

Detection Thresholds for Rotation and Translation Gains in 360° Video-based Telepresence Systems

Jingxin Zhang, Eike Langbehn, Dennis Krupke, Nicholas Katzakis and Frank Steinicke, Member, IEEE

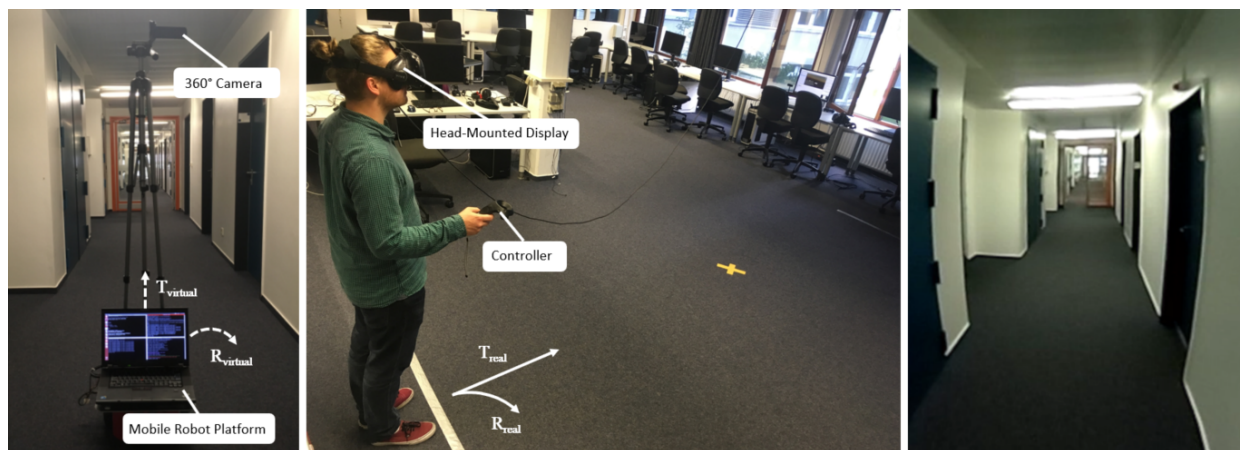


Fig. 1. Illustration of the concept of a redirected walking telepresence system based on translations and rotations: (left) the mobile platform is equipped with a 360° video camera moving in the remote environment (RE), (center) the user wears a virtual reality head-mounted display (HMD) walking in the local environment (LE), and (right) the user's view of the RE on the HMD.

Abstract—Telepresence systems have the potential to overcome limits and distance constraints of the real-world by enabling people to remotely visit and interact with each other. However, current telepresence systems usually lack natural ways of supporting interaction and exploration of remote environments (REs). In particular, single webcams for capturing the RE provide only a limited illusion of spatial presence, and movement control of mobile platforms in today's telepresence systems are often restricted to simple interaction devices. One of the main challenges of telepresence systems is to allow users to explore a RE in an immersive, intuitive and natural way, e. g., by real walking in the user's local environment (LE), and thus controlling motions of the robot platform in the RE. However, the LE in which the user's motions are tracked usually provides a much smaller interaction space than the RE. In this context, redirected walking (RDW) is a very suitable approach to solve this problem. However, so far there is no previous work, which explored if and how RDW can be used in video-based 360° telepresence systems. In this article, we conducted two psychophysical experiments in which we have quantified how much humans can be unknowingly redirected on virtual paths in the RE, which are different from the physical paths that they actually walk in the LE. Experiment 1 introduces a discrimination task between local and remote translations, and in Experiment 2 we analyzed the discrimination between local and remote rotations. In Experiment 1 participants performed straightforward translations in the LE that were mapped to straightforward translations in the RE shown as 360° videos, which were manipulated by different gains. Then, participants had to estimate if the remotely perceived translation was faster or slower than the actual physically performed translation. Similarly, in Experiment 2 participants performed rotations in the LE that were mapped to the virtual rotations in a 360° video-based RE to which we applied different gains. Again, participants had to estimate whether the remotely perceived rotation was smaller or larger than the actual physically performed rotation. Our results show that participants are not able to reliably discriminate the difference between physical motion in the LE and the virtual motion from the 360° video RE when virtual translations are down-scaled by 5.8% and up-scaled by 9.7%, and virtual rotations are about 12.3% less or 9.2% more than the corresponding physical rotations in the LE.

Index Terms—Virtual reality, telepresence, 360° camera, locomotion

1 INTRODUCTION

Telecommunication and remotely controlled operations are becoming increasingly common in our daily lives. Such telepresence technology has enormous potential for different application domains ranging from business, tourism, meetings, entertainment to academic conferences [35, 58], education [37, 45], and remote health care [1, 22]. The ideal goal for teleoperation is that users feel as if they were actually present at the remote site during the teleoperation task [54]. This illusion is referred to as the *sense of (tele-)presence* based on the so-called *place illusion* [44]. In currently available telepresence systems the sensation of presence is severely limited and therefore the presence illusion is often not evoked [54].

Among the many types of telepresence systems, our work focuses on systems for exploring remote sites, which aims to overcome the limitation of distance in order to allow people to interact and commu-

- Jingxin Zhang, Eike Langbehn and Dennis Krupke are doctoral students at the Human-Computer Interaction (HCI) group at the Universität Hamburg. E-mail: {jxzhang, langbehn, krupke}@informatik.uni-hamburg.de.
- Nicholas Katzakis is a Postdoctoral research associate at the HCI group at the Universität Hamburg. E-mail: nicholas.katzakis@uni-hamburg.de.
- Frank Steinicke is Full Professor and Head of the HCI group at the Universität Hamburg. E-mail: frank.steinicke@uni-hamburg.de.

Manuscript received 11 Sept. 2017; accepted 8 Jan. 2018.

Date of publication 19 Jan. 2018; date of current version 18 Mar. 2018.

For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.

Digital Object Identifier no. 10.1109/TVCG.2018.2793679

nicate over long distances and visit remote environments (REs) [36]. Telepresence systems required to achieve this usually consist of several technological components like cameras and microphones that capture live data in the RE and transfer it to the local user who can explore the RE, for example, by means of vision or hearing. Mobile platforms can carry these sensors and move through the RE under the control of the local user, who can change the position, orientation and perspective in the remote space [24]. At the local site, telepresence components usually consist of display devices, which enable users to perceive the streamed data from the RE, or input devices that can be used to control the remote mobile platform [54].

Despite advancements in the field of autonomous mobile robots, most of today's mobile robots still require the supervision of a human user. An important challenge related to the overall telepresence experience is the design of the user interface to control the mobile platform. Movement controls of mobile platforms in telepresence systems nowadays often rely on simple interaction devices such as joysticks, touchscreens, mice or keyboards [54]. As such operators have to use their hands in order to control the mobile platform and, therefore, the hands are not free to simultaneously perform other tasks. This may decrease the naturalness, task performance and overall user experience [36]. For example, even though it is known that real walking is the most presence-enhancing way of exploring a virtual space, real walking as a method to explore a RE is usually not possible, despite the general idea having been introduced more than a decade ago [36]. In addition, most current telepresence platforms consist of mobile webcams with speakers and microphones. As a result, the usage of single webcams for capturing the RE provides the users with a very narrow field of view and a limited illusion of spatial presence. Both issues limit an operator's sense of tele-presence [54]. Furthermore, the deficiency of visual information about the RE can lead to a high error-rate for teleoperation tasks and remote interactions [54].

In order to address these limitations and challenges, we introduce the idea of an improved telepresence system consisting of a head-mounted display (HMD), a 360° video camera and a mobile robot platform. The idea of this telepresence system is that the mobile robot, which is equipped with the 360° full-view video camera, can be remotely controlled by means of a real walking local user. The camera captures a 360° live stream from the RE and transfers this full-view live stream via a communication network in real-time to the user's LE. The received live stream is integrated into a spherical video, which is rendered in a 3D engine, and presented to the user via the HMD. This way, the user can experience the RE in real-time. A 360° camera and HMD form the basis of our telepresence system, which aims to increase sensation of presence and the user's spatial perception compared to a 2D narrow view.

To control the mobile base, we choose real walking in the local space as a travel technique because it is the most basic and intuitive way of moving within the real world compared to any other input device [25, 47]. When using real walking, a HMD user could literally walk through the local space and virtually explore the RE. In principle, movements of the user would be detected by a tracking system in the local space and transferred to the RE to control the mobile base. Since the position of the mobile base in the remote space would be determined and updated according to the position of the user in local space, this approach provides the most consistent and intuitive perception of motion in the target environment, releasing user's hands for other potential interactive teleoperation tasks as well. This walking approach is only feasible if the layouts of local and remote space are more or less identical. In most cases, however, the available local tracked space is smaller than the RE, which the user wants to explore, and furthermore, local and remote environments typically have completely different layouts. Redirected walking (RDW) is a technique to overcome the limits and confined space of the tracked room [42]. While RDW is based on real walking, the approach guides the user on a path in the real world, which might vary from the path the user perceives in the virtual environment (VE). This is done through manipulations applied to the VE, causing users to unknowingly compensate for scene motions by repositioning and/or reorienting themselves [53]. RDW without the

user's awareness is possible because the sense of vision often dominates proprioception [3, 9]; hence, slight differences between vision and proprioception are not noticeable in cases where the discrepancy is small enough [47]. While previous work has investigated the human sensitivity to such manipulations in computer-generated imagery only, so far, it is unknown how much manipulation can be applied to a mobile robot platform, which transfers 360° videos of real-world scenes rather than computer-generated images. Furthermore, it seems reasonable to assume that there are significant differences in the perception of self-motions in computer-generated images and 360° live streams from the real world, due to differences in visual quality, image distortion or stereoscopic disparity.

Therefore, we conducted two psychophysical experiments to investigate the amount of discrepancy between movements in the LE and the RE that can be applied without users noticing. We designed two experiments in order to find the thresholds for two basic motions, i. e., translation and rotation, in 360° video-based environments. The results of these experiments provide the basis for future immersive telepresence systems in which users can naturally walk around to explore remote places using a local space that has a different layout.

To summarize, the contributions of this article include:

- introduction of the concept of a redirected walking robotic platform based on a 360° video camera,
- a psychophysical experiment to identify detection thresholds for translation gains, and
- a psychophysical experiment to identify detection thresholds for rotation gains for controlling such a platform.

The remainder of this article is structured as follows: Section 2 summarizes previous related work. Section 3 explains the concept of using RDW for mobile 360° video-based telepresence systems. The two psychophysical experiments are described in Section 4. Section 5 provides a general discussion of the findings of the experiments. Section 6 concludes the article and gives an overview about further work.

2 RELATED WORK

In this section we summarize work related to telepresence systems, mobile robotic platforms, locomotion in general and detection thresholds in psychophysics.

2.1 Telepresence Systems

Telepresence refers to a set of technologies, which aim to convey the feeling of being in a different place than the space where a person is physically located [54]. Therefore, telepresence systems should allow humans to move through the remote location, interact with remote artifacts or communicate with the remote people. Such telepresence systems have been developed since the beginning of the 1990s [13]. In this context, the term presence [44] describes the *place illusion*, which denotes the illusion of being in a different environment, in which events occur in a plausible way, i. e., *plausibility illusion*. Telepresence systems therefore refer to the special case that the illusion of presence is generated in a spatially distant real-world environment [54].

Currently available telepresence systems are often based on the “window-on-a-world” metaphor, where a computer screen becomes a transparent window for video transmissions, through which two groups of participants in geographically different locations (usually rooms) can communicate with each other using video-based representations. In contrast to traditional video conferencing systems, such telepresence systems offer integrated tracking systems that enable participants to move their heads to explore the RE [29–32]. Recent approaches also support spatial 3D audio, which creates the impression that a participant actually speaks in a specific position/direction in the adjacent room. Thus, the most important natural forms of communication are supported in such face-to-face conferences. However, the spatial separation by the “window-on-a-world” metaphor cannot be canceled. Strictly speaking, these telepresence systems do not provide the impression of being in a

RE, but rather only allow two distant environments to be viewed with a certain degree of geometrical correctness.

The TELESAR V system [54] is a telepresence masterslave robot system that was developed to realize the concept of haptic telepresence. Motions of the full upper body are mapped between the local human operator and the remote system. Walking, however, is not possible with such a system. Nitzsche et al. [36] introduced a telepresence system, in which a HMD user could steer a remote robot by real walking in a confined space. Therefore, they introduced the concept of motion compression to arbitrarily walk in large RE without making use of scaling or walking-in-place metaphors. In contrast to the work presented in this article, motion compression maps both travel distances and turning angles with a ratio 1:1, where straight forward motions are bent to curves. Furthermore, their system was not equipped with a 360° camera, but used two regular cameras for stereoscopic imaging. Kasahara et al. [20] designed JackIn Head, a visual telepresence system with an omnidirectional wearable camera, which can share one's first-person view and immersive experience to the remote people via Internet. However, none of the previous work has considered detection thresholds for motions between LE and RE [20, 36, 54].

In order to provide a common space for telepresence systems, sometimes computer-rendered VEs are used, which in this case provide *virtual telepresence systems*, e. g., SecondLife¹, ActiveWorlds², Facebook Spaces³, AltspaceVR⁴ or OpenSimulator⁵. Such VEs may be implemented, for example, by immersive display technologies, e. g., Oculus Rift HMD, or stereoscopic projection screens [2]. While these systems make it possible for several participants to be present in a common virtual space, those environments are purely virtual and thus do not correspond to the original idea of telepresence [54]. In addition, such systems present a number of limitations, for example, these systems do not make it possible to explore real-world objects or environments without complex pre-processed digitalization or virtualization processes.

2.2 Robotic Platforms

Mobile camera systems on motion platforms—sometimes referred to as *video collaboration robots*—can be used to enable participants to control their viewing direction in the RE. Since the 1990s, remote motion platforms have been used in various fields of applications. Traditional applications can be found in military use as well as for fire fighting or other dangerous situations [4, 55], while recent applications also find their way into office spaces [28, 35]. For example, the Double Robotics⁶ serves as a modern video collaboration robot for office environments based on a Segway motion platform and an iPad-based video conferencing application. The system has been designed to enable remotely working cooperators to communicate with each other.

Although, such systems allow for controlling the camera's direction (i. e., the viewing direction of the user), there are various limitations. In particular, current solutions do not cater for an immersive experience when using the telepresence systems, since life-size 3D representations or calibrated geometric egocentric perspectives are not possible, and therefore, significantly reduce the sense of presence, space perception and social communication [54]. Moreover, the currently used directional control mechanisms (e. g., joysticks, mouse or keyboard) do not allow natural control by the head or body of the participants as in a real situation.

2.3 Locomotion

In recent years, different solutions are used to make it possible for users to explore VEs, which are significantly larger compared to the available tracking space in the real world. Several of these approaches are based

on specific hardware developments such as motion carpets [43], torus-shaped omni-directional treadmills [5, 6], or motion robot tiles [16–18]. As an cost-effective alternative to these hardware-based solutions, some techniques were introduced, which take advantage of imperfections in the human perceptual system. Examples include concepts such as virtual distractors [40], change blindness [52, 53], or impossible and flexible spaces [56, 57]. In their taxonomy [51], Suma et al. provide a detailed summary and classification of different kinds of redirection and reorientation solutions in a range from subtle to overt, as well as from discrete to continuous approaches. The solution adopted in our work belongs to the class of techniques that reorient users by continuous subtle manipulations. In this situation, when users explore a VE by walking in the tracked space manipulations (such as slight rotations) are applied to the virtual camera [41, 46]. Based on these small iterative rotating manipulations, the user is forced to adjust the walking direction by means of turning to the opposite direction of the applied rotation. As a result, the user walks on a curvature in the real space while she perceives the illusion to walk along a straight path in the VE. In other words, the visual feedback that the user sees on the HMD corresponds to the motions in the VE, whereas proprioception and vestibular system are connected to the real world. If the discrepancy between stimuli is small enough, it is difficult for the user to detect the redirection, which leads to the illusion of an unlimited natural walking experience [41, 46].

2.4 Detection Thresholds

Identifying detection threshold between motion in the real-world and those displayed in the VE has been in the topic of several recent studies. In his dissertation, Razzaque [41] reported that a 1 deg/s manipulation serves as lower detection threshold. Steinicke et al. [47] described a psychophysical approach to identify discrepancies between real and virtual motions. Therefore, they introduced gains to map users' movements from the tracked space in the real world to camera motion in the VE. In this context, they use three different gains, i. e., (i) rotation, (ii) translation and (iii) curvature gains, which scale a user's rotation angle, walking distance and bending of a straight path in the VE to a curved path in the real world respectively. In addition, they determined detection thresholds for these gains through psychophysical experiments, by which the noticeable discrepancies between visual feedback in the VE on the side and proprioceptive and vestibular cues on the other side in the real world are identified. For example, to identify detection thresholds for curvature gains, participants were asked to walk a straight path in the VE, while in the real world they actually walked a path, which was curved to the left or right using a randomized curvature gain. Participants had to judge whether the path they walked in the real world was curved to the left or to the right using a *two-alternative forced-choice (2AFC)* task. Using this method, Steinicke et al. [47] found that users can not reliably detect manipulations when the straight path in the VE is curved to a circular arc in the real world with a radius of at least 22m. In recent work it has been shown that these thresholds can be increased, for instance, by adding passive haptic feedback [33] or by constraining users to walk on curved paths instead of straight paths only [25].

Several other experiments have focused on identifying detection thresholds for such manipulations during head turns and full body turns. For instance, Jerald et al. [19] suggest that users are less likely to notice gains applied in the same direction as head rotation rather than against head rotation. According to their results, users can be physically turned approximately 11% more and 5% less than the virtual rotation. For full-body turns, Bruder et al. [8] found that users can be physically turned approximately 30% more and 16% less than the virtual rotation. In a similar way, Steinicke et al. [47] found that users can be physically turned approximately 49% more and 20% less than the virtual rotation.

Furthermore, Paludan et al. [38] explored if there is a relationship between rotation gains and visual density in the VE, but the results showed that the amount of visual objects in the virtual space had no influence on the detection thresholds. However, other work has shown that walking velocity has an influence on the detection thresholds [34]. Then, another study by Bruder et al. [7] found that RDW could be affected by cognitive tasks, or in other words, RDW induce some

¹www.secondlife.com

²www.activeworlds.com

³www.facebook.com/spaces

⁴www.altvr.com

⁵www.opensimulator.org

⁶www.doublerobotics.com

cognitive effort on users.

While the results mentioned above have been replicated and extended in several experiments, all the previous analyses have considered computer-generated VEs only, whereas video-based streams of real scenes have not been in the focus yet.

3 REDIRECTED WALKING TELEPRESENCE SYSTEMS

In this section, we describe the concept and the challenges of using redirected walking in the context of a mobile 360° video-based telepresence system.

3.1 Concept and Challenges

As described above, one of the main challenges of telepresence systems is to allow users to explore a RE by means of real walking in the user's LE, and thus controlling the motion of the robot platform in the RE. However, usually the available local tracked space is smaller than the RE that the user wants to explore, and furthermore, local and remote environments typically have dissimilar layouts.

For computer-generated VEs, RDW has been successfully used to guide users. Hence, RDW seems to be a very suitable approach to solve this problem also in the context of a mobile 360° video-based telepresence system. However, while in VR environments RDW is based on manipulating movement of the virtual camera, such approaches cannot be directly applied to manipulations of the real camera due to latency issues, mechanical constraints, or limitations in the precision and accuracy of robot control. Figure 1 illustrates the concept of using RDW for 360° video-based mobile robotic platforms. We suppose that both the tracking system in the LE and the coordinate systems in RE are calibrated and registered. When users wearing an HMD perform movements in the LE, their position change can be detected by a tracking system in real-time. Such a change in position can be measured by the vector $T_{real} = P_{cur} - P_{pre}$, where P_{cur} and P_{pre} mean the current position and the previous position. Normally, T_{real} is mapped to the RE by means of one-to-one mapping when a movement is tracked in the LE. With respect to the registration between the RE and the tracking coordinate system, the physical camera (attached to the robot platform) is moved by T_{real} distance in the corresponding direction in VEs. One essential advantage of using a 360° video stream for a telepresence system is that rotations of the user's head can be mapped one-to-one without the requirement that the robot needs to rotate. This is due to the fact that the 360° video already provides the spherical view of the RE.

In a computer-generated VE, the tracking system will update multiple times every second (e.g., with 90Hz), and the VE is rendered accordingly. However, due to the constraints and latency caused by network transmission of video streams and the robot platform, which needs to move, such constant real-time updates are not possible in telepresence setups. Instead, the current video data from the camera capturing the RE is transmitted and displayed with a certain delay to the HMD. However, the user can change the orientation and position of the virtual camera inside the spherical projection with an update rate of 90Hz, but needs to wait for the latest display from the RE until the robot platform has moved and re-sent an updated image again.

We implemented a first prototype of a RDW telepresence system using the above mentioned hardware and concept. This prototype is shown in Figure 1(left). The experiments described in Section 4 are based on this prototype and exploit the videos captured with the system. However, currently the prototype is not suitable for real-time use yet due to the latency of movement control and image update. However, we assume that future telepresence setups will allow lower latency communication similar to what we have today in purely computer-generated VEs.

3.2 RDW Gains for 360° Videos

As described in Section 2, Steinicke et al. [47] introduced translation and rotation gains for computer-generated VEs. In this section, we explain the usage of translation and rotation gains in our concept of a 360° video-based setup. Furthermore, we describe how the application of such gains can influence user movements.

3.2.1 Translation Gains

We refer to camera motions, which are used to render the view to the RE, as *virtual translations* and *virtual rotations*. The mapping between real and virtual motions can be implemented as follows: We define a translation gain as the quotient of the corresponding virtual translation $T_{virtual}$ and the tracked real physical translation T_{real} , i.e., $g_T = \frac{T_{virtual}}{T_{real}}$. When a translation gain g_T is applied to a real-world movement T_{real} , the virtual camera is moved by $g_T \cdot T_{real}$ in the corresponding direction in the VE. This approach is useful in many situations, especially, when the user needs to explore a RE, which is much smaller or larger than the size of the tracking space in the LE. For example, for exploring a molecular structure with a nano-scale robot by means of real walking, the movements in the real world have to be compressed a lot with a $g_T \approx 0$, whereas the exploration of a larger area on a remote planet with a robot vehicle by means of real walking may need a translation gain like $g_T \approx 50$.

Translation gains can be also denoted as $g_T = \frac{v_{virtual}}{v_{real}}$, where v_{real} means the speed of physical movement in LE and $v_{virtual}$ means the speed of virtual movements showing the RE. In addition, the position changes in the real world can be actually performed in three orientations at the same time [48], which includes fore-aft direction, lateral and vertical motions. In our experiments we focus on the translation gains in the direction of the actual walking direction, which means that only the movements in fore-aft direction are tracked, whereas the movements in lateral and vertical directions are filtered [14].

3.2.2 Rotation Gains

In a similar way, a rotation gain can be defined as the quotient of the mapped rotation in a virtual space and the real rotation in the tracked space: $g_R = \frac{R_{virtual}}{R_{real}}$, where $R_{virtual}$ is the virtual rotation and R_{real} is the real rotation. When a rotation gain g_R is applied to a real rotation R_{real} in the LE, the user sees the resulting virtual rotation of the RE given by $g_R \cdot R_{real}$. That means, when $g_R = 1$ is applied, the rendered view to the RE remains static during a user's change of the head orientation since the 360° already provides the spherical view. However, if $g_R > 1$, the remotely displayed virtual 360° scene that the user views on the HMD, will rotate against the direction of the head rotation and, therefore, the rotation will appear faster than normal. In the opposite case $g_R < 1$, the view to the RE rotates with the direction of head rotation, and will appear more slowly. For example, when a user rotates her head in the LE by 90°, a gain of $g_R = 1$ will be applied in a one-to-one mapping to the virtual camera, which makes the virtual camera also rotate 90° in the corresponding orientation. For $g_R = 0.5$ the user rotates 90° in the real world while they view only a 45° orientation change in the VE displayed on the HMD. Correspondingly, for the gain $g_R = 2$ a physical rotation of 90° in the real world is mapped to a rotation in the VE by 180°.

Again, rotations can be performed in three orientations at the same time in the real world, i.e., yaw, pitch and roll. However, in our experiment we focused on the rotation gain for yaw rotation only, since yaw manipulations are used most often in RDW as it allows to steer users towards the desired directions, for instance, in order to prevent a collision in the LE [19, 23, 39, 49].

3.2.3 Other Gains

In principle, all other gains introduced for RDW such as curvature gains [47] or bending gains [25] are possible with 360° videos as well. Nevertheless, the focus of this work was on evaluating the user's sensitivity to detecting thresholds for rotation and translation gains, we will not discuss those gains in more detail.

4 EXPERIMENTS

In this section, we describe the psychophysical experiments in which we analyzed the detection thresholds for translation and rotation gains in the 360° video environment. Since both experiments used similar material and methods, we describe the setup and procedure first, and then explain each experiment in detail.

4.1 Hardware Setup

The experiment was performed in a $12m \times 6m$ laboratory room (see Figure 2 and 5). During the experiment all participants wore an HTC Vive HMD, which displays the 360° video-based RE with a resolution of 1080×1200 pixels per eye. The diagonal field of view is approximately 110° and the refresh rate is 90Hz. For tracking the user's position, we use a pair of lighthouse tracking stations delivered with the HTC Vive. The lighthouse tracking system was calibrated in such a way that the system provides a walking space of $6m \times 4m$. During the experiments, the lab space was kept dark and quiet in order to reduce interference with the real-world. Experimental instructions were shown to the participants by means of slides displayed on the HMD only. Participants used an HTC Vive controller as input device to perform the operations described below and answer questions after each trial. For rendering the RE as well as system control, we used an Intel computer, which had a 3.5GHz Core i7 processor, 32GB of main memory, and two NVIDIA GeForce GTX 980 graphics cards. Furthermore, participants answered questionnaires on an iMac computer. The 360° experimental video-based RE was recorded by the RICOH THETA S camera, which was attached to the robot platform (see Figure 1(left)). It has a still image resolution up to 5376×2688 pixels and a live streaming resolution up to 1920×1080 pixels, which we used in Unity3D Engine 5.6. We connected the HMD with the link box using a HTC Vive 3-in-1 (HDMI, USB and Power) 5m cable in such a way that participants could move freely within the tracking space during the experiment. Considering the constraints and latency caused by network transmission of video streams, the 360° video of RE for the experiments was recorded with a 1280×720 resolution and a 15fps frame rate, which is consistent with the real use of the RDW telepresence system prototype as described in Section 3.

4.2 Two-Alternative Forced-Choice Task

In order to identify the amount of deviations between physical movements in the LE and the virtual movements as shown from the RE, which are unnoticeable to users, we used a standard psychophysical procedure based on the method of constant stimuli in a two-alternative forced-choice (2AFC) task. In this method, the applied gains are presented randomly and uniformly distributed instead of appearing in a specific order [25, 47].

After each trial, participants have to choose one of two possible alternatives such as in our case “smaller” and “larger”. Even though, in several situations, it is difficult to correctly identify the answer, participant would need to choose the answer randomly, and will be correct in 50% on average. The *point of subjective equality* (PSE) is defined as the gain for which the participants answer “smaller” in 50% of the trials. At the PSE, participants perceive the translation or rotation in the RE and in the LE as identical. When the gain decreases or increases from the PSE, it becomes easier to detect the discrepancy between movements in the RE and in the LE. Typically, this results in a psychometric curve. When the answers reach a chance level of 100% or 0%, it is obvious and easy for the participants to detect the manipulations. A threshold can be described at the gain at which participants can just sense the difference between physical motions in the LE and virtual motion displayed on the HMD. However, stimuli at values close to thresholds could be often perceptible. Hence, thresholds are determined by a series of gains where the participants can only sense the manipulations with some probability. Typically for psychophysical experiments, the point where the psychometric curve reaches the middle between the 0% chance level and 100% is regarded as a detection threshold (DT). Thus, the lower DT for gains smaller than the PSE value is defined as the gain where the participants answered in 75% of all trials with “smaller” on average. Similarly, the upper DT for gains larger than the PSE value is the gain where participants have just answered in 25% of all trials “smaller” on average.

In this article, we analyze the range of gains for which users are not able to reliably detect the discrepancy as well as the gain at which users perceive motions in the LE and in the RE as equal. The 25% to 75% DTs shows a gain interval of potential manipulations, which can be applied for RDW in 360° video-based REs. Moreover, the PSE values



Fig. 2. Illustration of the experimental setup: A user is walking straight forward in the LE to interact with the 360° video-based RE. Translation gains are applied to change the speed of displayed virtual movement from RE. The inset shows the users view to the 360° video environment, which shows a corridor from the RE.

indicate how to map the user motions in the LE to the movements of the telepresence robot in the RE, such that the visual information displayed on the HMD appears naturally to the users.

4.3 Experiment 1 (E1): Difference between Virtual and Physical Translation

In E1, we investigate the participant's ability to discriminate whether a physical translation in the LE was slower or faster than the virtual translation displayed in the 360° video-based RE. We instructed the participants to walk a fixed distance in the LE and mapped their movements to a pre-recorded 360° video-based RE.

4.3.1 Methods for E1

We pre-recorded a 360° video with the telepresence system prototype described in Section 3 showing a forward movement in the RE with a normal walking of speed of 1.4m/s [50]. The height displayed in the 360° video was recorded at a height of 1.75m.⁷ We manipulated the speed of the video based on the walking speed measured in the LE by applying the described translation gains in such a way that the speed in the video was manipulated accordingly. This means that when the user walked with 1.4m/s in the LE, the video was displayed in normal speed, whereas when the user decreased the speed and stopped, the video was slowed down with the gains until it was paused. The video showed a movement in the fore-aft direction, and all other micro head movements were implemented as micro motions of the virtual camera inside a 360° video-based spherical space. Changes of the head orientation were implemented using a one-to-one mapping.

Figure 2 illustrates the setup for E1. For each trial, participants were guided to the start line and held an HTC Vive controller. When participants were ready, they clicked the trigger button to display the 360° video, which presented the RE on the HMD, and started to walk in the LE. The play speed during walking was adjusted to the participant's physical speed in real-time. For instance, if the participants stopped, the scene of the RE displayed on the HMD would also pause. The walking velocity was determined by movements along the main direction of the corridor shown in the 360° video. During the experiment, we used different translation gains to control the play speed of the 360° video. For example, when walking with the translation gain g_T , the 360° video would be played in the speed of $g_T \cdot v_{real}$, where v_{real} is

⁷We could not find any significant effect of the deviation from the user's actual eye height and the recorded height on the estimation of the detection thresholds.

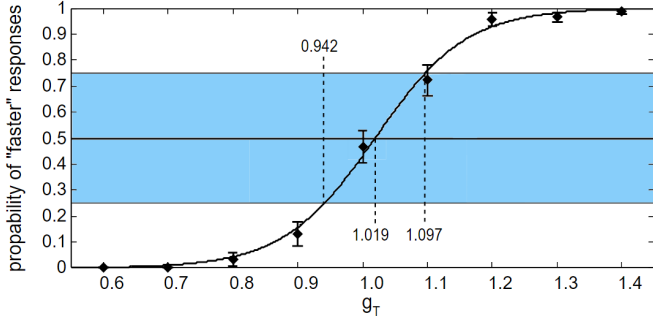


Fig. 3. Pooled results of the discrimination between movements displayed from the RE and movements performed in the LE. The x axis shows the applied translation gain g_T , the y axis shows the probability that participants estimated the virtual straightforward movement displayed as 360° video faster than the actually performed physical motion.

the participant's real-time speed along the fore-aft direction in the LE. When the participant traveled 5m in the LE and crossed the end line, the RE displayed on the HMD would automatically disappear. Then, the participant had to estimate whether the virtually displayed motion was faster or slower than the physical translation in the LE (in terms of distance, this corresponds to longer or shorter). Participants had to provide their answer by using the touch pad on the HTC Vive controller. After each trial, the participants walked back to the start line, while they were guided by visual markers displayed on the HMD, and then clicked the trigger again to start the next trial. For each participant we tested 9 different gains in the range of $\{0.6, 1.4\}$ in steps of 0.1 and repeated each gain 6 times. Hence, in total, each participant performed 54 trials in which they walked a 5m distance in the LE, while they viewed virtual distances within a range of $\{3m, 7m\}$ for each trial. All of the trials appeared in randomized order. After each trial, the participant turned back to the start orientation with the help of the markers displayed on the HMD, and clicked the trigger button again to continue with the next trial.

4.3.2 Participants of E1

16 participants (14 male and 2 female, age 19-37, $M=26.4$) participated in E1, in which we explored the participant's sensitivity to translation gains. One participant could not complete the experiment because of cyber sickness. All data from the remaining participants was included in the analyses. Most of the participants were students or members from our local department of computer science. All participants had normal or corrected to normal vision. 5 participants took part in the experiment with glasses. None of the participants suffered from a disorder of equilibrium. Four of the participants reported dyschromatopsia, strong eye dominance, astigmatism and night blindness, separately. There are no other vision disorders reported by the participants. The experience of the participants with 3D stereoscopic displays (such as cinema or games) was $M = 2.4$ within the range of 1 (no experience) to 5 (much experience). 14 participants have worn HMDs before. Most of them had experiences with 3D computer games ($M = 3.2$ with a range of 1 = no experience to 5 = much experience) and played 4.4 hours per week on average. The body height of the participants varied between 1.60m - 1.90m ($M = 1.80m$).

The experimental process for each participant included pre- an post-online-questionnaires, instructions, training trials, experiment, and breaks, the total time for each participant was about 40 - 50 minutes. The participants needed to wear the HMD for around 25-30 minutes. During the experiment, the participants were allowed to take breaks at any time.

4.3.3 Results of E1

Figure 3 shows the mean probability over all participants that they estimate the virtual straightforward movement shown on the HMD as faster than the physical motion for different translation gains. The

error bars show the standard errors. Translation gains g_T lead to faster virtual straightforward movements (relative to the physical movements) if $g_T > 1$. Then, participants would feel that they move a larger distance in the RE than in the LE. A gain of $g_T < 1$ results in a virtual translation movement, which is slower than the physical walking speed, resulting in a shorter distance displayed from the RE. We fitted a psychometric function of the form $f(x) = \frac{1}{1+e^{a-x+b}}$ with real numbers a and b .

From the psychometric function a slight bias for the PSE was determined at $PSE = 1.019$. In order to compare the found bias from the gain of 1.0, we performed a one sample t-test, which did not show any significant difference ($t=1.271$, $df=14$).

The results for the participant's sensitivity to translation gains show that gains from 0.942 to 1.097 (25% and 75% DT) cannot be reliably detected. This means that within this range participants were not able to reliably discriminate whether a physical translation in the LE was slower or faster than the virtual translation displayed from the 360° video RE.

4.3.4 Discussion of E1

The results show that participants could not discriminate the difference between physical translation performed in the LE and virtual translation perceived from the RE, when the movement is manipulated with a gain in a range from 5.8% slower to 9.7% faster than the real movement. From the definition of translation gains, a $PSE = 1.019$ indicates that the virtual translations displayed from the 360° video-based RE are slightly faster than the physical translation in the LE [12, 14, 15, 27]. A translation gain $g_T = 1.019$ appeared natural to the participants, which means that walking a distance of 4.91m in the LE felt like traveling 5m in the RE. Therefore, participants tended to travel a shorter physical distance in the LE when they tried to approach the same expected virtual distance in the 360° video-based REs. In addition, a PSE larger than 1 is consistent with the results from previous research on translations in fully computer-generated VEs [46, 47]. However, the bias reproduced in our experiment was not statistically significant.

According to the results from previous research for computer-generated VEs [46], we expected a slight shift of the psychometric function and detection thresholds towards the larger gains also for 360° video environments. Visual analysis from Figure 3 shows such a slight shift, and both the 25% and 75% detection thresholds are slightly shifted towards the larger gains. However, this shift is smaller than the ones reported in previous work. Furthermore, an interesting observation is that the 25% and 75% detection thresholds for the translation gains are both closer to the PSE value in the 360° video environment compared to the results from previous research in fully computer-generated VEs [46, 47].

4.4 Experiment 2 (E2): Difference between Virtual and Physical Rotation

4.4.1 Methods for E2

In E2, we analyzed the ability of participants to discriminate between virtual rotations displayed from the RE and physical rotations performed in the LE. Figure 4 shows the setup for the experiment. During the experiment, the participants wore an HMD and were placed inside the tracking system. They were instructed to perform rotations in the LE, which were tracked and displayed as virtual rotations in the 360° video-based RE.

We used a 360° full-view image to create a spherical projection in Unity3D, which presented a 360° outdoor RE. Rotation gains were applied to the yaw rotation only. Again, the view height in the RE was adjusted to 1.75m. At the beginning of the experiment, participants were instructed to stand in the center of the tracking space and hold a HTC Vive controller. The participants could start the next trial by clicking the trigger button on the controller. Now, they could see the video stream from the RE (Figure 4) to which we applied a randomized rotation gain, when they started to turn. Participants saw a green ball in front of their view at the eye-level that marked the start point for the rotation. An arrow showed the rotation direction that participants were required to follow. The participants were told to rotate in the



Fig. 4. Illustration of the experimental setup: a user is performing rotations in the LE to interact with the 360° video-based RE. Rotation gains are applied in the experiment to change the speed of virtual rotations displayed from the RE. The inset shows the users view to the 360° video environment, in which a start point and a directional arrow are displayed.

corresponding direction until a red ball appeared in the front of their view, which indicated the end point of the rotation. The angle between the start point (green ball) and the end point (red ball) was adjusted to 90°. Hence, the virtual rotation shown on the HMD from the RE was always 90°, but the physical rotation participants performed in the LE was different according to the corresponding rotation gains. During the experiment, different rotation gains were applied to the virtual rotations showing the RE. A rotation gain $g_R = 1$ shows a one-to-one mapping between physical rotation in the LE and virtual rotation displayed from the RE. However, for example, when a rotation gain satisfies $g_R < 1$, the virtual scene on the HMD rotates with the direction of the real physical rotation in the LE and slowed down the change in the RE. In the opposite case, for a rotation gain $g_R > 1$, the scene in the RE rotates against the direction of the real physical rotation in the LE, and accelerated the change of VE.

For each participant, we tested 9 different gains in a range of $\{0.6, 1.4\}$ by steps of 0.1 and repeated each gain 6 times. Hence, each participant performed a series of physical rotations in the LE with a range of $\{64.29^\circ, 150^\circ\}$ to achieve a 90° virtual rotation in the virtual RE. In order to study the effects of different rotation direction we considered rotations to the left and to the right. Therefore, in total, there were 108 trials for each participant. All of the trials appeared in randomized order. Then, participants had to choose whether the perceived rotation from the RE was smaller or larger than the physical rotation performed in the LE. Again, responses had to be given via the touchpad of HTC Vive controller. After each trial, the participant turned back to the start orientation with the help of the markers displayed on the HMD, and clicked the trigger button again to continue with the next trial.

4.4.2 Participants of E2

17 participants (13 male and 4 female, age 24 - 38, $M=29.5$) took part in the second experiment analyzing the sensitivity to rotation gains. Two participants stopped the experiment because of suffering from motion sickness. The data from the remaining 15 participants were included in the analyses.

Most of the participants were students or members of our local department of computer science. All participants had normal or corrected to normal vision. 3 participants wore glasses during the experiment, and 1 participant wore contact lenses. 1 participant reported to suffer from a disorder of equilibrium. 2 participants reported strong eye dominance and night blindness. No other vision disorders have been reported by the participants. Most of the participants had experiences

with 3D stereoscopic displays before with $M = 2.76$ in a range of 1 (no experience) to 5 (much experience). 14 participants had experiences using HMDs before, and 13 of them had experiences with 3D computer games with $M = 2.71$, and played games with an average time of 5.26 hours per week. The body height of the participants was in a range of 1.60m - 1.92m with $M = 1.75$ m.

The total time of the experimental procedure for each participant including pre-online-questionnaires, instructions, a few training trials, experiment, breaks and post-online-questionnaires, took almost 40 - 50 minutes. The participants wore the HMD for about 25 - 30 minutes. During the experiment, the participants were allowed to take breaks at any times.

4.4.3 Results of E2

To verify the influence of different rotation orientations, we analyzed the data of rotations to the left (cf. Figure 5(a)) as well as rotations to the right (cf. Figure 5(b)). In our experiment, a rotation gain g_R results in a smaller physical rotation than the virtual rotation if $g_R > 1$. This means that participants rotate less in the LE than in the RE. A rotation gain lead to a larger physical rotation than the virtual rotation if $g_R < 1$. In other words, participants would rotate more in the LE compared to the rotation they view in the RE. We fitted the data with the same psychometric function as in Experiment E1.

Figure 5(a) presents the mean probability over all participants that they estimated the virtual rotations to the left smaller in the RE than the physical rotations in the LE with different applied rotation gains. The error bars show the standard errors. The psychometric function determined a bias for the PSE at $PSE = 0.984$. The 25% and 75% detection thresholds for rotation gains were found at 0.877 and 1.092. Within this range of gains participants were not able to reliably discriminate whether a physical rotation to the left in the LE was smaller or larger than the corresponding virtual rotation displayed from the 360° RE. Figure 5(b) presents the situation in which rotations were performed to the right. For the PSE we derived $PSE = 0.972$, and the gains between the detection thresholds of 25% and 75% were from 0.892 to 1.054.

In order to compare the found bias from the gain of 1.0, we performed a one sample t-test, which did not show any significant difference for rotations to the left ($t=-0.429$, $df=14$) or rotations to the right ($t=-1.466$, $df=14$). Furthermore, no significant differences between rotations to the left and right were found ($t=0.472$, $df=14$).

4.4.4 Discussion of E2

For a physical rotation to the left, the participants could not discriminate the difference between physical rotations in the LE and perceived virtual rotations from the RE when rotation gains were within a range of $\{0.877, 1.092\}$. That means that the virtual rotation in the RE is 12.3% less and 9.2% more than the physical rotation in the LE. A rotation gain $g_R = 0.984$ appeared most natural, indicating that participants have to rotate for 91.46° in the LE to perceive the illusion that they actually rotated by 90° in the RE.

For a physical rotation to the right, the range of rotation gains that participants could not reliably detect as manipulation between physical rotations in the LE and virtual rotations in the RE is $\{0.892, 1.054\}$. In other words, for the virtual rotation in the RE the participants could accept a 10.8% smaller or 5.4% larger physical rotation in the LE without noticing the discrimination. The most natural rotation gain for rotations to the right is $g_R = 0.972$, which indicates that participants need to rotate for 92.59° in the LE to feel that they rotate 90° in the remote space.

As described above, independent from direction of the rotation, the most natural rotation gains for the participants are slightly smaller than 1, which suggest that the participants need to rotate more in the LE to perceive the illusion that they have already rotated the same expected angle in the 360° REs. However, this bias was not statistically significant. These results show an opposite effect to what we have found from the translation experiment, but a PSE smaller than 1 appears to be consistent with the results from previous research on the rotation in fully computer-generated VEs [10, 19, 47]. Moreover, our results indicate that the range of gains, which can be applied to 360° environments and be

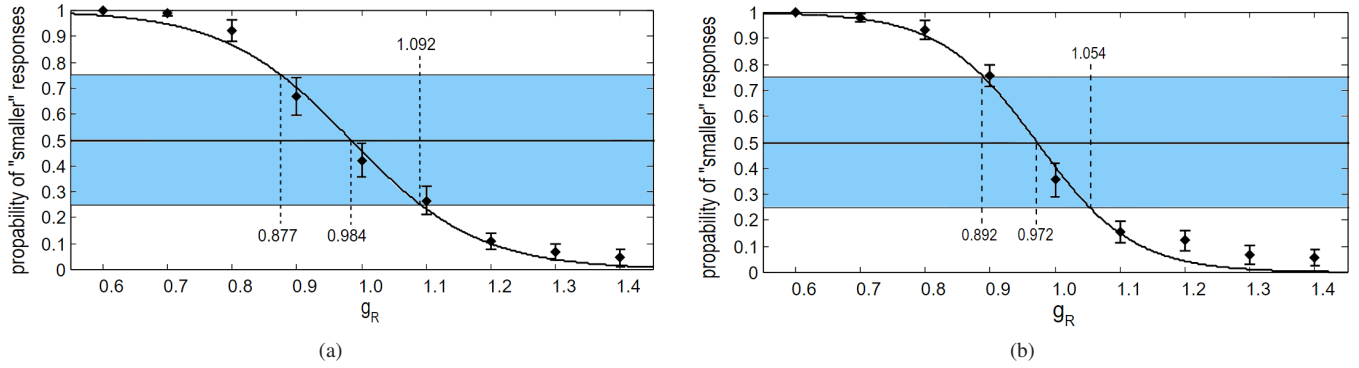


Fig. 5. Pooled results of the discrimination between remote virtual and local physical rotations towards (a) left and (b) right. The x axis shows the applied rotation gain g_R , the y axis shows the probability that participants estimated the virtual rotation as smaller than the physical rotation.

unnoticeable to the participants, are narrower than the results reported in earlier work for purely VEs. Hence, our results suggest again that participants have a better discrimination ability for manipulations of rotations in a 360° RE compared to rotations in a purely computer-generated VE. We will discuss further on this point in the general discussion.

Furthermore, the results indicate that the interval of detection thresholds for manipulations of rotations to the right is smaller than the manipulations of rotations to the left in the 360° RE. This means that participants have provided more accurate estimations for rotating to the right than to the left. Such a finding has not been reported in earlier work. One possible explanation of the observed phenomenon might be related to the structure of the brain and hand dominance; since most of our participants were right-handed, however, this has to be verified in further research.

In summary, there is a range of rotation gains, in which participants could not reliably discriminate between physical rotations in the LE and virtual rotations in the 360° REs.

4.5 Post-Questionnaires

After the experiments, the participants answered further questionnaires in order to identify potential drawbacks of the experimental design. Participants estimated whether they feel that the 360° RE surrounded them (0 corresponds to fully disagree, 7 corresponds to fully agree). For the translation experiment E1 the mean value was 4.4 (SD = 1.76), and for the rotation experiment E2 the average value was 5.2 (SD = 1.29). Hence, most of the participants agree that when using our telepresence system they perceived a high sense of presence. Furthermore, we asked the participants how confident they were that they chose the correct answer (0 corresponds to very low, 4 corresponds to very high). The average value for answers to this questions was 2.53 (SD = 0.83) for the translation experiment and 2.29 (SD = 1.06) for the rotation experiment.

After both experiments, we also measured simulator sickness by means of Kennedys Simulator Sickness Questionnaire (SSQ). For the translation experiment, the average Pre-SSQ score for all participants was 6.23 (SD = 9.34) before the experiment, and an average Post-SSQ score of 26.68 (SD = 27.80) after the experiment. For the rotation experiment, the average Pre-SSQ score for all participants was 9.23 (SD = 21.06) before the experiment, and the average Post-SSQ score was 55.60 (SD = 68.03) after the experiment. The results show that the average Post-SSQ score after the rotation experiment was larger than after the translation experiment. This finding can be explained by the sensory-conflict theory, since continuous rotations provide more vestibular cues than constant straightforward motions. Hence, manipulations during such rotations induce more sensory conflicts [26].

5 GENERAL DISCUSSION

Our results show that participants cannot distinguish discrepancies between physical translations in the LE and perceived virtual translations in the 360° RE when the virtual translation is down-scaled by 5.8%

and up-scaled by 9.7%. A small bias for the PSE was determined in $PSE = 1.019$ indicating that slightly up-scaled virtual translations in the RE appear most natural to the users, which means that users believe they have already walked a 5m distance in the 360° video-based RE after only walking a 4.91m distance in the LE. These results are consistent with most previous findings in the fully computer-generated VEs [12, 14, 15, 27]. However, the strong asymmetric characteristic of the psychometric function, which was found in previous research on RDW in VEs could not be replicated in our experiment in which we used realistic 360° video environments.

The rotation experiment results show that when virtual rotations in the 360° video-based RE are applied within a range of 12.3% less or 9.2% more than the corresponding physical rotation in the LE, the users cannot reliably detect the difference between them. For rotations to the left, a rotation gain of $PSE = 0.984$ appears most natural to the participants, meaning that they have to rotate 91.46° in the LE to have the illusion that they have already rotated 90° in the RE. The most natural rotation gain for the right rotation is $PSE = 0.972$, which means that they need to rotate 92.59° in the LE to have the impression that they have already rotated 90° in the RE. These results also confirm previous findings to some extent [10, 19, 47]. Again, the asymmetric characteristic of the psychometric function is not so obvious for our experiment results in 360° video REs compared to previous findings for fully computer-generated VEs.

The data as well as the analysis we presented in Section 4 suggest that manipulations in 360° video-based REs have similar influence on users as manipulations in fully computer-generated VEs [47], i. e., users tend to travel a slightly shorter distance but rotate a slightly larger angle in the LE when they try to approach the same expected motion in the 360° video-based REs. However, some differences in PSE values and distribution of detection thresholds between 360° video-based REs and computer-generated VEs should also be noted. On the one hand, the PSE value for translations in 360° video-based REs is 1.019 which is much closer to a one-to-one mapping compared to previous results in VEs. The same situation can also be found from the results of the rotation experiment with a PSE value of 0.984 to the left and 0.972 to the right in 360° video REs and 0.9594 in fully computer-generated VEs. Conversely, the ranges between 25% and 75% detection thresholds for translation and rotation in 360° video REs are both smaller than the results in VEs, which indicates a smaller range for users in which they are not able to reliably discriminate the difference between the motions in a 360° video RE and in the real world. All these differences suggest that users have a more accurate ability to judge the difference between physical motions in the LE with corresponding virtual motions in a 360° video RE than in a fully computer-generated VE. However, future work is required to explore these differences in more depth.

There are a few possible explanations for this finding. First, the scenes shown to the users during our experiments are 360° videos of a remote environment in the real world rather than a computer-generated

VE. Therefore, a 360° video-based RE, which is displayed in the LE might appear more realistically and accurately with respect to size, texture or other characteristics in contrast to a fully computer-generated VE. Furthermore, since the objects in our 360° video RE are projected as spherical textures in the HMD, there was a lack of stereoscopic disparity and motion parallax, and hence users might have perceived translation and rotation based on the whole environment rather than on one or two specific objects. A lack of stereoscopic disparity and motion parallax has little influence on users concerning distance and angle perception in 360° video REs. Furthermore, we believe that the resolution of the virtual scenes presented on the HMD can also lead to different results. Usually, humans perceive the angle of a physical rotation via two sensory channels: visual information from the environment and the proprioception and vestibular information. When people perform a rotation in a VE displayed with high resolution, their sensory system will likely weight the visual information as more reliable in the multi-sensory integration process compared to scenarios with low visual resolution [11, 21]. Cues from vision will take a leading role in the perception of rotating angle, which has been shown in the literature [11, 21]. In contrast, when the scenes displayed on the HMD have low resolutions, the reliability of vision will be reduced to a certain extent, whereas the vestibular and proprioceptive cues might be evaluated with a higher weight in the multi-sensory integration process [11, 21]. The extreme situation occurs when participants close the eyes. In this case, no visual information from the environment provide cues about rotating angles, and users could only perceive information about the rotation using the vestibular system and proprioception. Hence, we believe the resolution of scenes on the HMD can also be a possible reason that caused different results between a fully computer-generated VE and a 360° video-based RE.

Similar to RDW in computer-generated VEs, in most applications a greater range of gains can also be accepted by users without noticing that they are manipulated. The detection thresholds of our experiments are conservatively estimated, since for most actual telepresence systems, the users will not be able to easily recognize the discrimination between real motion and remote virtual motion, because they need to focus on other tasks in the virtual space like object selections, manipulations etc. Hence, the upper and lower detection thresholds for translation and rotation gains found in this article serve as lower and upper bounds in the actual telepresence application.

6 CONCLUSION

In this article, we proposed the idea of a redirected walking telepresence system based on 360° videos. For this concept and the prototype, we investigated the users' ability of recognizing RDW manipulations for translation and rotation in two separate experiments. The results show that participants were not able to reliably discriminate the difference between physical motion in the LE and perceived virtual motion from the 360° RE when virtual translations are down-scaled by 5.8% and up-scaled by 9.7%, and virtual rotations are about 12.3% less or 9.2% more than the actual physical rotations. These findings provide interesting implications for future implementations of immersive telepresence systems in which users can explore remote places by means of controlling a remote robot platform by real walking.

As described above, a telepresence system based on robotic systems and on a 360° camera introduces larger latency compared to typical VR environments. In our experiment, we implemented micro movements of the user's head by micro movements of the virtual camera inside the virtual sphere on which the 360° camera images are projected. While these movements were considerably small in our experiment, for real-world applications it might be required to implement larger movements, which will then get noticeable. Again, it is an interesting question how much deviation from the projection center of the virtual sphere can be reliably detected. In addition, we would like to explore other VR setups in the LE for the described RDW telepresence system. For example, we have explored already CAVE-like setups, and are interested to see if the detection thresholds and the biases of the PSEs can be replicated in such setups. Furthermore, we will test different REs and application domains like exploration of hallways, cooperation in business meeting

rooms or inspections of outdoor scenarios to explore if the thresholds proposed in this article can be generalized to different scenarios.

ACKNOWLEDGMENTS

The authors wish to thank the participants of the experiment for their valuable contribution and feedback. Furthermore, the reviewers deserve special thanks for their helpful advice. Authors of this work receive financial support in part from the German Research Foundation (DFG) and the European Union (EU).

REFERENCES

- [1] D. Anton, G. Kurillo, A. Y. Yang, and R. Bajcsy. Augmented telemedicine platform for real-time remote medical consultation. In *International Conference on Multimedia Modeling*, pp. 77–89. Springer, 2017.
- [2] S. Beck, A. Kunert, A. Kulik, and B. Froehlich. Immersive group-to-group telepresence. *IEEE Transactions on Visualization and Computer Graphics*, 19(4):616–625, 2013.
- [3] A. Berthoz. *The brain's sense of movement*, vol. 10. Harvard University Press, 2000.
- [4] S. Bhat and M. Meenakshi. Vision based robotic system for military applications—design and real time validation. In *Signal and Image Processing (ICSIP), 2014 Fifth International Conference on*, pp. 20–25. IEEE, 2014.
- [5] L. Bouguila, M. Ishii, and M. Sato. Realizing a new step-in-place locomotion interface for virtual environment with large display system. In *Proceedings of the workshop on Virtual environments 2002*, pp. 197–207. Eurographics Association, 2002.
- [6] L. Bouguila, M. Ishii, and M. Sato. Virtual locomotion system for human-scale virtual environments. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, pp. 227–230. ACM, 2002.
- [7] G. Bruder, P. Lubas, and F. Steinicke. Cognitive resource demands of redirected walking. *IEEE transactions on visualization and computer graphics*, 21(4):539–544, 2015.
- [8] G. Bruder, F. Steinicke, P. Wieland, and M. Lappe. Tuning self-motion perception in virtual reality with visual illusions. *IEEE Transactions on Visualization and Computer Graphics*, 18(7):1068–1078, 2012.
- [9] J. Dichgans and T. Brandt. Visual-vestibular interaction: Effects on self-motion perception and postural control. In *Perception*, pp. 755–804. Springer, 1978.
- [10] D. Engel, C. Curio, L. Tcheang, B. Mohler, and H. H. Bühlhoff. A psychophysically calibrated controller for navigating through large environments in a limited free-walking space. In *Proceedings of the 2008 ACM symposium on Virtual reality software and technology*, pp. 157–164. ACM, 2008.
- [11] M. O. Ernst and M. S. Banks. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870):429–433, 2002.
- [12] H. Frenz, M. Lappe, M. Kolesnik, and T. Bührmann. Estimation of travel distance from visual motion in virtual environments. *ACM Transactions on Applied Perception (TAP)*, 4(1):3, 2007.
- [13] H. Fuchs, A. State, and J.-C. Bazin. Immersive 3d telepresence. *Computer*, 47(7):46–52, 2014.
- [14] V. Interrante, B. Ries, and L. Anderson. Distance perception in immersive virtual environments, revisited. In *Virtual Reality Conference, 2006*, pp. 3–10. IEEE, 2006.
- [15] V. Interrante, B. Ries, J. Lindquist, M. Kaeding, and L. Anderson. Elucidating factors that can facilitate veridical spatial perception in immersive virtual environments. *Presence: Teleoperators and Virtual Environments*, 17(2):176–198, 2008.
- [16] H. Iwata. The torus treadmill: Realizing locomotion in ves. *IEEE Computer Graphics and Applications*, 19(6):30–35, 1999.
- [17] H. Iwata, H. Yano, H. Fukushima, and H. Noma. Circulafloor [locomotion interface]. *IEEE Computer Graphics and Applications*, 25(1):64–67, 2005.
- [18] H. Iwata, H. Yano, and H. Tomioka. Powered shoes. In *ACM SIGGRAPH 2006 Emerging technologies*, p. 28. ACM, 2006.
- [19] J. Jerald, T. Peck, F. Steinicke, and M. Whitton. Sensitivity to scene motion for phases of head yaws. In *Proceedings of the 5th symposium on Applied perception in graphics and visualization*, pp. 155–162. ACM, 2008.
- [20] S. Sasahara, S. Nagai, and J. Rekimoto. Jackin head: Immersive visual telepresence system with omnidirectional wearable camera. *IEEE*

- transactions on visualization and computer graphics*, 23(3):1222–1234, 2017.
- [21] N. Katakis, J. Tong, O. Ariza, L. Chen, G. Klinker, B. Röder, and F. Steinicke. Stylo and handfact: Modulating haptic perception through visualizations for posture training in augmented reality. In *Proceedings of the 5th Symposium on Spatial User Interaction, SUI '17*, pp. 58–67. ACM, New York, NY, USA, 2017. doi: 10.1145/3131277.3132181
 - [22] S. Koceski and N. Koceska. Evaluation of an assistive telepresence robot for elderly healthcare. *Journal of medical systems*, 40(5):121, 2016.
 - [23] L. Kohli, E. Burns, D. Miller, and H. Fuchs. Combining passive haptics with redirected walking. In *Proceedings of the 2005 international conference on Augmented tele-existence*, pp. 253–254. ACM, 2005.
 - [24] S. Kratz and F. R. Ferriera. Immersed remotely: Evaluating the use of head mounted devices for remote collaboration in robotic telepresence. In *Robot and Human Interactive Communication (RO-MAN), 2016 25th IEEE International Symposium on*, pp. 638–645. IEEE, 2016.
 - [25] E. Langbehn, P. Lubos, G. Bruder, and F. Steinicke. Bending the curve: Sensitivity to bending of curved paths and application in room-scale vr. *IEEE transactions on visualization and computer graphics*, 23(4):1389–1398, 2017.
 - [26] J. J. LaViola Jr. A discussion of cybersickness in virtual environments. *ACM SIGCHI Bulletin*, 32(1):47–56, 2000.
 - [27] J. M. Loomis and J. M. Knapp. Visual perception of egocentric distance in real and virtual environments. *Virtual and adaptive environments*, 11:21–46, 2003.
 - [28] W.-Y. G. Louie, T. Vaquero, G. Nejat, and J. C. Beck. An autonomous assistive robot for planning, scheduling and facilitating multi-user activities. In *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, pp. 5292–5298. IEEE, 2014.
 - [29] A. Maimone, J. Bidwell, K. Peng, and H. Fuchs. Enhanced personal autostereoscopic telepresence system using commodity depth cameras. *Computers & Graphics*, 36(7):791–807, 2012.
 - [30] A. Maimone and H. Fuchs. Encumbrance-free telepresence system with real-time 3d capture and display using commodity depth cameras. In *Mixed and Augmented Reality (ISMAR), 2011 10th IEEE International Symposium on*, pp. 137–146. IEEE, 2011.
 - [31] 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.
 - [32] A. Maimone and H. Fuchs. Real-time volumetric 3d capture of room-sized scenes for telepresence. In *3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON), 2012*, pp. 1–4. IEEE, 2012.
 - [33] K. Matsumoto, Y. Ban, T. Narumi, Y. Yanase, T. Tanikawa, and M. Hirose. Unlimited corridor: redirected walking techniques using visuo haptic interaction. In *ACM SIGGRAPH 2016 Emerging Technologies*, p. 20. ACM, 2016.
 - [34] C. T. Neth, J. L. Souman, D. Engel, U. Kloos, H. H. Bulthoff, and B. J. Mohler. Velocity-dependent dynamic curvature gain for redirected walking. *IEEE transactions on visualization and computer graphics*, 18(7):1041–1052, 2012.
 - [35] C. Neustaedter, G. Venolia, J. Procyk, and D. Hawkins. To beam or not to beam: A study of remote telepresence attendance at an academic conference. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, pp. 418–431. ACM, 2016.
 - [36] N. Nitzsche, U. Hanebeck, and G. Schmidt. Extending telepresent walking by motion compression. *Tagungsbeiträge*, 1:83–90, 2002.
 - [37] E. Okamura and F. Tanaka. A pilot study about remote teaching by elderly people to children over a two-way telepresence robot system. In *Human-Robot Interaction (HRI), 2016 11th ACM/IEEE International Conference on*, pp. 489–490. IEEE, 2016.
 - [38] A. Paludan, J. Elbaek, M. Mortensen, M. Zobbe, N. C. Nilsson, R. Nordahl, L. Reng, and S. Serafin. Disguising rotational gain for redirected walking in virtual reality: Effect of visual density. In *Virtual Reality (VR), 2016 IEEE*, pp. 259–260. IEEE, 2016.
 - [39] T. C. Peck, H. Fuchs, and M. C. Whitton. Evaluation of reorientation techniques and distractors for walking in large virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, 15(3):383–394, 2009.
 - [40] T. C. Peck, H. Fuchs, and M. C. Whitton. An evaluation of navigational ability comparing redirected free exploration with distractors to walking-in-place and joystick locomotion interfaces. In *Virtual Reality Conference (VR), 2011 IEEE*, pp. 55–62. IEEE, 2011.
 - [41] S. Razzaque. *Redirected walking*. University of North Carolina at Chapel Hill, 2005.
 - [42] S. Razzaque, Z. Kohn, and M. C. Whitton. Redirected walking. In *Proceedings of EUROGRAPHICS*, vol. 9, pp. 105–106. Manchester, UK, 2001.
 - [43] M. Schwaiger, T. Thümmel, and H. Ulbrich. Cyberwalk: Implementation of a ball bearing platform for humans. *Human-Computer Interaction. Interaction Platforms and Techniques*, pp. 926–935, 2007.
 - [44] M. Slater, B. Lotto, M. M. Arnold, and M. V. Sanchez-Vives. How we experience immersive virtual environments: the concept of presence and its measurement. *Anuario de psicología*, 40(2), 2009.
 - [45] D. Smith and N. Louwagie. Delivering advanced technical education using online, immersive classroom technology. *Community College Journal of Research and Practice*, 41(6):359–362, 2017.
 - [46] F. Steinicke, G. Bruder, J. Jerald, H. Frenz, and M. Lappe. Analyses of human sensitivity to redirected walking. In *Proceedings of the 2008 ACM symposium on Virtual reality software and technology*, pp. 149–156. ACM, 2008.
 - [47] F. Steinicke, G. Bruder, J. Jerald, H. Frenz, and M. Lappe. Estimation of detection thresholds for redirected walking techniques. *IEEE transactions on visualization and computer graphics*, 16(1):17–27, 2010.
 - [48] F. Steinicke, G. Bruder, L. Kohli, J. Jerald, and K. Hinrichs. Taxonomy and implementation of redirection techniques for ubiquitous passive haptic feedback. In *Cyberworlds, 2008 International Conference on*, pp. 217–223. IEEE, 2008.
 - [49] F. Steinicke, G. Bruder, T. Ropinski, and K. Hinrichs. Moving towards generally applicable redirected walking. In *Proceedings of the Virtual Reality International Conference (VRIC)*, pp. 15–24. IEEE Press, 2008.
 - [50] F. Steinicke, Y. Visell, J. Campos, and A. Lécuyer. *Human walking in virtual environments*. Springer, 2013.
 - [51] E. A. Suma, G. Bruder, F. Steinicke, D. M. Krum, and M. Bolas. A taxonomy for deploying redirection techniques in immersive virtual environments. In *Virtual Reality Short Papers and Posters (VRW), 2012 IEEE*, pp. 43–46. IEEE, 2012.
 - [52] E. A. Suma, S. Clark, S. L. Finkelstein, and Z. Wartell. Exploiting change blindness to expand walkable space in a virtual environment. In *Virtual Reality Conference (VR), 2010 IEEE*, pp. 305–306. IEEE, 2010.
 - [53] E. A. Suma, S. Clark, D. Krum, S. Finkelstein, M. Bolas, and Z. Warte. Leveraging change blindness for redirection in virtual environments. In *Virtual Reality Conference (VR), 2011 IEEE*, pp. 159–166. IEEE, 2011.
 - [54] S. Tachi. Telexistence: Enabling humans to be virtually ubiquitous. *IEEE computer graphics and applications*, 36(1):8–14, 2016.
 - [55] A. ur Rehman, D.-S. Neculescu, and J. Sasiadek. Robotic based fire detection in smart manufacturing facilities. *IFAC-PapersOnLine*, 48(3):1640–1645, 2015.
 - [56] K. Vasylevska, H. Kaufmann, M. Bolas, and E. A. Suma. Flexible spaces: A virtual step outside of reality. In *Virtual Reality (VR), 2013 IEEE*, pp. 109–110. IEEE, 2013.
 - [57] K. Vasylevska, I. Podkosova, and H. Kaufmann. Walking in virtual reality: Flexible spaces and other techniques. In *The Visual Language of Technique*, pp. 81–97. Springer, 2015.
 - [58] G. Venolia, J. Tang, R. Cervantes, S. Bly, G. Robertson, B. Lee, and K. Inkpen. Embodied social proxy: mediating interpersonal connection in hub-and-satellite teams. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1049–1058. ACM, 2010.