



**HAL**  
open science

# What Can I Do There? Controlling AR Self-Avatars to Better Perceive Affordances of the Real World

Adélaïde Genay, Anatole Lécuyer, Martin Hachet

► **To cite this version:**

Adélaïde Genay, Anatole Lécuyer, Martin Hachet. What Can I Do There? Controlling AR Self-Avatars to Better Perceive Affordances of the Real World. ISMAR 2022 - 21st IEEE International Symposium on Mixed and Augmented Reality, Oct 2022, Singapore, Singapore. pp.450-459, 10.1109/ISMAR55827.2022.00061 . hal-03820756

**HAL Id: hal-03820756**

**<https://hal.science/hal-03820756>**

Submitted on 19 Oct 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# What Can I Do There? Controlling AR Self-Avatars to Better Perceive Affordances of the Real World

Adélaïde Genay\*  
Inria, Univ. Bordeaux, LaBRI, CNRS

Anatole Lécuyer†  
Inria, Univ. Rennes, IRISA, CNRS

Martin Hachet‡  
Inria, Univ. Bordeaux, LaBRI, CNRS



Figure 1: Control of self-avatars visualized through an Augmented Reality headset to better perceive interactions and affordances in the physical surroundings. *Left* Testing fire exit paths with a gamepad. *Center* Planning and testing a route before climbing by controlling the avatar’s limbs with gestures. *Right* Evaluating possible actions on a distant step stool with body-tracking mapping.

## ABSTRACT

This work explores a new usage of Augmented Reality (AR) to extend perception and interaction within physical areas ahead of ourselves. To do so, we propose to detach ourselves from our physical position by creating a controllable “digital copy” of our body that can be used to navigate in local space from a third-person perspective. With such a viewpoint, we aim to improve our mental representation of distant space and understanding of action possibilities (called affordances), without requiring us to physically enter this space. Our approach relies on AR to virtually integrate the user’s body in remote areas in the form of an avatar. We discuss concrete application scenarios and propose several techniques to manipulate avatars in the third person as a part of a larger conceptual framework. Finally, through a user study employing one of the proposed techniques (puppeteering), we evaluate the validity of using third-person embodiment to extend our perception of the real world to areas outside of our proximal zone. We found that this approach succeeded in enhancing the user’s accuracy and confidence when estimating their action capabilities at distant locations.

**Index Terms:** Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Mixed/augmented reality;

## 1 INTRODUCTION

Humans perceive the physical world through action [9]. By moving their bodies, they provide their sensory organs with continuous access to new data that, combined with experience, allows carrying out decisions successfully. However, acting may be impossible when the environment is inaccessible, distant, or dangerous and experience may not be sufficient to fill the missing information. In such situations, one could wish to have the ability to be free from one’s bodily envelope and explore the world from a distance.

\*e-mail: adelaid.genay@inria.fr

†e-mail: anatole.lecuyer@inria.fr

‡e-mail: martin.hachet@inria.fr

Recent progress in Mixed Reality (MR) technologies offers to do so. Whereas our physical body is intrinsically limited by its material characteristics, Merleau-Ponty argues that our perception and experience of the world cannot be reduced to material properties and may therefore be extended [22]. By modifying the inputs of our perception, research has shown that MR has the potential to enable such an extension. For example, it was previously used to extend the reach of one’s arms by virtually modifying their length [7], or to duplicate one’s body and interact with it [16].

“Extending” our body’s physical limits through MR has many promising use cases. In particular, users could employ MR to send a virtual version of their body (called *self-avatar*) nearby a distant object to get a better idea of its size or to simulate actions and observe them performed in relation to a physical space. We think such ability also opens the door to new types of explorations aimed at better understanding the relationship between our body, motor actions, environment, and thus cognition. However, MR research on how virtually pushing the limits of our body can enhance our perception of the real world is still preliminary. Additionally, the control of an avatar in the third person within a physical environment is neither innate nor easy to implement. Using this ability in studies or concrete applications first requires setting up the appropriate technology and providing suitable means to act through self-avatars in the real world. In this paper, we focus on this challenge.

More specifically, the contribution of this paper is threefold. First, we discuss the concept of using a self-avatar in the third person to improve the real world’s perception by leveraging existing cognitive mechanisms. Second, we describe a concrete implementation of this concept through an AR system allowing one to manipulate an avatar from a remote place through three interaction techniques: Physical Control, Puppeteering, and Body Tracking. Third, we present two user experiments employing one of these interaction techniques (Puppeteering) to provide feedback on the validity of our approach. Although the embodiment in the third-person perspective (3PP) was investigated before, we believe it had never been explored to enhance the perception of real-world surroundings with such a method. The results build towards a new way of using MR displays to better perceive what is already present in the physical world through a “virtual twin”, instead of augmenting the world with virtual objects while staying constrained to our bodily envelope.

## Use cases

To exemplify the possible advantages of using a self-avatar in assisting the perception of real-world situations, we present concrete scenarios. In each case, users may improve their understanding of the physical environment, test, and refine their strategies at a distance before acting for real.

**Climbing.** The sport of climbing requires anticipating a route from the ground by imagining one's body in a place where it is not [32]. Identifying which holds can be grasped from a distant position can be difficult for climbers that lack experience. These climbers could use an AR system to send their self-avatar onto the wall and plan their ascent from a vantage point on the ground (see Fig. 1, center). While controlling their avatar, they may anticipate which holds can be reached next by extending the virtual limbs of their avatar and trying different postures. Having the same body size as their users, self-avatars might also help to correct false affordances that occur when observing others successfully reaching holds that are too far for their own arms.

**Rescue.** Misperceptions of one's abilities also occur in situations that present risks or that engage certain mental states. For example, people may underestimate their ability to reach objects through small apertures when feeling anxious [2]. Providing the means to test one's capabilities virtually could help correct the effects of emotions in real-life situations, e.g. before entering a building threatening to collapse, or when training to face a fire. A firefighter trainee that is not confident about their ability to crouch under a beam could check whether their body would fit or not by sending their self-avatar in their place first. This might help them to combat misperceptions linked to their fear and gain more confidence in future real interventions.

**Observing a piece of art.** It is difficult to realize how tall an actor is on a theater scene while sitting far away in the hall. Similarly, estimating the size of a very large statue while standing at its very feet is also hard. In these situations, controlling a virtual double of oneself could provide a familiar and reliable scale to grasp dimensions more accurately. Users could bring their virtual avatar to a point of interest and observe the size of its body in relation to that of the object from different angles.

These simple scenarios can be generalized to many other real-life tasks that require imagining oneself in a distant place. We believe MR embodiment in the third person has the potential to relieve a part of the mental effort demanded by this process.

## 2 RELATED WORK

The concept presented in this work is based on the vision that avatar embodiment can enhance real-world perception. In this section, we first discuss the roots of this vision in cognitive science. Second, we describe how others used MR before us to improve the physical environment's perception. Lastly, we outline the research on the 3PP embodiment that inspired us when designing the presented system.

### Theoretical Foundations: Embodied Cognition

According to Gibson's ecological theory of perception, the perception of environments is directly linked to the actions that one is capable of performing within it [9]. The term *affordance* refers to the compatibility of environmental (as perceived by the senses) and individual characteristics (e.g. size of the body) [23, 34]. For instance, a tree branch set sufficiently high may afford walking under, but not sitting onto or stepping over [2].

People can usually determine if an environment allows them to perform an action without having to try it [2, 19, 37]. For example, Warren and Whang showed that participants estimated correctly that apertures needed to be at least 1.16 times their shoulder width to be able to pass through them without having to rotate one's shoulders [37]. Affordances can also be recalibrated to meet new skills or

situations [11, 35]. Ishak et al. [11] notably found that participants were able to adjust their decisions about whether or not their hands could fit through an aperture after having enlarged their hands.

However, Mark et al. [20] showed that such recalibration can only occur if participants are allowed to move their point of view: their capacity to adjust information and judge affordances was considerably diminished when visual input was limited to vision through a peephole or when mobility was restricted by having them rest their heads against a wall. The system designed in this paper is built on these observations and seeks to take advantage of our natural ability to understand things through action and locomotion.

### Improving the Perception of Real Environments with MR

One way to improve the perception of physical space is to allow the user to access new information by letting them adopt artificial viewpoints [4, 13]. Systems implementing such viewpoints use cameras to reconstruct the environment in 3D, and then immerse their users in the resulting virtual environment where any perspective can be displayed [16, 24, 29]. The experience of such systems is close to Virtual Reality — even in the case of Remixed Reality, a system developed by Lindbaueur et al. [16] where real-time photogrammetry of the physical world is displayed.

Rather than substituting the user's sight, some research proposed using situated visualization to help users imagine the effect of their actions on objects [12, 38]. For example, Leigh et al. [15] developed a mobile see-through AR system letting users see the consequences of their potential actions, predicted by a model. Other papers have looked into making already existing information easier to perceive by augmenting various sensory channels, including vision [3, 5, 40], audition [33], and touch [39]. We draw upon this set of examples to improve the user's perception of affordances. Unlike them, we propose to let users explore and sense their real environment at will through self-initiated action in see-through AR.

### Increasing Spatial Awareness through Third-person Avatars

Often used in games, the 3PP provides a wide field of view enabling one to quickly perceive elements around oneself. Previous work investigating this view in MR usually implemented it by moving the user's camera viewpoint outside of their body's location (no avatars) [8, 14]. In MR, Salamin et al. [31] showed that moving the user's viewpoint behind their bodies reduced the training required for a ball-catching task. Liu et al. [17] also showed that the 3PP resulted in slightly less precision during a measurement task in AR, but it allowed being three times faster. In VR, the effects of the 3PP are contrasting. While several papers [1, 10] showed positive effects on spatial awareness in various kinds of tasks, Medeiros et al. [21] found that these effects varied with the avatar's appearance. It is not clear at this point whether this also applies outside of virtual worlds.

A second approach to provide a 3PP is to display a duplicate of the user's body ahead (i.e. an avatar), observed from a first-person viewpoint. This visualization is similar to autoscopic experiences, as two bodies are visible. Although papers implementing such a perspective exist [25, 30], we are not aware of studies testing it to increase the spatial awareness of physical spaces. We propose to start exploring this approach in this paper.

In summary, what fundamentally differentiates our work is: (i) our users do not change their visualization perspective, (ii) they see their real environment rather than a remote/virtual one, and (iii) they have control over the exploration of their surroundings through a virtual avatar, matching their body dimensions. We explored how to design a system to assist real-world perception by using a fully rigged abstract model directly registered in the physical environment with see-through AR. To our knowledge, our work is the first to explore such new directions.

### 3 RATIONALE AND DESIGN CONSIDERATIONS

Our final goal is to improve the perception of environmental properties in distant spaces to make better-informed decisions and prepare for actions. We do not seek automatic methods that could analyze on the fly the physical environment and try to optimize ideal body movements. Instead, our approach is to leverage existing cognitive mechanisms by providing people with the ability to simulate their actions outside of their peripersonal space. To do so, we propose to rely on a virtual avatar that represents the user, that is embedded in the real world, and that can be easily manipulated.

The 3D registration of the avatar in the real world requires the use of an MR system. Head-Mounted Displays (HMDs) appear well suited for our objective as users can observe virtual objects while keeping their hands free for interaction. To safeguard natural perception of the real world, we opted for an optical see-through (OST) HMD. Compared to video see-through systems, OST HMDs provide an unmediated view of the real world and therefore ensure that visual and proprioception information is synchronized [8]. Current OST-HMDs also have scanning capabilities, which favor the consistent integration of the virtual avatar within the real environment.

Of course, for the user to perceive the real world as if they were actually experiencing it, the sizes of the avatar's limbs have to be similar to the user's body. Beyond limb sizes, reproducing the user's traits with fidelity and realism does not seem essential for this system. Realistic avatar appearances may additionally provoke Uncanny Valley effects that can negatively impact user experience [36]. Therefore, we decided to personalize the avatar's body, but cover it with an abstract and generic texture.

Lastly, as the actions of the users in the real world should be as varied as possible, we explored and identified three main potential interaction needs.

- **Travelling and wayfinding:** First, it may be interesting for users to stay in place and explore possible paths in the real world by moving their avatars from one location to another, as if they were walking themselves, e.g. to better perceive the dimensions of a room.
- **Posture editing:** Second, beyond global movement, individual limb manipulation may be valuable. An example is when trying to figure out which holds can be grasped before climbing onto a boulder. In this case, independent and fine control over each body limb is necessary.
- **One-to-one mapping:** Third, it may be interesting to project, through the avatar, a particular body gesture in the real world. For example, a dancer could wish to check if they have enough space to perform a particular figure on a stage with cluttered and fragile decor by actually performing the figure at a distance, in a safer zone.

To accommodate for the variety of tasks related to these different needs, we have explored three interaction categories that are described in Table 1. Depending on the environment and goal, one may choose the best-suited approach, or combine them for comprehensive exploration. The choice of the interaction method may also come from users' specific needs. For example, an elderly user may have difficulties with precise motor input but may be able to control the avatar with a controller instead.

### 4 IMPLEMENTATION

As a proof of concept, we implemented a prototype enabling the control of a self-avatar as described in previous sections. We implemented three modes to manipulate this avatar which is personalized to match the user's body. This section details how the different components of the overall system were implemented. The code is available at: <https://gitlab.inria.fr/agenay/ISMAR22-whatCanIDoThere>.

#### 4.1 AR Self-Avatar Visualization

As we meant to propose several control techniques to animate the avatar, we decided to opt for a rigged mesh model rather than a point cloud avatar which only affords body tracking.

**Display.** We used a Microsoft HoloLens 2 to display the avatar in OST AR. This HMD has an approximate field of view of 54 degrees diagonally and is equipped with 4 visible-light cameras, 2 infrared cameras, 1-MP time-of-flight depth sensor, and inertial measurement units allowing real-time surface detection, hand tracking, and positional tracking with six degrees of freedom.

**AR Module.** We exploit the HoloLens 2 sensors in a C# implementation to register the avatar in 3D space and to detect user gestures. Environment detection is also used to implement occlusion and collisions of the avatar with real surfaces. To do so, we use Unity3D 2019.4.16f1 and Mixed Reality Tool Kit (MRTK) v2.6.1. to build an application for HoloLens 2. This application is also in charge of processing the user inputs of all three modes and of managing the changes in mode.

**Avatar Generation.** Medeiros et al. [21] found that mesh models resulted in lower accuracy during navigation tasks in 3PP VR compared to point cloud avatars. It is unclear whether this also occurs when exploring physical environments too. However, to avoid potential discrepancies that might have caused such negative effects, we personalize the avatar to match the user's morphology, gender, and limb sizes with the free avatar creation tool Virtual Caliper [27] (based on the SMPL model [18]). We use all 6 of the proposed input parameters to generate user-matching avatars before testing: height, weight, arm span, inseam height, inseam width, and wrist-to-shoulder distance. The model generated by Virtual Caliper is rigged and skinned but does not include the user's real body texture (hair, clothes, etc.). Once imported in Unity, we used a generic abstract texture to cover the avatar (see Fig. 1).

#### 4.2 Control Modes

We implemented three control modes corresponding to the categories described in Table 1. Depending on the task, the best-suited mode can be chosen. One can also use a combination of the three modes by switching between the control modes through a virtual menu attached to one's hand (see Fig. 2.2).



Figure 2: Example usage of our system. 1) A user is trying to figure out how to reach a hold on a boulder. 2) To better visualize her possibilities, she puts on an AR headset and launches the Puppeteering mode. 3) She then sets the position of her avatar with hand interactions. 4) The avatar as seen by the user (photo shot from the headset).



Table 1: Overview of the interaction modes that we have explored to allow manipulation of an avatar from a distance, in AR.

Category	Scale	Implementation	Perks	Limits
<b>Travelling, wayfinding</b>	World	Keypad, game controller, joystick, tactile displays, etc.	Requires minimum physical effort. Allows making the avatar walk over distances without actually moving. Can be used eyes-off after little training.	Remembering the mapping between buttons becomes difficult after only 2-3 buttons are used. Control over the avatar is limited to a set of pre-recorded animations.
<b>Posture editing</b>	Limb	Pinching, dragging, pointing, puppeteering, raycasting, etc.	Provides the finest control of the avatar's posture. Metaphore-based interactions are easy to learn.	It's difficult to manipulate several limbs at the same time. Gesture recognition is not always reliable and can be physically tiring. Implementing postures can be slow.
<b>One-to-one mapping</b>	Body	Optical, inertial, mechanical, magnetic tracking, etc.	Most direct and natural control (one-to-one mapping). Provides vestibular cues (inertia and balance) and a strong sense of agency [8].	Multiple technological constraints, including sensor range, cost, and portability. Noise is introduced in movements due to tracking errors. Achieving certain postures can be impossible from a distance (e.g. climbing on a wall).

### Physical Control Mode (travelling, wayfinding)

For this mode, we used a wireless XBOX controller (X/S series). It was paired in Bluetooth to the Hololens 2 and its button mapping was managed by a Unity application. Since this mode is dedicated to providing navigation, buttons were mostly mapped with actions linked to locomotion through pre-recorded animations. The button layout we chose follows conventional controls of western platform games: left joystick for moving and turning, (A) button for jumping upwards, and left trigger button for crouching. We additionally use (B) for sitting, (Y) for extending arms in T-pose, and the down pad button for laying down. When not moving, the avatar was animated with an idle animation making it appear to breathe slowly.

### Puppeteering Mode (posture editing)

In this mode, the avatar's behavior is set to that of an idle active ragdoll whose limbs can be moved by dragging around transparent spheres attached to them (see Fig. 2.4). These spheres respond to input gestures detected by the Hololens 2 (pinching, dragging, and ray-casting). To implement this, we use MRTK and inverse kinematics scripts with an active ragdoll configuration. Colliders and joint limits of the avatar's bones are generated automatically with the help of the PuppetMaster v1.1 package [28]. We let users drag the avatar's position without affecting its posture by selecting its body. They may also rotate it or its individual limbs by making a twisting movement with their wrists. To facilitate placement, we froze the avatar's body rotation to only the vertical axis by default. To enable other rotation axes, users may press a "free rotation" button.

### Body Tracking Mode (one-to-one mapping)

The Body Tracking mode employs a Microsoft Azure Kinect to track the position and rotation of 32 body joints. Tracking data is extracted with Microsoft's Body Tracking SDK (v.1.0.1), and streamed to the Hololens 2 through a PC (NVIDIA GeForce RTX 2080, Intel Core i9-9900K, 32 GB of RAM). To avoid having to build a client-server network, we used the Holographic Remoting tool provided by MRTK. There was a total of 32 joints streamed by the tracker. For the tracking to function properly, the user needs to stay in the camera's sensor range. The user may face any direction while in this mode, but body parts that are not in the field of view of the Kinect cannot be tracked. For this reason, the camera should be placed in a manner that minimizes self-occlusion.

## 5 EXPLORATORY STUDY

We conducted two experiments exploring if a self-avatar could enhance the perception of affordances. We set up a controlled indoor environment where subjects had to explore physical spaces through a rich set of movements while using one avatar control technique. We chose to focus on the Puppeteering mode as it allowed testing various and complex postures. The other modes limited the poses we could test (see Table 1). We used the same system as presented

before but containing the Puppeteering mode only. Before the testing session, we collected the body measures of the volunteers to generate their avatars. No compensation or course credits were issued, and all participants were unaware of the purpose of either experiment.

### 5.1 Experiment 1

The first experiment aimed to validate that 3PP avatar manipulation could effectively help users when assessing their action possibilities. This study was run in a between-subjects design with 18 participants from age 22 to 45 ( $\bar{m} = 30.5, SD = 8.8$ , 10 identified as males). Among them, 10 had never used an AR headset before.

The experimental task consisted in judging affordances and the space occupied by one's body within a real environment while being seated 4 meters away from it. The environment was a spatial arrangement of diverse objects (blocks, holds, chains), laid onto or above a platform made with tables (see Fig. 3.1, left). The arrangement was revealed to the participants at the last minute. Then, they had to go through one of the following conditions:

- **Condition "R" (Reality):** the participants had to answer questions asked orally about their ability to perform actions (e.g. touching an object) or about where some body parts would arrive if they were at specific locations, in specific postures (e.g. head position when sitting on a block).
- **Condition "AR":** the participants had to implement the actions and postures that were inquired about in condition R with their self-avatar (Fig. 3.2-3). After each action, they had to answer the question related to this action that was asked in condition R, this time seeing their avatar in place while still being able to manipulate it.

Participants remained seated the entire time in either condition. The questions asked after each action aimed to evaluate the participant's judgment of affordances and their accuracy when mentally projecting themselves at the observed locations (see Appendix 1 for details). They were of the following types:

- **Yes-or-no questions:** "By looking at [*some place*], do you think you could [*do some action*] if you were [*located in some place, in some posture*]?" What is your level of confidence from 1 to 10?". Example of answer: "— Yes, 8"
- **Estimation questions:** For these questions, we projected a scale with linear non-standard units (i.e. not part of an existing metric system) on the wall behind, as in Fig. 3.4. "By looking at [*some place*], which unit do you think [*some body part*] would reach if you were [*located in someplace, in some posture*]?" With what margin of error?". Example of answer: "— Unit 31, plus or minus 2 units".

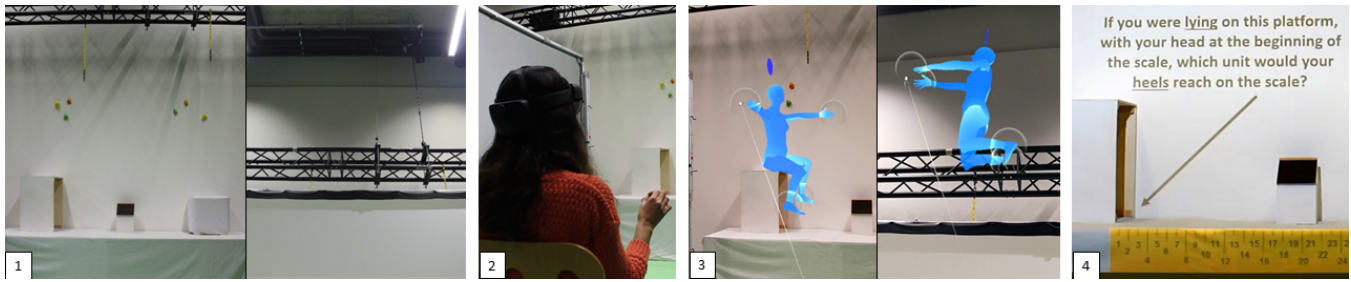


Figure 3: Illustration of the user study. 1) Testing environments on the ground and the beam. 2) Subject manipulating their self-avatar during the AR condition. 3) The view of this subject in the two environments. Note that the “free rotation” button was set above the avatar’s head for the study and the virtual menu allowing to switch modes was removed to avoid confusion. 4) Example of estimation-type question that we asked.

Because the participants had various body sizes, we adapted the placement of the objects so that the difficulty would not vary across users. We did not compare the user’s accuracy when using their avatar to when using other measurement tools (e.g. AR ruler). Indeed, our objective was not to measure distances remotely, as we could do with a telemeter, but really to better understand the extent to which an externalization of our body may help us to perceive our possibilities for action in a distant physical surrounding. Similarly, we did not measure completion time as it was out of our research scope.

In the AR condition, the participants were told that their avatars had the same body proportions as them. This condition was preceded by an eye calibration for the AR headset and by a short training session (10 minutes). During training, participants were presented with their self-avatar and instructed to manipulate it as dictated by the experimenter. They were also asked to observe its similarity with their body shape and size by walking around it and comparing the lengths of their limbs. Before starting the experiment, the participants were invited to sit on a chair and close their eyes. The experimenter would then take the AR headset back, uncover the environment of the experiment and scan it again to ensure it was properly detected by the HoloLens 2. They then returned the HMD to the participant, sat out of sight, and instructed them to open their eyes again. The hands-on time with the system lasted about 25 minutes during which the experimenter could see the participants’ viewpoint and their interactions with the avatar via a live video stream from the HMD.

The same number of participants experienced each condition (9). After going through their condition, the participants were invited to fill out a questionnaire assessing their subjective experience of the system. The whole session lasted about 45 minutes.

## 5.2 Experiment 2

The second user study was designed to further assess the strength of such embodiment experiences by checking whether one’s perception of real environments could also be improved **after** having used the system, and not only while using it.

This study was run in a between-subjects design with 16 participants from age 22 to 58 ( $\bar{m} = 30.6, SD = 10.5$ , 9 identified as males). Most had previous experience with AR headsets (12 of them). The apparatus, task, evaluation, and conditions were the same as in Experiment 1. The only difference is that the entire set of evaluation questions was asked at the end of the condition, after having performed all of the instructed actions. In the R condition, participants were first instructed to mentally visualize themselves performing all of the actions. In the AR condition, the HMD was removed from the participants before starting the evaluation so that they could no longer see their avatar when answering. None of the participants were aware of the type of questions that they would get, and they were not specifically instructed to memorize what they saw when implementing the actions with their avatar.

The hands-on time with the system was about 15 minutes. All participants went through both the R and AR conditions in counterbalanced order, but within different environments: either on the same grounded platform as in Experiment 1 or on a metallic beam suspended above this platform (see Fig. 3.1, right). The goal of this design was to prevent potential learning effects. The attribution of these environments was also distributed to either condition in a counterbalanced order. After going through both conditions, the participants were invited to fill out a subjective questionnaire similar to the one in Experiment 1 (see Section Sect. 5.4). The whole session lasted approximately 1 hour.

## 5.3 Results

To better understand the inherent strengths of this system for judging affordances and projecting oneself mentally, we assessed three main aspects: (i) the quantitative accuracy of user judgments, (ii) their level of confidence during evaluation, and (iii) the qualitative appreciation of the system. The first two were evaluated from the answers to the evaluation questions for each separate study, whereas the last was evaluated with the subjective questionnaires. One participant was removed from the analysis of each study due to incorrect body measurements that impacted the avatar’s perception.

### 5.3.1 Accuracy assessment

To evaluate errors, we used a theoretical ground truth for each question. This ground truth was obtained by using the body measures that participants had provided for Virtual Caliper and complementary measures taken at the end of the experiment. We computed individual scores for each condition and question type.

- The scores of the yes-or-no questions were computed by averaging their answers, coded with 1 or 0 (for true or false). Scores ranged from 0 to 1, with 1 meaning that all answers were true.
- The score of the estimation questions was computed by averaging the difference between the participant’s answer and the ground truth (i.e. the unit they should have been able to reach). A low score means the participant was accurate.

We then performed a descriptive analysis of these scores whose results are summarized with boxplots in Fig. 4. Shapiro-Wilk tests show that the scores did not follow a normal distribution, so we used non-parametric tests to evaluate the significance of differences.

**Experiment 1** The mean scores of the yes-or-no questions were 0.73 for the R condition and 0.95 for the AR condition. This means that the average success rate was close to 100% in the AR condition. The mean scores of the estimation questions were 4.93 for the R condition and 2.01 for the AR condition. Being closer to 0, the mean of the AR condition indicates that participants made smaller errors

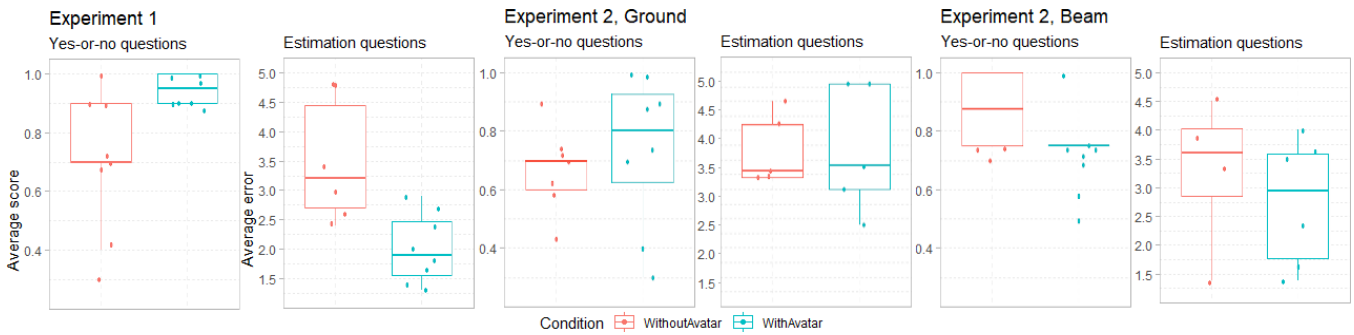


Figure 4: Boxplots representing the scores of the yes-or-no and estimation questions for both Experiment 1 and 2.

than in the R condition. Looking at the error values, it appears that the participants of the R condition tended to underestimate their body sizes, whereas those of the AR condition were closer to the ground truth and sometimes slightly overestimated their sizes.

We applied Wilcoxon rank-sum tests which showed that the scores were significantly different across the R and AR conditions for both types of questions (yes-or-no:  $p = 0.049^*$ , estimation:  $p = 0.003^{**}$ ), with a **moderate** effect size for the yes-or-no questions ( $r = 0.49$ ) and a **large** effect size for the estimation questions ( $r = 0.74$ ). We conclude that the participants who used an avatar to answer the evaluation questions were more accurate than those who did not have an avatar to do so.

**Experiment 2** Generally speaking, participants seem to have underestimated their body sizes in both R and AR conditions, regardless of the environment (beam or ground). Our statistical analysis did not show clear evidence that having used the avatar led participants to have a more accurate perception after removing the AR headset for either environments (Wilcoxon rank sum tests, yes-or-no questions:  $p_{ground} = 0.34$ ,  $p_{beam} = 0.19$ ; estimation questions:  $p_{ground} = 0.69$ ,  $p_{beam} = 1$ ). Further study will be needed to determine if manipulating a self-avatar in unexplored distant locations allows assimilating an experience that can be used from memory.

### 5.3.2 Confidence level

We ran a between-subjects analysis to compare the confidence ratings and margins of errors given by the participants in the R and AR conditions. Shapiro-Wilk tests show that the data of Experiment 1 follow a normal distribution, but not Experiment 2. We still used non-parametric tests for both as the data is not continuous and the number of participants is rather small.

**Experiment 1** The average confidence rating was 7.78 for the R condition and 8.66 for the AR condition. This means that the average confidence ratings were closer to the maximum confidence level (value of 10) in the AR condition. The average margin was 1.6 for the R condition and 1.05 for the AR condition. Being closer to 0, the mean of the AR condition indicates that participants estimated they made smaller errors than in the R condition.

The results of the Wilcoxon rank sum tests show that the margin size given during the estimation questions was significantly different across the R and AR conditions ( $p = 0.017^*$ ) with a **large** effect size, but not the confidence rating of the yes-or-no questions ( $p = 0.135$ ). We conclude that the participants who used an avatar to answer the estimation-type questions were more accurate than those who did not have an avatar.

**Experiment 2** During the evaluation, participants frequently accounted for the difficulty of the questions and communicated they were very unsure of their answers. Despite frequent subjective feedback suggesting that they were more confident in the AR condition, we found no significant difference in the average levels of confidence

and error margins between the R and AR conditions (Wilcoxon rank sum tests, yes-or-no questions:  $p_{ground} = 1$ ,  $p_{beam} = 0.81$ ; estimation questions:  $p_{ground} = 0.40$ ,  $p_{beam} = 0.30$ ).

## 5.4 Subjective feedback

The post-experiment subjective questionnaire of Experiment 1 contained 11 items on a 7-point Likert scale and 5 comment boxes letting participants write their thoughts on the avatar’s appearance, integration in the real world, control, the help it provided, and their general appreciation of the system. The questionnaire of Experiment 2 contained the same comment boxes and questions, except for item 7 which was reformulated, and for items 8 and 9 which were removed. Fig. 5 shows the results for the questions of Experiments 1 and 2.

### 5.4.1 Help provided by the avatar

We received quite positive feedback regarding the help that the avatar provided in Experiment 1: 89% of the participants judged that the avatar helped them gain confidence when answering (item n°7). Written comments included: “*Without the avatar, it would have been difficult to evaluate the answers to the questions*”, “*The avatar is of great help*”. Interestingly, the subjective feedback of Experiment 2 also shows that 80% of the participants found the avatar helpful despite not seeing it during the evaluation. Several participants of this experiment mentioned that the avatar had allowed them to correct their perception of distance or size: “*I saw myself much smaller, it allowed me to raise my estimates*”, “*I could better realize the relative position of each object. For example, ah, the distance between the chain and the block is not that big*”.

### 5.4.2 Trust in the avatar

Although many of the participants confessed that they were surprised by the places they could reach with their avatars in Experiment 1, the majority seems to have trusted what they saw and relied on it to answer. One participant commented the following: “*I felt I could completely trust the avatar as I had compared myself to it, so I also felt it was a good representation of me in space*”, and “*felt quite sure of my answers*”. In Experiment 2, one participant wrote that they also relied on their experience with the avatar despite not being explicitly instructed to. Surprisingly, two participants from Experiment 1 decided not to use their avatars to answer the evaluation questions. They explained that it seemed to have the same body size as them when standing up close to it, but that it appeared bigger when it was farther away. They preferred to rely on their own impression, leading them to answer with lower estimations than what the avatar indicated. Still, they may have been influenced by the embodiment of their avatars as their scores were higher than those who did not have one. Further investigation is thus necessary to clarify whether such behavior is due to depth perception issues or a general mistrust of the technology.

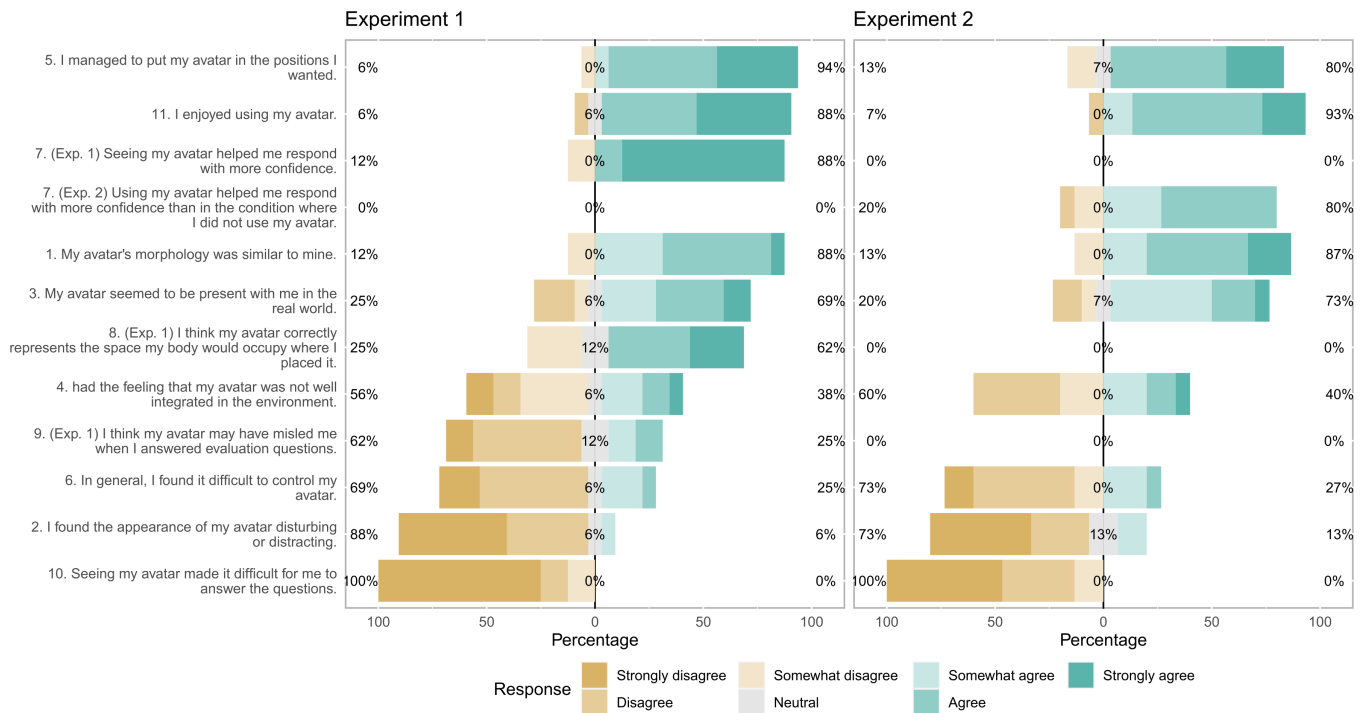


Figure 5: Results of the subjective post-experiment questionnaires of Experiments 1 and 2 (translated from French). Note that some questions were asked only for Experiment 1 or 2. Detailed percentages are available in Appendix 2

### 5.4.3 Avatar appearance

Most participants reported that the avatar's body resembled their own in both experiments (item n°1). One participant shared the following: *"Initially, I thought she was bigger than me in terms of scale [...] but when I walked close and compared my arm length and height, etc. I felt I could confirm she was very similar to me"* (Exp. 1). This feeling was often expressed orally during the training session by other participants. Among the two that did not agree with item n°1, one participant from Experiment 2 observed that the avatar's distribution of fat was quite different from their own. This is probably linked to the limited number of input parameters of Virtual Caliper, which does not include traits like muscularity. Lastly, the participants did not find their avatar's appearance to be disturbing or distracting (all except 2 in Exp. 2). One participant shared that they *"enjoyed the neutrality of the appearance"* (Exp. 1). Among those that did not like it, one mentioned that the idle animation made them uncomfortable (Exp. 2).

### 5.4.4 3D registration

The questionnaire also seems to show that the avatar was usually perceived as being well registered in the environment. However, some participants experienced environment detection issues that led them to have more mitigated answers. In Experiment 2, these issues occurred when the spatial mesh built by the HoloLens 2 was updated inaccurately, which happened more often when the AR condition was performed on the beam due to its angle of view. The device sometimes interpreted the environment to be closer to the participant than it was. As a result, the avatar appeared occluded by a virtual wall or residual artifacts. Nevertheless, 75% of the participants reported that the avatar felt "present" with them in the real world (item n°3), which suggests that they were usually able to ignore the detection errors when they occurred.

### 5.4.5 Avatar control

Regarding the control of the avatar, 88% of the participants reported being successful in putting it in the positions they wanted (item n°5), and 75% reported they did not find controlling it difficult (item n°6). We collected comments such as *"quite intuitive"*, *"really surprising"*, or *"easy to take in hand"*. Room for improvement was pointed out regarding the rotation of the avatar: *"It would be better if we could choose the axis of rotation"*. Another participant suggested including more feedback to better perceive when the avatar is in contact with real surfaces. The HMD's limitations seem to have added difficulty to the avatar's manipulation: participants sometimes lowered their hands too much for the headset to see them and their gestures were no longer detected. It also happened regularly that the gesture was not understood despite being detected, and this led participants to repeat their movements several times before succeeding. The frequency of these errors usually decreased as they progressed through the experiment, which suggests that more training might have been required. This was accounted for by one of the participants: *"I, for sure, had a learning curve, but towards the end, I found it actually quite easy to manipulate her"*.

### 5.4.6 User engagement

Lastly, all participants except one reported that they enjoyed using the avatar. The only participant that disliked using it had done the AR condition in the beam environment in Experiment 2, and commented the following: *"It is more the location of the avatar (far and high) that is a pain rather than the control of the avatar itself"*. This feeling was shared by several others who reported they found it hard to select the spheres depending on their angle of view and that this slowed them down. The comment boxes included positive feedback such as *"quite fun"*, *"very playful"*, or *"strangely pleasant"* despite these difficulties. We expect the next generation of AR technology and improvements in the proposed control modes may solve the usability issues that they described.



## 6 LESSONS LEARNED AND DISCUSSION

We explored how to take advantage of 3PP virtual embodiment to access locations without physically entering them. By matching the avatar's body size and morphology with its user, we provided a visual reference that can be manipulated and used as a means of comparison and simulation to better understand one's environment. We ran an exploratory study with the Puppeteering mode of our proof-of-concept and found that 3PP AR embodiment could successfully enhance the perception of physical space and estimation confidence: (i) as expected, participants were more accurate when estimating their ability to act (moderate effect size) and the space occupied by their body (large effect size) with the help of their self-avatar, and (ii) participants were more confident when performing mental projections of their body size (large effect size). These results obtained in Experiment 1 allowed validating our approach and making sure that AR technology was reliable enough for such usage despite well-known issues linked to them.

The majority of participants found the Puppeteering mode to be useful, usable, and fun. We learned that the interaction technique we used to rotate the avatar needs to be improved as it was considered laborious by some participants. One participant suggested that it would be convenient to have a physical doll to put the avatar in the desired posture, as previously proposed in research in other contexts [26]. It could be interesting to study such a technique as it could allow for more direct and efficient enacting, as it does not rely on the detection of gestures that can be faulty.

Moreover, as the participants did not have the possibility to stand up and change their angle of view during the experiments, they couldn't see parts of the avatar that were occluded by the avatar itself. Although none of the evaluation questions relied on this, it made it hard to tell when the avatar was in contact with surfaces behind. To remedy this problem, multimodal feedback could be provided (e.g. sound, visual cues, vibratory feedback) to inform the user when and where their avatar comes in contact with real surfaces that are occluded by other parts of its body. This could be useful in situations in which it is impossible to walk and change one's angle of view due to contextual or environmental restrictions.

Through the exploratory study, we also wanted to verify whether a self-avatar could go beyond being a visual reference and provide a way to gain experience that can be used after manipulating it. We could not find clear evidence that it was the case with our participants as we found no significant difference between the confidence and accuracy of their estimations, with and without having used their avatars. This can be explained by several factors.

First, the design of our experiment has probably played a role: to avoid a transfer effect, we decided not to provide a clear goal by asking them to memorize the position of their avatar and to use this memory to answer the questions afterward. Being uninformed of the type of questions we would ask, it seems that most participants focused their entire attention on achieving the instructed actions with their avatar and did not pay further attention to its position relative to the environment. Some participants confirmed this fact: *"I was more focused on manipulating the avatar than on its size in space"*. Experiment 2, therefore, raises several questions on memory encoding and attentional tunneling for future work to explore.

Second, it seems that the environment detection provided by the AR headset was not perfect, and it is likely that it negatively impacted the perception of the avatar's position relative to the environment. To better quantify the error introduced by AR headsets like the HoloLens 2, future work could put in place a Just Noticeable Difference (JND) study [6]. This would also allow determining the impact of visualizing self-avatars on the affordance perception and mental projection with more precision.

## Future Work

The exploratory study we presented only investigated the use of the Puppeteering mode to improve affordance perception. As a follow-up, future work could assess and compare the usability of all three control modes by measuring time performance and the user's sense of embodiment. Additionally, it would be interesting to reproduce a similar experiment with more participants and with different types of avatars. Medeiros et al. [21] previously identified that the avatar's appearance could impact spatial awareness in VR, but this fact was not studied in AR. We recommend pursuing research on this effect.

Secondly, AR displays still have limitations that impede exploiting avatar embodiment to its full potential. In particular, they do not offer the range necessary to interact with faraway content. It is likely that using the avatar in distant locations will not allow gaining as precise information. It would be interesting for future work to evaluate this aspect when progress in AR will allow it. Additionally, contact identification between real and virtual surfaces is a well-known perceptual issue of XR interfaces. This is partly because rendering light and shadows on virtual objects in real-time is still a hard problem. Without shadows, holograms seem to float in mid-air instead of resting against surfaces. This is an ongoing problem, and we did not evaluate the impact of such an issue on physical affordance perception. It will need to be investigated in future work.

Finally, the system we implemented can be expanded with other modes and improved with countless other techniques to adapt to more various and specific situations. Creating an ultimate system implementing a myriad of modes was outside the scope of the present study, but we seek to bring light to the many possibilities that are available for future work. The following list provides some examples of features that could complement the core system we proposed:

**Record feature.** It can be difficult to observe the avatar's body movement if one has to perform the same movement at the same time. Therefore, it might be useful to record and rewind this movement at a different pace as previously proposed in Remixed Reality [16].

**Contact feedback.** As suggested by some participants, identifying when the avatar is colliding with physical surfaces hidden behind the avatar itself can be hard. We imagine that the inclusion of haptics could allow for the user to feel surfaces that the avatar touches remotely. Visual highlights or sound cues could also be implemented [33, 40].

**Affordance detection.** Areas of the environment that afford specific actions could be highlighted and labeled (e.g. "grabbable", "walkable") to help the user visualize all possibilities at once.

**Physical abilities calibration.** By measuring and modeling user traits such as flexibility, strength, and stamina, one could more precisely calibrate the avatar to the user's body capabilities. This would allow making it more representative of its user.

**Duplication of the avatar.** It could be interesting to allow the user to manipulate several copies of their avatar all at once, or individually e.g. in authoring scenarios where the user needs to have a side-by-side visualization of different actions.

## 7 CONCLUSION

In this paper, we explored how self-avatars and MR can be used to extend our perception of physical environments by expanding our range of action to areas outside of our peripersonal space. We designed an AR embodiment system allowing users to control a self-avatar from a third-person perspective with three control modes. We used one of them to evaluate the validity of our approach through an exploratory study. Our results highlight how such use of avatars has the potential to improve the user's understanding of their options during decision-making. Lastly, we contribute with lessons learned from this design process and provide guidelines for future work seeking to implement virtual embodiment in third-person AR. The presented work may serve as a starting point for future research aiming to explore this promising potential of embodiment experiences.

## REFERENCES

- [1] S. Cmentowski, A. Krekhov, and J. Krüger. Outstanding: A multi-perspective travel approach for virtual reality games. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play, CHI PLAY '19*, p. 287–299. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3311350.3347183
- [2] S. H. Creem-Regehr, K. T. Gagnon, M. N. Geuss, and J. K. Stefanucci. Relating spatial perspective taking to the perception of other's affordances: providing a foundation for predicting the future behavior of others. *Frontiers in Human Neuroscience*, 7:596, 2013. doi: 10.3389/fnhum.2013.00596
- [3] F. Daiber, F. Kosmalla, and A. Krüger. Boulder: Using augmented reality to support collaborative boulder training. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems, CHI EA '13*, p. 949–954. Association for Computing Machinery, New York, NY, USA, 2013. doi: 10.1145/2468356.2468526
- [4] C. Dede. Immersive Interfaces for Engagement and Learning. *Science*, 323(5910):66–69, Jan. 2009. doi: 10.1126/science.1167311
- [5] M. R. Everingham, B. T. Thomas, and T. Troscianko. Head-Mounted Mobility Aid for Low Vision Using Scene Classification Techniques. *International Journal of Virtual Reality*, 3(4):1–10, Jan. 1998. doi: 10.20870/IJVR.1998.3.4.2629
- [6] G. Fechner. *Elements of psychophysics. Vol. I*. New York, 1966.
- [7] T. Feuchtnner and J. Müller. Extending the Body for Interaction with Reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*, pp. 5145–5157. ACM Press, Denver, Colorado, USA, 2017. doi: 10.1145/3025453.3025689
- [8] A. C. S. Genay, A. Lecuyer, and M. Hachet. Being an avatar “for real”: a survey on virtual embodiment in augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–1, 2021. doi: 10.1109/TVCG.2021.3099290
- [9] J. J. Gibson. *The ecological approach to visual perception: classic edition*. Psychology Press, New York, 2014.
- [10] G. Gorisse, O. Christmann, E. A. Amato, and S. Richir. First-and third-person perspectives in immersive virtual environments: presence and performance analysis of embodied users. *Frontiers in Robotics and AI*, 4:33, 2017.
- [11] S. Ishak, K. E. Adolph, and G. C. Lin. Perceiving affordances for fitting through apertures. *Journal of Experimental Psychology: Human Perception and Performance*, 34(6):1501, 2008.
- [12] Y. Itoh, J. Orlosky, K. Kiyokawa, and G. Klinker. Laplacian Vision: Augmenting Motion Prediction via Optical See-Through Head-Mounted Displays. In *Proceedings of the 7th Augmented Human International Conference 2016*, pp. 1–8. ACM, Geneva Switzerland, Feb. 2016. doi: 10.1145/2875194.2875227
- [13] A. Klippel, P. Sajjadi, J. Zhao, J. O. Wallgrün, J. Huang, and M. M. Bagher. Embodied digital twins for environmental applications. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 4:193–200, 2021.
- [14] R. Komiya, T. Miyaki, and J. Rekimoto. Jackin space: Designing a seamless transition between first and third person view for effective telepresence collaborations. In *Proceedings of the 8th Augmented Human International Conference, AH '17*. Association for Computing Machinery, New York, NY, USA, 2017. doi: 10.1145/3041164.3041183
- [15] S.-w. Leigh and P. Maes. AfterMath: Visualizing Consequences of Actions through Augmented Reality. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, CHI EA '15*, pp. 941–946. Association for Computing Machinery, New York, NY, USA, Apr. 2015. doi: 10.1145/2702613.2732695
- [16] D. Lindbauer and A. D. Wilson. Remixed Reality: Manipulating Space and Time in Augmented Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, pp. 1–13. ACM Press, Montreal QC, Canada, 2018. doi: 10.1145/3173574.3173703
- [17] F. Liu and S. Seipel. On the precision of third person perspective augmented reality for target designation tasks. *Multimedia tools and applications*, 76(14):15279–15296, 2017.
- [18] M. Loper, N. Mahmood, J. Romero, G. Pons-Moll, and M. J. Black. Smpl: A skinned multi-person linear model. *ACM transactions on graphics (TOG)*, 34(6):1–16, 2015.
- [19] L. S. Mark. Eyeheight-scaled information about affordances: a study of sitting and stair climbing. *Journal of experimental psychology: human perception and performance*, 13(3):361, 1987.
- [20] L. S. Mark, J. A. Balliett, K. D. Craver, S. D. Douglas, and T. Fox. What an actor must do in order to perceive the affordance for sitting. *Ecological Psychology*, 2(4):325–366, 1990.
- [21] D. Medeiros, R. K. dos Anjos, D. Mendes, J. M. Pereira, A. Raposo, and J. Jorge. Keep my head on my shoulders! why third-person is bad for navigation in vr. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology, VRST '18*. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3281505.3281511
- [22] M. Merleau-Ponty. *Phénoménologie de la perception*, p. page 160. Gallimard, Paris, 1945.
- [23] C. Michaels. Care llo, c.(1981). direct perception, 1981.
- [24] T. Miyaki and J. Rekimoto. LiDARMAN: reprogramming reality with egocentric laser depth scanning. In *ACM SIGGRAPH 2016 Emerging Technologies on - SIGGRAPH '16*, pp. 1–2. ACM Press, Anaheim, California, 2016. doi: 10.1145/2929464.2929481
- [25] C. Nimcharoen, S. Zollmann, J. Collins, and H. Regenbrecht. Is that me?—embodiment and body perception with an augmented reality mirror. In *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 158–163. IEEE, 2018.
- [26] N. Numaguchi, A. Nakazawa, T. Shiratori, and J. K. Hodgins. A puppet interface for retrieval of motion capture data. In *Proceedings of the 2011 ACM SIGGRAPH/Eurographics Symposium on Computer Animation - SCA '11*, p. 157. ACM Press, Vancouver, British Columbia, Canada, 2011. doi: 10.1145/2019406.2019427
- [27] S. Pujades, B. Mohler, A. Thaler, J. Tesch, N. Mahmood, N. Hesse, H. H. Bühlhoff, and M. J. Black. The Virtual Caliper: Rapid Creation of Metrically Accurate Avatars from 3D Measurements. *IEEE Transactions on Visualization and Computer Graphics*, 25(5), May 2019. doi: 10.1109/TVCG.2019.2898748
- [28] Pärtel Lang. Puppet master, 2022.
- [29] J. S. Roo and M. Hachet. One Reality: Augmenting How the Physical World is Experienced by combining Multiple Mixed Reality Modalities. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, pp. 787–795. ACM, Québec City QC Canada, Oct. 2017. doi: 10.1145/3126594.3126638
- [30] N. Rosa, J.-P. van Bommel, W. Hurst, T. Nijboer, R. C. Veltkamp, and P. Werkhoven. Embodying an Extra Virtual Body in Augmented Reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 1138–1139. IEEE, Osaka, Japan, 2019. doi: 10.1109/VR.2019.8798055
- [31] P. Salamin, T. Tadi, O. Blanke, F. Vexo, and D. Thalmann. Quantifying effects of exposure to the third and first-person perspectives in virtual-reality-based training. *IEEE Transactions on Learning Technologies*, 3(3):272–276, 2010. doi: 10.1109/TLT.2010.13
- [32] X. Sanchez, P. Lambert, G. Jones, and D. J. Llewellyn. Efficacy of pre-ascent climbing route visual inspection in indoor sport climbing. *Scandinavian Journal of Medicine & Science in Sports*, 22(1):67–72, 2012. doi: 10.1111/j.1600-0838.2010.01151.x
- [33] E. Schoop, J. Smith, and B. Hartmann. HindSight: Enhancing Spatial Awareness by Sonifying Detected Objects in Real-Time 360-Degree Video. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–12. ACM, Montreal QC Canada, Apr. 2018. doi: 10.1145/3173574.3173717
- [34] T. A. Stoffregen. Affordances as properties of the animal-environment system. *Ecological psychology*, 15(2):115–134, 2003.
- [35] J. B. Wagman and K. R. Taylor. Perceiving affordances for aperture crossing for the person-plus-object system. *Ecological Psychology*, 17(2):105–130, 2005.
- [36] I. Wang, J. Smith, and J. Ruiz. Exploring Virtual Agents for Augmented Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–12. ACM, Glasgow Scotland UK, May 2019. doi: 10.1145/3290605.3300511
- [37] W. H. Warren Jr and S. Whang. Visual guidance of walking through

apertures: body-scaled information for affordances. *Journal of experimental psychology: human perception and performance*, 13(3):371, 1987.

- [38] E. Wu and H. Koike. FuturePong: Real-time Table Tennis Trajectory Forecasting using Pose Prediction Network. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–8. ACM, Honolulu HI USA, Apr. 2020. doi: 10.1145/3334480.3382853
- [39] A. Zoran and J. A. Paradiso. FreeD: a freehand digital sculpting tool. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2613–2616. ACM, Paris France, Apr. 2013. doi: 10.1145/2470654.2481361
- [40] M. Zöllner, S. Huber, H.-C. Jetter, and H. Reiterer. NAVI – A Proof-of-Concept of a Mobile Navigational Aid for Visually Impaired Based on the Microsoft Kinect. In D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, M. Naor, O. Nierstrasz, C. Pandu Rangan, B. Steffen, M. Sudan, D. Terzopoulos, D. Tygar, M. Y. Vardi, G. Weikum, P. Campos, N. Graham, J. Jorge, N. Nunes, P. Palanque, and M. Winckler, eds., *Human-Computer Interaction – INTERACT 2011*, vol. 6949, pp. 584–587. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011. Series Title: Lecture Notes in Computer Science. doi: 10.1007/978-3-642-23768-3\_88