that **H2: Interacting with Magnoramas is inferior in terms of usability compared to context coherent interactions.** Finally, considering the importance of co-location, joint attention, communicative cues, and collaborative verbal and nonverbal interaction (broadly discussed e.g., [5, 54, 55, 8, 6]), one could fear that with our method **H3: The perception of the interaction in terms of co-presence and social presence aspects is inferior when using Magnoramas** since the remote user would change focus to other parts in the scene when modifying the Magnorama, or completely lose the context (Magnorama-only).

3 METHODS

We present a solution for the simultaneous view on the original sized, virtual depiction of the real-world environment and a user-controlled and rigidly transformable duplicate of a region of interest (ROI). As seen in Figure 1, the ROI is visualized as a transparent cube with opaque edges. A duplicate of the same region is created in front of the user, which can be moved, rotated, and scaled. We call this duplicated cut-out, which the user can interact with, a *Magnorama*, as a portmanteau of “magnified” and “diorama”. Magnoramas allow the users to focus on their actions but still be aware of their surroundings at different scales and points of view in the remote space. This is especially true in their interaction with other users in the same space.

For further addressing, we refer to the consultant working in the VR environment as the remote expert (RE) and the consultee in the AR environment as the local user (LU). They represent both sides of the teleconsultation system in our study and are subject to measuring their perceived co-presence, telepresence, and social presence of their partner.

3.1 Implementation

Although the implementation can be done using different techniques, we present our solution to this concept, implemented in Unity3D. We used an HTC Vive Pro Eye HMD, together with three HTC Vive trackers, a tracked HTC Vive Controller, and a Logitech VR Ink pen for the RE to realize inverse kinematic tracking [56]

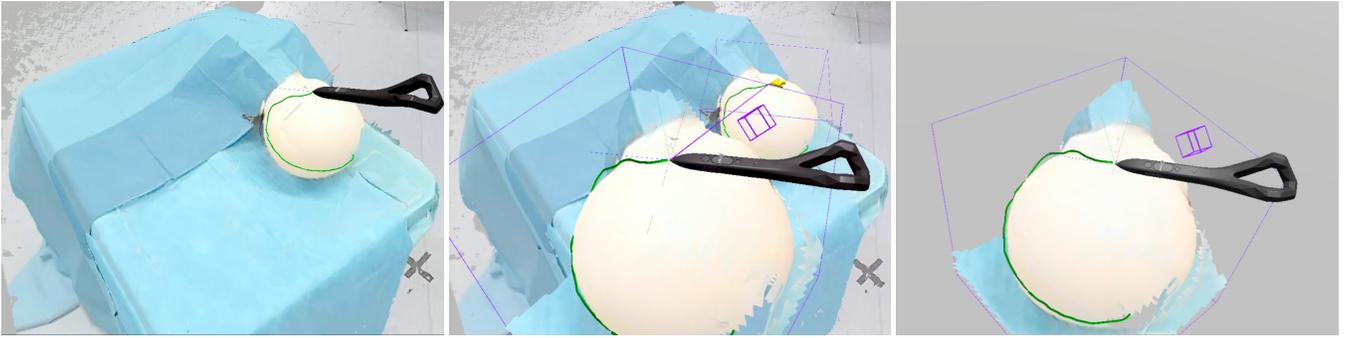


Figure 3: **Three experimental conditions.** From left to right: Baseline (left), Magnorama (center), Magnorama-Only (right). In Magnorama-Only, all context is masked, including the collaboration partner. The purple cube on the controller indicates that the Magnorama is active.

on the VR side. We used a Microsoft HoloLens 2 for the LU. The Magnorama implementation consists of these components:

- **ROI:** A transparent cube with the transformation wT_{ROI} which encapsulates the ROI inside the virtual environment.
- **Magnorama:** A placeholder object with the transformation $wT_{Magnorama}$ to provide the user with an interactable object.
- **Magnorama Camera:** A camera positioned with the same relation to the Magnorama as the rendering camera wT_{Cam} , but in relation to the ROI:

$$wT_{MagnoramaCam} = wT_{ROI} \cdot (wT_{Cam}^{-1} \cdot wT_{Magnorama})^{-1}$$

The rendering of the Magnorama camera has to accommodate our method. Objects and reconstructed point clouds need to be scaled around the center and clipped at the ROI's border, which can be done inside their shader used for rendering. The camera itself should be rendered using the same depth buffer as used by the camera of the HMD for the correct occlusion with the scene.

In our implementation, as seen in Figure 1, it appears that annotations drawn in the Magnorama are also directly and in real-time drawn at the original position inside the ROI. However, the opposite is the case. By drawing inside the Magnorama, the pose of the pen is transformed into the coordinate system of the ROI where the line is drawn. Since the Magnorama is a detached camera view of the ROI, the newly drawn line appears simultaneously in the Magnorama. This approach of implementing Magnoramas avoids duplicating objects in the scene since any interactions are directly performed at the original location.

3.2 Digital Representation of the Remote Expert

The RE can directly see the LU in the real-time captured point cloud (see Figure 2); however, the LU cannot see the RE without a virtual representation. For this reason, the RE is represented as a generic male or female avatar, to himself and to the LU, to allow for avatar embodiment [3, 4, 57] and (social)-presence [55]. The avatar's pose is transmitted to the LU and visualized as seen in Figure 4 and calculated in real-time through inverse kinematics. Parallel to the avatar representation, both participants were able to discuss the task using an external audio communication channel.

3.3 Appearance of the Magnorama for the Local User

As soon as the RE creates a Magnorama and proceeds to annotate the ROI, the user simultaneously detaches himself from the region at the on-site location. To communicate the use of the Magnorama for the LU, we added visual indicators. Two boxes depicting the selected ROI and the Magnorama are rendered for the LU while

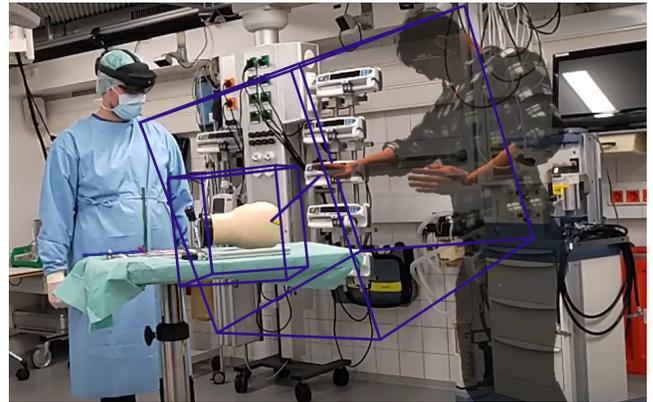


Figure 4: **Magnoramas as seen in shared Augmented Reality.** This view is captured at the local site from an additional HoloLens 2. The local user (left) observes the avatar of the remote expert (right) while drawing annotations using the Magnorama. The image was post-processed for better visibility of the Magnorama (purple).

the VR pen is inside the Magnorama. For the LU, the reconstruction inside of the boxes is not visualized because transmitting the content of the Magnorama as seen in VR would occupy an excessive amount of network bandwidth, memory capacity, and compute capabilities of the HoloLens 2. A link is rendered between both boxes that connect the location of the pen tip within the Magnorama and the corresponding back-transformed position inside the ROI. This link aids the LU to find the RE's avatar representation, even if it moves away from the scene during the annotation process. This link is also visible in Figure 4. We measure potential adverse effects from this solution by including the role of the LU.

3.4 Asymmetric Teleconsultation System

The proposed interaction methods were implemented in an asymmetric telepresence system inspired by Maimone et al. [24]. The system consists of three stationary Azure Kinect RGB-D cameras attached to dedicated camera PCs (MSI Trident 3, 16GB RAM, NVidia Geforce RTX 2060 GPU) and a high-performance rendering workstation (Intel Core i7, 64GB RAM, NVidia Geforce RTX 2080Ti). The computers communicate via a dedicated 1Gbps local area network. Each camera PC captures the color-image (BGRA, 1536p) and depth-image (NFOV Unbinned) with 30 FPS, encodes both image streams to H264 using the hardware encoders on the GPU (color: lossy compression/RGBA, depth: lossless compression/UINT16), and provides these streams with low latency as RTSP endpoints. Furthermore, the sensor calibration (intrinsic and

extrinsics) is supplied as Capnproto RPC endpoint from each camera PC. The image streams and calibration data are then received by the rendering workstation using a custom, native Unity3D plugin, decoded using the hardware decoders of the GPU and directly streamed to DirectX textures on the GPU to achieve low latency on the receiver side as well. First, each depth-image is unprojected into a structured point-cloud using the respective sensors' intrinsic parameters. Next, the individual point-clouds are converted to surface meshes [58] in a geometry shader by creating triangles from neighbored values of each depth-image and textured using the respective color images. Edges inside the depth image are handled by only allowing triangles to be generated if all three corners have at most a 2 centimeters difference in depth. The resulting meshes are positioned using their respective camera extrinsic parameters.

The extrinsics of the three RGB-D cameras for 3D reconstruction are estimated using a procedure similar to the room calibration of commercial optical tracking systems. In this process, we use the infrared images from the Azure Kinect sensors since they correlate directly with the generated depth-image for best precision. We use an L-shaped target with four reflective spheres placed on the floor to define the world origin and roughly estimate the camera poses. Next, we collect a video sequence using a calibration wand with two reflective spheres and use bundle-adjustment to refine the estimation of extrinsics. We register the Microsoft HoloLens 2 into the same world coordinate frame using a fiducial marker that is calibrated within the room using an external tracking system.

4 USER STUDY

An extensive user study was performed for the evaluation of our methods on Magnoramas. In the following, we describe the design of our user study and its related components.

4.1 Design

The experiment was designed as a one-factor (*Experimental Condition*) within-subjects experiment. Pairs of two participants performed a semi-collaborative task in an asymmetric VR/AR telepresence setting. The situation reflects a medical scenario with a LU requiring assistance for a surgical task and a RE assisting by annotating procedure steps. Participants experienced both the AR side as a LU and the VR side as a RE in three trials each, differing in the experimental condition. Our primary research goal was to confirm our hypothesized benefits of improved precision of the annotations and investigate potential downsides regarding presence and usability arising from the new methods and communicative inconsistencies that emerge from the two proposed novel interaction concepts. The object of interest for the study is a model of a head that is rigidly fixated in the room.

4.2 Experimental Conditions

We compare three conditions which we refer to as “baseline”, “Magnorama”, and “Magnorama-Only”. We theorize that each condition has advantages and disadvantages regarding the perceived presence and precision of the drawing task.

Baseline When the RE draws annotations using the baseline method for our comparison, it refers to the act of directly drawing on the visualized head in its original pose and size, as seen in Figure 3 (left). This represents the drawing methodology of similar telepresence systems with annotations with no option for magnification. In this condition, the user in VR can only see the 3D reconstruction but no magnification.

Magnorama The RE draws annotations inside the Magnorama but can still see the annotations on the real head. The RE is still able to see the body language of the LU in the point cloud. The RE can use the controller of their non-dominant hand to grab, rotate, and scale the Magnorama. In this condition, the user in VR can see

both the 3D reconstruction and the magnification. This method can be seen in Figure 3 (center).

Magnorama-Only Similar to the previous condition, the user draws the annotations inside the Magnorama. However, the user cannot see the original point cloud that is depicting the real-world, as seen in Figure 3 (right). Again, the user can use the controller of their non-dominant hand to grab, rotate, and scale the Magnorama. In this condition, the user in VR cannot see the 3D reconstruction but only the magnification.

4.3 Three Tasks Performed Per Condition

Our user study imposes a simplified scenario of a craniectomy. Craniectomy was identified as one of many potential use-cases for life-supporting remote telepresence systems in exchange with doctors and medical specialists. For this procedure, the surgeon must act both quickly and precisely in order to prevent life-critical damage. In medical terms, a craniectomy describes the removal of a part of the skull for releasing built-up pressure from a swelling of the brain after a traumatic brain injury. Three main tasks are necessary during the procedure: (1) Cut open the scalp of the injured person, (2) use a medical-grade drill to prepare holes in the skull (craniotomy), (3) use a medical-grade saw to disconnect the bone tissue between the holes.

For this study, we reduced the complexity of the tasks into abstracted color-coded tasks. The colors green, blue, and red each indicate one of the craniectomy tasks: a green line for outlining the cut on the scalp, blue pins for marking the drilling spots, and a red line for outlining the saw paths on the skull. The green line task covers a large area from the forehead to the ear. Users only require a single tap on the controller to place a pin during the blue pin placement task, which may provide insight into the precision of single-action tasks. In the red circle task, the guiding line covers a relatively small area, which is also passing through the positions of the pins. The guiding lines appear as blue lines, as seen in Figure 5, and not in the color assigned to the task to avoid confusion during the drawing procedure. All guiding elements are visible inside the Magnorama to the RE. Therefore medical expertise was not required for participation in the study as the participants were only required to redraw predefined guiding elements, as seen in Figure 5. The tasks will be referred to as 'line', 'pin', and 'circle' task further in this work.

4.4 Study Procedure

The user study was conducted in pairs. Each participant experienced both parts of the study paradigm (i.e., RE and LU). We welcomed participants separately and guided them to separate rooms. The study began with the visual tests and an initial demographics questionnaire, followed by the mental rotation questionnaire further described in subsection 4.7. The first participant on VR dons three Vive trackers for controlling their digital representation, which is visible for both RE and LU as described in subsection 3.2. The participants hold the VR pen for drawing annotations in their dominant hand, while they use their non-dominant hand for the controller to move the Magnorama. Each participant had the chance to become acquainted with the system for a maximum of 10 minutes, including creating annotations and interacting with the VR Ink pen and the Magnorama. No participant exhausted the full 10 minutes of familiarization to feel confident with the interactions. The order of the three experimental conditions (Baseline, Magnorama, Magnorama-Only) and the order of the color-coded tasks are randomized. The LU communicates the order of the tasks to the RE over an audio-communication channel. Additionally, the LU decides on a preference for the drawing direction of the annotation. This is done to encourage communication between both parties.

COVID-19 measures: Experimenters wore masks during the experiment and kept a safe distance from the participants. Partici-



Figure 5: **Guiding elements of the three tasks.** 1. Cut on the scalp (left), 2. drilling locations marked with cross-hairs (center), and 3. saw paths to disconnect bone tissue (right). All guiding elements are also visible in the Magnorama.

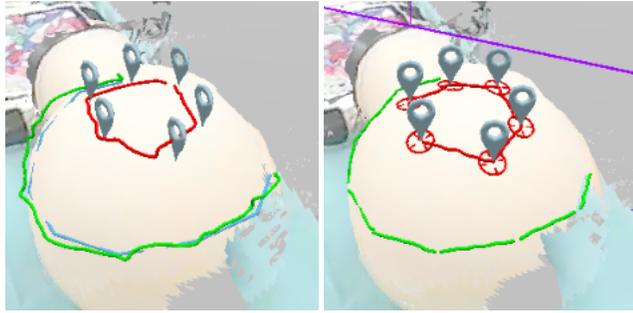


Figure 6: **Exemplary hand-drawn annotations** of the green line, light-blue pins, and red circle as seen in VR. Annotations are drawn by one of the participants based on the guiding elements for a baseline condition (left) and a Magnorama condition (right).

pants wore masks except for the time of the task. All equipment and contact surfaces were carefully disinfected after each trial block, and exchange devices were prepared for the participant switch. Rooms were sufficiently ventilated and participants were located in separate rooms. Strict exclusion criteria for the study were previous visits to risk areas and any symptoms or contact with infected persons. Participants were clarified of these conditions, and all participants consented. The study was conducted in accordance with the local COVID-19 regulations with necessary precautions and in accordance with the declaration of Helsinki.

4.5 Objective Performance Measures

The simulation logged error measurements of the drawings. In the green line and red circle task, the user redraws guiding lines. The error is calculated as the distance between the pen-tip and the closest line segment. In the blue pin task, the error is calculated using the distance between the pin and the closest target cross-hair. Inputs with an error greater than five centimeters are discarded during the evaluation. This excludes the annotations created by accident or for testing. Additionally, we recorded the time to task completion between the first and last valid user input for each task.

4.6 Subjective Measures

Participants are asked to complete a questionnaire consisting of five parts after completing each experimental condition. We assess co-presence, telepresence, and social presence based on the factors proposed by Nowak & Biocca [55]. The scales are adjusted to a 7-point Likert scale to ease the interpretation. We assessed the perceived usability by including the system usability scale (SUS) [59]. The SUS was evaluated using a 5-point scale (1 - strongly disagree, 5 - strongly agree). Further, we assessed the perceived task load using the NASA task load index (TLX) [60]. We evaluated the raw TLX total score (see [61]) and the sub-scores. A single question regarding the potential symptoms of cyber-sickness was added (Fast Motion Sickness Scale (FMS) [62]). After each condition, we

asked free-text answers for specific advantages and disadvantages of the method. At the end of the study, we asked participants for their method preference, the underlying reason, and comments.

4.7 Participants

In total, $N = 24$ participants ($M_{age} = 23.63$, $SD_{age} = 3.03$) were recruited via mailing lists and campus announcements. Of those, 23 were students of various fields, including medicine (3) and computer science (2). 8 participants were female, 16 male. Participants stated to spend time with digital media (PC, mobile phone, etc.) for about 34.21 hours per week ($SD = 3.85$). 19 participants noted to have used VR systems before, and 8 participants noted to have used AR systems before. The average amount of previous VR usage was $M = 4.46$ times, ranging between 0 and 30. The average amount of AR usage was $M = 2.17$ times, ranging between 0 and 30. 6 participant pairs have known each other before, 6 pairs did not know each other and were matched together on a first-come-first-serve basis.

To avoid any bias from visual impairments, we assessed a Landolt C-Test (EN ISO 8596) for acuity, an Ishihara Color test for color deficiency [63], and a Titmus test for stereo vision. All participants had normal or corrected-to-normal vision regarding acuity. One participant had slightly reduced stereo vision. Two participants had a slight red-green color weakness. Since there were no color mixtures involved in the experiment, we decided to include these in the analysis. We found that all participants were capable of performing the experiment. The average inter-pupillary distance of the sample was $M = 62.66$ mm, measured by the HoloLens 2 device. The mental rotation test [64] confirmed that none of the participants had severe mental rotation deficits.

5 RESULTS

5.1 Objective Performance Measures

The annotation performance was analyzed by calculating the minimum, maximum, and mean error of the deviation from the performed annotations from target shapes/pin positions and their standard deviations. We analyzed the annotation performance by the participants using a one-way repeated measures analysis of variance (ANOVA), with the method of annotation as the factor. Greenhouse-Geisser corrected values are reported in the case the assumption of sphericity was violated. Bonferroni corrected pairwise comparisons are reported for significant main effects. Descriptive results are depicted in Figure 7.

Pin Task Performance The results showed a significant main effect for the mean error of the pin placement measure; $F(1.44, 33.08) = 3.89$, $p = .043$, $\eta_p^2 = .145$. Pairwise comparisons revealed a significant difference between the baseline method ($M = 5.22$ mm, $SD = 5.24$ mm) and the Magnorama method, which resulted in a statistically significant smaller error ($M = 2.56$ mm, $SD = 2.32$ mm; $p < .05$). The Magnorama-Only method ($M = 3.02$ mm, $SD = 2.73$ mm) outperformed the baseline, but not to a significant level.

The analysis revealed a significant main effect for the minimal error of the pin placement measure; $F(2, 46) = 6.57$, $p = .003$, $\eta_p^2 = .222$. Bonferroni corrected pairwise comparisons revealed that the minimal error was significantly lower in the Magnorama-Only condition ($M = 0.98$ mm, $SD = 0.75$ mm), compared to the baseline condition ($M = 1.18$ mm, $SD = 1.38$ mm; $p = .012$). The Magnorama condition ($M = 1.18$ mm, $SD = 1.38$ mm) showed a lower error than the baseline condition, but not to a significant level.

Circle Task Performance Greenhouse Geisser corrected results for the main effect of the mean error of the circle task were statistically significant $F(1.55, 28.28) = 3.93$, $p = .038$, $\eta_p^2 = .146$. Pairwise comparisons showed that both the Magnorama condition ($M = 4.17$ mm, $SD = 4.44$ mm) as well as the Magnorama-Only

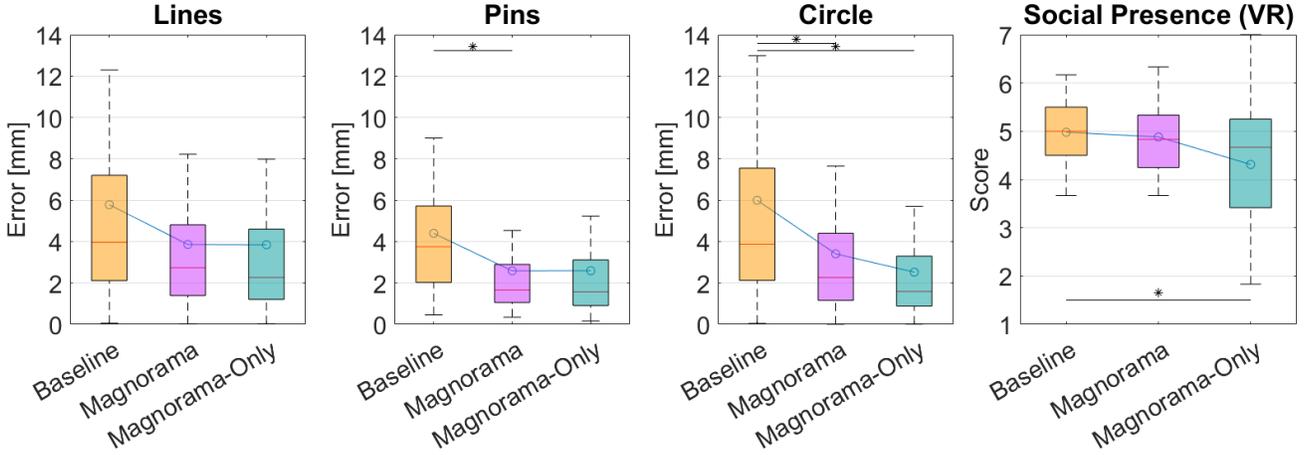


Figure 7: **Subjective and objective results of the study.** From left to right: (1-3) Box plots for the annotation errors from all participants. The red line indicates the median. The lower limit and upper limit of the box represent the 25th and 75th percentile. (4) Social presence as perceived by the participants when being the RE. Significance between conditions ($p < 0.05$) are marked with *.

condition ($M = 3.01 \text{ mm}$, $SD = 2.33 \text{ mm}$) significantly outperformed the baseline ($M = 6.43 \text{ mm}$, $SD = 5.72 \text{ mm}$; all $p \leq .003$). In this task, the Magnorama-Only condition performed significantly better than the Magnorama condition ($p = .001$).

Line Task Performance Greenhouse Geisser corrected values for the main effect on the mean error of the line error measurement showed no significant difference $F(1.14, 28.28) = 3.49$, $p = .068$, $\eta_p^2 = .132$. The baseline resulted in the highest mean error ($M = 6.79 \text{ mm}$, $SD = 6.25 \text{ mm}$), following the Magnorama condition ($M = 4.55 \text{ mm}$, $SD = 3.42 \text{ mm}$). The Magnorama-Only condition showed the lowest mean error ($M = 4.09 \text{ mm}$, $SD = 2.60 \text{ mm}$). No further significant effects were observed.

In summary, both Magnorama methods outperformed the baseline in all assessments, partly to a significant level. Regarding the mean error for drawing related tasks, the Magnorama-Only condition seems to outperform the Magnorama condition. However, the pin placements were more successful in the Magnorama condition.

Timing Results We recorded the time in which the participants performed each annotation task. We found a significant main effect for the line task; $F(2, 46) = 10.66$, $p < .001$, $\eta_p^2 = .317$. Pairwise comparisons revealed that the baseline method ($M = 18.43$, $SD = 9.14$) outperformed the Magnorama method ($M = 30.59$, $SD = 12.05$) as well as the Magnorama-Only method ($M = 31.35$, $SD = 19.61$, all $p \leq .003$). There was no significant difference between Magnorama and Magnorama-Only.

This main effect was similarly present for the circle task with a slightly smaller effect size; $F(2, 46) = 3.78$, $p = .030$, $\eta_p^2 = .141$. Pairwise comparisons showed that the timing for the baseline ($M = 13.49$, $SD = 8.76$) was lower than for the Magnorama method ($M = 18.70$, $SD = 8.70$) as well as lower than the Magnorama-Only method ($M = 20.51$, $SD = 12.93$), but not to a statistically significant level.

Interestingly in the pin placement task, this effect was not present; $F(1.42, 32.66) = 1.30$, $p = .282$, $\eta_p^2 = .054$. Baseline ($M = 11.56$, $SD = 5.88$), Magnorama ($M = 15.06$, $SD = 13.82$), and Magnorama-Only ($M = 13.03$, $SD = 7.09$) were almost at the same level.

5.2 Subjective Results

We performed Friedman tests with consecutive Bonferroni adjusted pairwise comparisons for the subjective measures. We were mainly

Table 1: SUS Score (0-100) and Raw total TLX (0-100), ($M \pm SD$)

	Baseline	Magnorama	Magnorama-Only
SUS RE	74.17 ± 12.74	73.96 ± 11.86	74.17 ± 9.37
SUS LU	71.04 ± 13.91	71.35 ± 13.70	70.62 ± 13.40
TLX RE	26.00 ± 16.25	22.19 ± 16.03	21.90 ± 14.50
TLX LU	16.00 ± 11.22	16.60 ± 10.58	16.56 ± 9.93

interested in the VR side (executing the annotation actions through the different methods) of the telepresence system.

For the VR side (RE), we found that the three conditions significantly impacted the level of social presence perceived by the participants; $\chi^2(2) = 6.66$, $p = .036$. Bonferroni corrected pairwise comparisons revealed that the baseline condition ($MDN = 5.00$) showed a significantly higher social presence than the Magnorama-Only condition ($MDN = 4.67$, $p = 0.30$). Differences between the baseline and the Magnorama condition ($MDN = 4.83$) or between the two Magnorama conditions were not significant. No significant differences were observed for the co-presence or telepresence measures. Further, no significant impacts on the presence factors on the AR side resulting from the different methods were found.

A Friedman's test for the SUS showed no significant difference in the usability assessment. All techniques were rated above a score of 70 for both the AR and the VR assessments (see Table 1). Friedman tests for the raw NASA TLX score (see [61] for a discussion), the NASA TLX subscales, and the FMS measure did not show significant differences between the conditions.

5.3 Preference and Comments

In VR, 12 participants preferred the Magnorama condition, 7 are undecided or did not answer, 4 preferred the baseline, and 1 preferred the Magnorama-Only condition. The participants liked the Magnorama condition because they can annotate on the magnified head while still being able to see the LU. Participants preferring the baseline condition perceived it as more natural compared to the other conditions. 18 participants in AR were unsure to pinpoint differences between the three conditions. 6 participants liked the baseline because the avatar directly worked on the head.

6 DISCUSSION

Our study compared the proposed Magnorama technique to a baseline and a Magnorama variant that masks the situative context. Our

results support *H1: The magnification of details in the form of Magnoramas increases free-hand drawing precision*, in the sense that the Magnorama conditions clearly outperformed the baseline conditions in many evaluated aspects to a significant level. The results show an improvement of the drawing precision using the magnified view of Magnoramas. This does not contradict one finding of Arora et al. [27], that drawing errors of larger objects in VR are higher, as the user-inputs are scaled back while using Magnoramas.

Magnoramas increased the time required for the line-tracing annotation tasks. We did not specifically draw a hypothesis on this aspect. Still, we suspect the reason for the increased time to be the increased length of the guiding lines inside the magnified region while users draw at similar speed in all conditions. This interpretation is backed by the fact that the time for placing pins did not show any significant difference between conditions. The green line task did not show significant improvement in precision, although having overall lower mean errors. This may be caused by the large region spanning from the forehead to the back of the ear and forces the user to change the point of view multiple times, whereas both, pin and circle tasks covered smaller regions.

The Magnorama-Only condition tends to yield lower error values. We assume by anecdotal observation that by hiding the environment, users are more likely to choose a larger scale of the Magnorama as space occupied by the original point cloud can be used to place the Magnorama. The precision can be further increased by choosing a smaller ROI and a larger Magnorama scale. This may be another starting point for further investigation.

Based on the TLX scores, our findings did not support *H2: Interacting with Magnoramas is inferior in terms of usability compared to context coherent interactions*. This was surprising since we would not have expected the Magnorama condition to be perceived similarly usable. One argument for the result may be that the users also perceived increased performance and, therefore, higher usability. For tasks where continuous lines need to be drawn (such as the line and circle task), users should consider a trade-off between an increased time-on-task and the magnification value.

The evaluation of the questionnaires showed further that among the three types of perceived presence, social presence perception was impacted by the conditions and found to be significantly lower for the RE in the Magnorama-Only condition compared to the baseline. This only partially supports *H3: The perception of the interaction in terms of co-presence and social presence aspects is inferior when using Magnoramas*. In addition, the Magnorama condition was able to maintain its perceived social presence while increasing the precision during the annotation tasks. We interpret this as the cause of the partial remaining coherence.

During the study, we observed that both participants exceedingly focus on the head during the drawing task and rarely look up at the participant. Therefore, we propose the use of Magnoramas when precise annotations or interactions are required. Before and after each task, automatic mechanisms could be incorporated to toggle the visualization of Magnoramas to regain a better perception of the communication partner. Magnoramas have a positive aspect on synchronicity, as the users can gesture, utilize non-verbal communication, etc., compared to the Magnorama-Only condition. On the other hand, in tasks requiring utmost concentration, such as the craniectomy, the Magnorama-Only setting can provide intentional concealment of the periphery as fewer distractions divert attention from the precision task and thus allowing the focus on the region of interest.

6.1 Limitations

There are some limitations. First, our study measured only the drawings from virtual ground truth to virtual space annotations. Therefore we cannot conclude the precision of annotations between virtual and physical relations. However, this was a conscious de-

sign choice since we did want to exclude additional noise from tracking and calibration errors from the experiment. For the same reason, only the precision for RE annotations was measured but not the precision of the drawings at the LU. Further research should investigate the error of LU annotations and the physical-to-virtual discrepancies. The number of left-handed participants was low (3 out of 24) for concluding its impact on the measurements. However, the randomized drawing direction dictated by the LU should mitigate the effect of handedness when reproduced on a larger sample. We did not explicitly measure the correlation between the degree of magnification and the time-on-task. Future studies using Magnoramas should monitor the drawing speed in the combination with the magnification. We are also aware that the quality of the real-time captured point cloud may introduce artifacts. Therefore, our findings with regard to the presence measures should be subject to further validation. Finally, for simplification and experimental control, we pre-defined the position of both the ROI and Magnorama. Users were neither required to choose the position and the initial sizes of the ROI by themselves. This may partially explain the usability results. In the desired target use-case, the ROI can be automatically selected through object detection based on the point cloud or opened manually by the user.

7 FUTURE WORK

Future work could integrate more compatible interactions for Magnoramas besides creating annotations, e.g., selecting and manipulating objects. We imagine Magnoramas hold potential as interactive second viewpoints, similar to Augmented Mirrors [65] to perform specific tasks, such as alignment or multi-modal visualization, more efficiently than are otherwise difficult. In the present work, we focused on the interaction of the VR user. In the future, we would like to compare new approaches in representing the avatar of the expert in AR since 25% of the participants in AR preferred it when the avatar directly annotated on the head. An exciting solution includes the attachment of the avatar at the real head in combination with the scaling of the avatar corresponding to the scale of the Magnorama. A similar approach has been investigated by Piumsomboon et al. [8] under the name of Mini-me. Consequently, rotating the Magnorama could have the avatar fly through the scene with a “jet-pack” inside the AR view, presented by Piumsomboon et al. [66]. In a scenario with more than two users, the perceived coherence and social presence may be impacted. Future work could therefore consider augmenting both, social behavior [67, 11] and appearance [8] of the avatars, to potentially compensate for missing coherence.

8 CONCLUSION

We proposed Magnoramas as a selective magnifier of a region of interest inside a 3D telepresence system using a real-time captured point-cloud. In our study we found that the magnification through a Magnorama allows a user to draw annotations more precisely in trade for a lower perceived social presence of the communication partner. This effect was mostly mitigated when using the Magnorama along-side the original point-cloud. The increased precision from Magnoramas can be incredibly impactful for any teleconsultation system which allows freehand interactions. Moreover, they can be generalized to manifold use-cases but could be specifically beneficial for medical or industrial scenarios. We conclude that the value of Magnoramas is substantial for our scenario of a craniectomy and successful in increasing the precision and quality of the annotations, which opens a path for future endeavors.

ACKNOWLEDGEMENTS

The authors wish to thank Andreas Keller and Gleb Gorbachev for their support during the software development process. This work was funded by the German Federal Ministry of Education and Research (BMBF), No.: 16SV8092.

REFERENCES

- [1] S. Beck, A. Kunert, A. Kulik, and B. Froehlich, "Immersive Group-to-Group Telepresence," *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 4, pp. 616–625, 2013.
- [2] N. Weibel, D. Gasques, J. Johnson, T. Sharkey, Z. R. Xu, X. Zhang, E. Zavala, M. Yip, and K. Davis, "ARTEMIS: Mixed-Reality Environment for Immersive Surgical Telementoring," in *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–4, 2020.
- [3] M. Slater, B. Spanlang, M. V. Sanchez-Vives, and O. Blanke, "First Person Experience of Body Transfer in Virtual Reality," *PLoS one*, vol. 5, no. 5, p. e10564, 2010.
- [4] K. Kilteni, R. Groten, and M. Slater, "The Sense of Embodiment in Virtual Reality," *Presence: Teleoperators and Virtual Environments*, vol. 21, no. 4, pp. 373–387, 2012.
- [5] J. Hollan and S. Stornetta, "Beyond Being There," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 119–125, 1992.
- [6] D. Roth, C. Kleinbeck, T. Feigl, C. Mutschler, and M. E. Latoschik, "Beyond Replication: Augmenting Social Behaviors in Multi-user Virtual Realities," in *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 215–222, IEEE, 2018.
- [7] R. Fribourg, N. Ogawa, L. Hoyet, F. Argelaguet, T. Narumi, M. Hirose, and A. Lécuyer, "Virtual Co-Embodiment: Evaluation of the Sense of Agency while Sharing the Control of a Virtual Body among Two Individuals," *IEEE Transactions on Visualization and Computer Graphics*, 2020.
- [8] T. Piumsomboon, G. A. Lee, J. D. Hart, B. Ens, R. W. Lindeman, B. H. Thomas, and M. Billinghurst, "Mini-me: An Adaptive Avatar for Mixed Reality Remote Collaboration," in *Proceedings of the 2018 CHI conference on human factors in computing systems*, pp. 1–13, 2018.
- [9] M. Slater and A. Steed, "Meeting People Virtually: Experiments in Shared Virtual Environments," in *The social life of avatars*, pp. 146–171, Springer, 2002.
- [10] I. Podkosova, K. Vasylevska, C. Schoenauer, E. Vonach, P. Fikar, E. Bronederk, and H. Kaufmann, "ImmersiveDeck: A large-scale Wireless VR System for Multiple Users," in *2016 IEEE 9th Workshop on Software Engineering and Architectures for Realtime Interactive Systems (SEARIS)*, pp. 1–7, IEEE, 2016.
- [11] D. Roth, G. Bente, P. Kullmann, D. Mal, C. F. Purps, K. Vogeley, and M. E. Latoschik, "Technologies for social augmentations in user-embodied virtual reality," in *25th ACM Symposium on Virtual Reality Software and Technology, VRST '19*, (New York, NY, USA), Association for Computing Machinery, 2019.
- [12] M. E. Latoschik, F. Kern, J.-P. Stauffert, A. Bartl, M. Botsch, and J.-L. Lugin, "Not Alone Here?! Scalability and User Experience of Embodied Ambient Crowds in Distributed Social Virtual Reality," *IEEE transactions on visualization and computer graphics*, vol. 25, no. 5, pp. 2134–2144, 2019.
- [13] S. Orts-Escolano, C. Rhemann, S. Fanello, W. Chang, A. Kowdle, Y. Degtyarev, D. Kim, P. L. Davidson, S. Khamis, M. Dou, et al., "Holoportation: Virtual 3D Teleportation in Real-time," in *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, pp. 741–754, 2016.
- [14] K. Kiyokawa, H. Takemura, and N. Yokoya, "A Collaboration Support Technique by Integrating a Shared Virtual Reality and a Shared Augmented Reality," in *IEEE SMC'99 Conference Proceedings. 1999 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No. 99CH37028)*, vol. 6, pp. 48–53, IEEE, 1999.
- [15] M. Billinghurst, H. Kato, and I. Poupyrev, "The MagicBook: A Transitional AR Interface," *Computers & Graphics*, vol. 25, no. 5, pp. 745–753, 2001.
- [16] O. Oda, C. Elvezio, M. Sukan, S. Feiner, and B. Tversky, "Virtual Replicas for Remote Assistance in Virtual and Augmented Reality," in *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, pp. 405–415, 2015.
- [17] A. Steed, W. Steptoe, W. Oyekoya, F. Pece, T. Weyrich, J. Kautz, D. Friedman, A. Peer, M. Solazzi, F. Tecchia, et al., "Beaming: An Asymmetric Telepresence System," *IEEE computer graphics and applications*, vol. 32, no. 6, pp. 10–17, 2012.
- [18] M. Le Chénéchal, T. Duval, V. Gouranton, J. Royan, and B. Arnaldi, "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)*, pp. 9–12, IEEE, 2016.
- [19] T. Piumsomboon, Y. Lee, G. Lee, and M. Billinghurst, "CoVAR: A Collaborative Virtual and Augmented Reality System for Remote Collaboration," in *SIGGRAPH Asia 2017 Emerging Technologies*, pp. 1–2, 2017.
- [20] R. A. J. de Belen, H. Nguyen, D. Filonik, D. Del Favero, and T. Bednarz, "A Systematic Review of the Current State of Collaborative Mixed Reality Technologies: 2013–2018," *AIMS Electronics and Electrical Engineering*, vol. 3, no. 2, p. 181, 2019.
- [21] P. Ladwig and C. Geiger, "A Literature Review on Collaboration in Mixed Reality," in *International Conference on Remote Engineering and Virtual Instrumentation*, pp. 591–600, Springer, 2018.
- [22] H. Fuchs and U. Neumann, "A Vision of Telepresence for Medical Consultation and other Applications," in *Sixth International Symposium of Robotics Research*, pp. 555–571, 1993.
- [23] H. Fuchs, G. Bishop, K. Arthur, L. McMillan, R. Bajcsy, S. Lee, H. Farid, and T. Kanade, "Virtual Space Teleconferencing Using a Sea of Cameras," in *Proc. First International Conference on Medical Robotics and Computer Assisted Surgery*, vol. 26, 1994.
- [24] A. Maimone and H. Fuchs, "A First Look at a Telepresence System with Room-sized Real-time 3D Capture and Life-sized Tracked Display Wall," *Proceedings of ICAT 2011, to appear*, pp. 4–9, 2011.
- [25] T. Piumsomboon, A. Dey, B. Ens, G. Lee, and M. Billinghurst, "The Effects of Sharing Awareness Cues in Collaborative Mixed Reality," *Frontiers in Robotics and AI*, vol. 6, p. 5, 2019.
- [26] I. Podkosova and H. Kaufmann, "Mutual Collision Avoidance During Walking in Real and Collaborative Virtual Environments," in *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*, pp. 1–9, 2018.
- [27] R. Arora, R. H. Kazi, F. Anderson, T. Grossman, K. Singh, and G. W. Fitzmaurice, "Experimental Evaluation of Sketching on Surfaces in VR.," in *CHI*, vol. 17, pp. 5643–5654, 2017.
- [28] E. Wiese, J. H. Israel, A. Meyer, and S. Bongartz, "Investigating the Learnability of Immersive Free-hand Sketching," in *Proceedings of the seventh sketch-based interfaces and modeling symposium*, pp. 135–142, 2010.
- [29] C. Moore, P. J. Dunham, and P. Dunham, *Joint Attention: Its Origins and Role in Development*. Psychology Press, 2014.
- [30] R. Stoakley, M. J. Conway, and R. Pausch, "Virtual Reality on a WIM: Interactive Worlds in Miniature," in *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '95*, (Denver, Colorado, United States), pp. 265–272, ACM Press, 1995.
- [31] A. Bluff and A. Johnston, "Don't Panic: Recursive Interactions in a Miniature Metaworld," in *The 17th International Conference on Virtual-Reality Continuum and its Applications in Industry, VRCAI '19*, (New York, NY, USA), pp. 1–9, Association for Computing Machinery, Nov. 2019.
- [32] R. Trueba, C. Andujar, and F. Argelaguet, "Complexity and Occlusion Management for the World-in-Miniature Metaphor," in *International Symposium on Smart Graphics*, pp. 155–166, Springer, 2009.
- [33] R. Trueba, C. Andujar, and F. Argelaguet, "Multi-scale Manipulation in Indoor Scenes with the World in Miniature Metaphor," in *Proceedings of the 15th Joint virtual reality Eurographics conference on Virtual Environments*, pp. 93–100, 2009.
- [34] A. Mulloni, H. Seichter, and D. Schmalstieg, "Indoor Navigation with Mixed Reality World-in-Miniature Views and Sparse Localization on Mobile Devices," in *Proceedings of the International Working Conference on Advanced Visual Interfaces*, pp. 212–215, 2012.
- [35] F. Bacim, D. Bowman, and M. Pinho, "Wayfinding Techniques for Multiscale Virtual Environments," in *2009 IEEE Symposium on 3D User Interfaces*, pp. 67–74, IEEE, 2009.
- [36] D. Coffey, N. Malbraaten, T. Le, I. Borazjani, F. Sotiropoulos, and D. F. Keefe, "Slice WIM: a Multi-surface, Multi-touch Interface for Overview+ Detail Exploration of Volume Datasets in Virtual Reality,"

- in *Symposium on Interactive 3D Graphics and Games*, pp. 191–198, 2011.
- [37] C. A. Wingrave, Y. Haciahetoglu, and D. A. Bowman, “Overcoming World in Miniature Limitations by a Scaled and Scrolling WIM,” in *3D User Interfaces (3DUI’06)*, pp. 11–16, IEEE, 2006.
- [38] J. S. Pierce, B. C. Stearns, and R. Pausch, “Voodoo Dolls: Seamless Interaction at Multiple Scales in Virtual Environments,” in *Proceedings of the 1999 symposium on Interactive 3D graphics*, pp. 141–145, 1999.
- [39] R. S. Renner, B. M. Velichkovsky, and J. R. Helmert, “The Perception of Egocentric Distances in Virtual Environments - A Review,” *ACM Computing Surveys (CSUR)*, vol. 46, no. 2, pp. 1–40, 2013.
- [40] H. Elsayed, M. D. Barrera Machuca, C. Schaarschmidt, K. Marky, F. Müller, J. Riemann, A. Matvienko, M. Schmitz, M. Weigel, and M. Mühlhäuser, “VRSketchPen: Unconstrained Haptic Assistance for Sketching in Virtual 3D Environments,” in *26th ACM Symposium on Virtual Reality Software and Technology*, pp. 1–11, 2020.
- [41] M. D. B. Machuca, W. Stuerzlinger, and P. Asente, “Smart3DGuides: Making Unconstrained Immersive 3D Drawing More Accurate,” in *25th ACM Symposium on Virtual Reality Software and Technology*, pp. 1–13, 2019.
- [42] M. D. Barrera-Machuca, A. Cassinelli, and C. Sandor, “Context-Based 3D Grids for Augmented Reality User Interfaces,” in *Adjunct Publication of the 33rd Annual ACM Symposium on User Interface Software and Technology*, pp. 73–76, 2020.
- [43] M. D. B. Machuca, P. Asente, W. Stuerzlinger, J. Lu, and B. Kim, “Multiplanes: Assisted Freehand VR Sketching,” in *Proceedings of the Symposium on Spatial User Interaction*, pp. 36–47, 2018.
- [44] T. Drey, J. Gugenheimer, J. Karlbauer, M. Milo, and E. Rukzio, “VRSketchIn: Exploring the Design Space of Pen and Tablet Interaction for 3D Sketching in Virtual Reality,” in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–14, 2020.
- [45] M. D. Barrera Machuca, W. Stuerzlinger, and P. Asente, “The Effect of Spatial Ability on Immersive 3D Drawing,” in *Proceedings of the 2019 on Creativity and Cognition*, pp. 173–186, 2019.
- [46] M. Garau, M. Slater, V. Vinayagamoorthy, A. Brogni, A. Steed, and M. A. Sasse, “The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment,” in *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 529–536, 2003.
- [47] D. Roth, J.-L. Lugin, D. Galakhov, A. Hofmann, G. Bente, M. E. Latoschik, and A. Fuhrmann, “Avatar Realism and Social Interaction Quality in Virtual Reality,” in *2016 IEEE Virtual Reality (VR)*, pp. 277–278, IEEE, 2016.
- [48] D. Roth, K. Waldow, M. E. Latoschik, A. Fuhrmann, and G. Bente, “Socially immersive avatar-based communication,” in *2017 IEEE Virtual Reality (VR)*, pp. 259–260, 2017.
- [49] P. Stotko, S. Krumpfen, M. B. Hullin, M. Weinmann, and R. Klein, “Slamcast: Large-scale, Real-time 3D Reconstruction and Streaming for Immersive Multi-client Live Telepresence,” *IEEE transactions on visualization and computer graphics*, vol. 25, no. 5, pp. 2102–2112, 2019.
- [50] T. Pejisa, J. Kantor, H. Benko, E. Ofek, and A. Wilson, “Room2room: Enabling Life-size Telepresence in a Projected Augmented Reality Environment,” in *Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing*, pp. 1716–1725, 2016.
- [51] B. Thoravi Kumaravel, F. Anderson, G. Fitzmaurice, B. Hartmann, and T. Grossman, “Loki: Facilitating Remote Instruction of Physical Tasks using bi-directional Mixed-Reality Telepresence,” in *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, pp. 161–174, 2019.
- [52] O. Oda and S. Feiner, “3D Referencing Techniques for Physical Objects in Shared Augmented Reality,” in *2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 207–215, IEEE Computer Society, 2012.
- [53] J. Kolkmeier, E. Harmsen, S. Giesselink, D. Reidsma, M. Theune, and D. Heylen, “With a Little Help from a Holographic Friend: The OpenIMPRESS Mixed Reality Telepresence Toolkit for Remote Collaboration Systems,” in *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, pp. 1–11, 2018.
- [54] R. L. Daft and R. H. Lengel, “Organizational Information Requirements, Media Richness and Structural Design,” *Management science*, vol. 32, no. 5, pp. 554–571, 1986.
- [55] K. L. Nowak and F. Biocca, “The Effect of the Agency and Anthropomorphism on Users’ Sense of Telepresence, Copresence, and Social Presence in Virtual Environments,” *Presence: Teleoperators & Virtual Environments*, vol. 12, no. 5, pp. 481–494, 2003.
- [56] D. Roth, J.-L. Lugin, J. Büser, G. Bente, A. Fuhrmann, and M. E. Latoschik, “A Simplified Inverse Kinematic Approach for Embodied VR Applications,” in *2016 IEEE Virtual Reality (VR)*, pp. 275–276, IEEE, 2016.
- [57] D. Roth and M. E. Latoschik, “Construction of the virtual embodiment questionnaire (veq),” *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 12, pp. 3546–3556, 2020.
- [58] G. Turk and M. Levoy, “Zippered Polygon Meshes from Range Images,” in *Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, pp. 311–318, 1994.
- [59] J. Brooke, “SUS: a ‘Quick and Dirty’ Usability,” *Usability evaluation in industry*, p. 189, 1996.
- [60] S. G. Hart and L. E. Staveland, “Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research,” in *Advances in psychology*, vol. 52, pp. 139–183, Elsevier, 1988.
- [61] S. G. Hart, “NASA-Task Load Index (NASA-TLX); 20 years later,” in *Proceedings of the human factors and ergonomics society annual meeting*, vol. 50, pp. 904–908, Sage publications Sage CA: Los Angeles, CA, 2006.
- [62] B. Keshavarz and H. Hecht, “Validating an Efficient Method to Quantify Motion Sickness,” *Human factors*, vol. 53, no. 4, pp. 415–426, 2011.
- [63] S. Ishihara, “Series of Plates Designed as Tests for Colour-blindness,” 1936.
- [64] S. G. Vandenberg and A. R. Kuse, “Mental Rotations, a Group Test of Three-dimensional Spatial Visualization,” *Perceptual and motor skills*, vol. 47, no. 2, pp. 599–604, 1978.
- [65] A. Martin-Gomez, A. Winkler, K. Yu, D. Roth, U. Eck, and N. Navab, “Augmented Mirrors,” in *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 217–226, IEEE, 2020.
- [66] T. Piumsomboon, G. A. Lee, B. Ens, B. H. Thomas, and M. Billinghurst, “Superman vs Giant: A Study on Spatial Perception for a Multi-scale Mixed Reality Flying Telepresence Interface,” *IEEE transactions on visualization and computer graphics*, vol. 24, no. 11, pp. 2974–2982, 2018.
- [67] D. Roth, M. E. Latoschik, K. Vogeley, and G. Bente, “Hybrid avatar-agent technology – a conceptual step towards mediated ‘social’ virtual reality and its respective challenges,” *i-com*, vol. 14, no. 2, pp. 107–114, 01 Aug. 2015.