



HAL
open science

Can Retinal Projection Displays Improve Spatial Perception in Augmented Reality?

Etienne Peillard, Yuta Itoh, Jean-Marie Normand, Ferran Argelaguet Sanz, Guillaume Moreau, Anatole Lécuyer

► **To cite this version:**

Etienne Peillard, Yuta Itoh, Jean-Marie Normand, Ferran Argelaguet Sanz, Guillaume Moreau, et al.. Can Retinal Projection Displays Improve Spatial Perception in Augmented Reality?. ISMAR 2020 - 19th IEEE International Symposium on Mixed and Augmented Reality, Nov 2020, Recife, Brazil. pp.124-133, 10.1109/ISMAR50242.2020.00028 . hal-02911740

HAL Id: hal-02911740

<https://hal.science/hal-02911740>

Submitted on 4 Aug 2020

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 - NonCommercial - NoDerivatives 4.0 International License

Can Retinal Projection Displays Improve Spatial Perception in Augmented Reality?

Etienne Peillard^{*†}

Yuta Itoh[‡]
Guillaume Moreau^{*}

Jean-Marie Normand^{*}
Anatole Lécuyer[†]

Ferran Argelaguet[†]

^{*}École Centrale de Nantes, AAU, Inria Hybrid

[†]Inria, University Rennes, IRISA, CNRS

[‡]Augmented Vision Lab, Tokyo Institute of Technology, Japan and RIKEN AIP, Japan

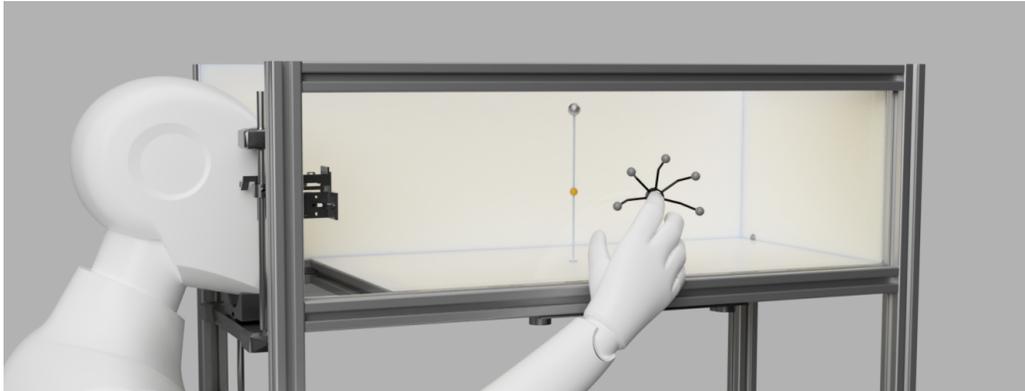


Figure 1: Overview of the experimental setup. The participant had his head on a chinrest and was able to see virtual elements with an OST device or a RPD. He was asked to point a target (virtual or real) with his finger (wearing a tracked ring). The frame was covered by white panels to prevent the participant from seeing his hand. (Here, one panel is transparent for illustration purpose).

ABSTRACT

Commonly used Head Mounted Displays (HMDs) in Augmented Reality (AR), namely Optical See-Through (OST) displays, suffer from a main drawback: their focal lenses can only provide a fixed focal distance. Such a limitation is suspected to be one of the main factors for distance misperception in AR. In this paper, we studied the use of an emerging new kind of AR display to tackle such perception issues: Retinal Projection Displays (RPDs). With RPDs, virtual images have no focal distance and the AR content is always in focus. We conducted the first reported experiment evaluating egocentric distance perception of observers using Retinal Projection Displays. We compared the precision and accuracy of the depth estimation between real and virtual targets, displayed by either OST HMDs or RPDs. Interestingly, our results show that RPDs provide depth estimates in AR closer to real ones compared to OST HMDs. Indeed, the use of an OST device was found to lead to an overestimation of the perceived distance by 16%, whereas the distance overestimation bias dropped to 4% with RPDs. Besides, the task was reported with the same level of difficulty and no difference in precision. As such, our results shed the first light on retinal projection displays' benefits in terms of user's perception in Augmented Reality, suggesting that RPD is a promising technology for AR applications in which an accurate distance perception is required.

Keywords: Perception, Distance, Augmented Reality, User Experiment, Psychophysical Study

^{*}surname.lastname@ec-nantes.fr

[†]surname.lastname@inria.fr

[‡]yuta.itoh@c.titech.ac.jp

Index Terms: Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Virtual reality; Human-centered computing—Human computer interaction (HCI)—HCI design and evaluation methods—User studies

1 INTRODUCTION

Augmented Reality (AR) devices are taking a growing place in both industrial and academic research, being used for many applications, such as design, control, teaching, training, or entertainment. Recent technological advances on tracking and display techniques make such devices practical for far more than prototyping purposes. With sharper displays, bigger field of view and resolution, shorter latency, better tracking and calculation power, modern devices are reaching a maturity in terms of AR technology.

However, AR displays still suffer from important issues highlighted in particular by perceptual evaluations. Several studies concur and show a strong and systematic bias: distance misestimation [24]. Distances perceived by the user of AR device tend to be overestimated at near-field distances (less than 2 meters) and underestimated at far-field distances (more than 30 meters). A similar phenomenon, which causes an underestimation of distances in Virtual Reality (VR), has been the source of many studies [5, 36] and is significantly reduced with modern Head Mounted Displays (HMDs) [4, 22]. However, even with modern AR displays, distance misestimation remains a significant perceptual issue [9, 14]. Recent studies tend to show that defective accommodation cue is one of the main factors of this misestimation [42, 45].

Indeed, most Optical See-Through (OST) headsets use fixed screens and lenses. This leads to a fixed focal distance for the displayed virtual images, which is then incoherent with the other perceptual cues used for distance estimation (binocular disparity and vergence, pictorial cues, etc.). This incoherence between depth cues usually leads to a misestimation of the distance. To tackle this

issue, new AR displays are emerging. Some can change their focal distance to provide a proper accommodation cue [1, 10], and even to provide varifocal occlusion [15]. However, such devices are still prototypes and need further development to be used as “all-in-one,” compact and wearable, AR headsets.

On the other hand, a different kind of usable and consumer-ready AR rendering technology is emerging: Retinal Projection Displays (RPDs). Those displays rely on a laser that projects the virtual images directly onto the user’s retina. While this type of rendering has been studied for a while, recent improvements in the miniaturization and precision of these devices have led to the emergence of the first commercial rendering devices. Since this technology does not rely on screens and lenses to refocus the image, the depth of field is almost infinite, and the image can thus be considered as to be always in focus, no matter where the user is looking at. As a consequence, the observer should not suffer from accommodation issues since; whenever the eyes focus on a specific point, the virtual image is perceived sharp. This technology could significantly improve the perception in AR and even solve distance misestimation in mixed environments.

To evaluate the effect of RPDs on distance perception in AR, we conducted an experiment where participants had to assess the distance of a given object placed at arms reach in front of them. This object can be real or virtual, displayed with an OST display or a RPD. Participants were asked to perform a blind-reaching task to evaluate their distance perception (see Fig. 1). The distance estimation of virtual objects seen with RPDs can then be compared with the distance estimation using OST devices or with real objects.

The remainder of this paper is structured as follows: Sect. 2 describes the related work in the field of distance perception, focusing on egocentric distance perception in AR and the consequences of the Vergence-Accommodation Conflict. Sect. 3 details some technological highlights on RPDs. Sect. 4 presents the user experiment comparing egocentric distance perception in RPDs with OST AR. Our results are presented in Sect. 5 and discussed, along with future research leads in Sect. 6. Finally a conclusion is drawn in Sect. 7.

2 RELATED WORK

2.1 Depth Perception Theory

Distance perception is usually divided into three distance fields [47]: (1) near-field or personal space, from 0 to 2 meters, (2) mid-field or action space, from 2 to 30 meters and (3) far field or vista space, after 30 meters. Since distance perception is closely linked with interaction, those distance fields can also be defined by interaction types feasible in each field. Direct interaction by hand defines personal space, interactions requiring movements does action space, and if no interaction with objects placed is possible, it defines the vista space.

To perceive depths, observers usually use several depth cues, which, when fused, allow to evaluate the distance between the observer and the object. Depth cues fall into two groups [17]: pictorial (atmospheric haze, motion parallax, linear perspective and foreshortening, occlusion, height in the visual field, shading, texture gradient) and non-pictorial ones, which rely on the oculomotor system (binocular disparity, binocular convergence, accommodation cues). Moreover, those cues relative importance are different according to the distance field considered [6]. While binocular disparity and motion perspective are the most important cues in personal space, relative size, and aerial perspective become preeminent in the vista space.

While these cues are usually coherent in real environments, Mixed Reality devices’ limitations imply that some cues are often lacking or are incoherent. As such, depth perception for virtual objects becomes more difficult, and those inconsistencies may bias depth estimation [24].

2.2 Egocentric Distance Perception in AR

To display virtual images, several types of displays are commonly used in AR. In particular, Head Mounted Displays are generally categorized in two groups [2]: (1) Optical See-Through (OST) displays, (2) Video See-Through (VST) displays. OST displays use optical elements to superimpose the view of the real world and the virtual augmentation onto semi-transparent displays. Examples of such displays are the Microsoft HoloLens or the Magic Leap One. In VST displays, the real environment is captured by cameras, and the augmentation is superimposed on the video stream from those cameras. Then, the final image is rendered using similar devices as in VR, such as Windows Mixed Reality HMDs.

Regarding egocentric distance perception in AR in the near distance field, Rolland et al. [37] showed, using a two-alternative forced choice protocol, that the depth of virtual objects is overestimated. Moreover, McCandless et al. [32] proved that within this distance field (between 0.75 m and 1.13 m) and using a matching task, depth judgment errors increase systematically along with distance and latency. Swan et al. [45] compared depth perception for real and virtual targets in the near distance field using two different protocols (matching and “blind reaching”). They found that participants were able to accurately match the distance of a real target while systematically overestimating the distance of virtual targets. They also evaluated depth perception using a “blind reaching” protocol. While the matching protocol, using sliders underneath the table, did not reveal any significant difference between real and virtual targets, the second one, using participants’ fingertips, showed an overestimation of virtual targets distances. Regarding the middle and far distance fields, Swan et al. [44] showed that depth is underestimated until 23 m before switching to overestimated at greater distances.

Livingston et al. [29] reproduced this work and compared the effect of indoor or outdoor environments on depth perception. The indoor environment induced an overall pattern of underestimation while, in the outdoor environment, observers overestimated depth. Still regarding the environment, the height of the perceived target is also known to bias perceived distances, off the ground targets being perceived farther than on-the-ground ones [39, 44].

The rendering of the virtual elements is also an important cue for distance perception [25, 35, 36]. Moreover, in OST HMDs, the representation of the augmentations depends on the background due to the transparency of the virtual display. Livingston et al. [28] evaluated the effect of the contrast of the displayed targets on depth perception and found that low-contrast targets are perceived farther. These findings were also confirmed by a more recent study [42], which showed that brighter targets are perceived closer.

Since they provide an anchor on the floor, shadows are also an important depth cue and, as such, an important rendering issue. Diaz et al. [9] showed that using a matching task in the middle distance field (between 2.5 m and 5.5 m) the use of dropped and cast shadows as well as “billboarding” (i.e., using a 2D object always facing the observer instead of a real 3D object) improved depth perception. Another study [14] showed that virtual objects’ shadows improve distance judgment accuracy but reported no significant differences when using different shadows’ shapes (exact shadow or ellipsoids), even with non-coherent ones.

It has to be mentioned that all these studies were conducted with fixed-focal OST HMDs. However, the dissociation of accommodation and vergence was reported to lead to a biased perception toward the accommodation distance [46]. Singh et al. [42] conducted a study to evaluate the effect of focal distance on depth matching in AR. They found that when the optics were collimated (i.e., displaying the virtual objects at an infinite focal depth), participants overestimated the distances, while they were rather accurate when virtual objects were displayed in the correct focal plane. Interestingly, when the focal plane of the optics was fixed in the middle of the range of tested distances, no significant bias was observed.

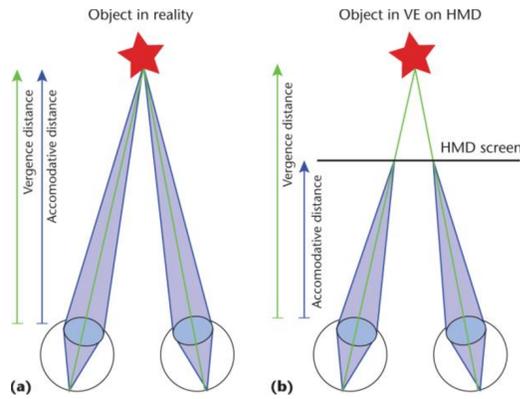


Figure 2: Vergence-Accommodation Conflict: (a) natural vision versus (b) screen based-vision.

To conclude, although no systematic model of depth perception arose from past studies, both the environment as well as the optical limitations of the OST HMDs seem to play an important role in distance perception in AR.

2.3 A Focus on Focal Distance

As detailed in Sect. 2.1, accommodation and vergence are two important cues for depth perception [2]. Under normal circumstances, these cues are coherent: they both provide the same depth information when looking at a specific distant object. The human vision system is familiar with this coherence and acts accordingly. When the eyes converge at a certain distance, they accommodate simultaneously to see the image sharp at this distance [13, 31] (see Fig. 2.a). In AR OST HMDs as in stereoscopic screens, the distance of vergence and that of accommodation are different as shown in Fig. 2.b. This phenomenon is known as the Vergence-Accommodation Conflict (VAC) [16, 21] and has been extensively studied in real, virtual, and augmented reality. It is known to induce visual fatigue and in a more general way to hinder visual performance [16]. However, the VAC may be even more impacting in AR OST devices since an augmented environment contains both real objects, which provide coherent depth cues, and virtual objects, which do not.

To overcome the VAC and to provide a more accurate accommodation cue, some AR devices propose other techniques to display virtual images, which can be categorized as follows:

1. *Varifocal displays*. Images provided by these devices are displayed at a certain depth, but the focal plane is movable. Varifocal displays can be implemented in several ways, each having some limitations, e.g., field of view (FoV), bulkiness, or eye-box size [1, 10, 18, 27].
2. *Multifocal devices* use multiplanar volumetric projection displays based on a stack of laminated planes. This technique can create a more plausible accommodation depth cue and, if dense enough, can provide a continuous and accurate depth cue [19, 38]. This technique is used in the Magic Leap AR headset¹ to provide a better depth sensation. However, this device only provides two distinct focal planes and, to date, there is no perceptual study conducted to assess the effect of this specific feature on depth perception.
3. Other devices try to reproduce the exact light-field, which would have been created by the virtual images with *holographic rendering*. These devices can provide a three-dimensional image with a per-pixel focal control. If the final image is precise enough, the depth cues provided by the virtual

¹<https://www.magicleap.com/>

object are the same as the ones which would have been created by a real object [23, 30].

All these technologies are more accurate than standard fixed-focal OST displays. However, they still suffer from other major drawbacks and limitations, such as a limited resolution, a high space requirement, and high cost. Overall, studies presenting those technologies usually focus on a specific rendering technique and are never compared to each other. Furthermore, these studies focus on technological challenges, yet the techniques they are presenting remain untested regarding their impact on perception.

The next section will focus on presenting a type of display that does not have a fixed focal plane since they do not have any focal plane: Retinal Projection Displays.

3 WHAT ARE RETINAL PROJECTION DISPLAYS?

In 1980, Webb et al. [48] developed the Scanning Laser Ophthalmoscope (SLO) and used it to provide retinal images using a scanning light beam. This prototype of a retinal scanning device was then iteratively improved over the years to include modulation of the light beam to project a whole image in the retina eventually. As a result, the light beam stimulates the retina like an ordinary light ray seen by the eye, allowing the device to act as a virtual display.

Retinal scanning technology relies on the Maxwellian view. This specific configuration of an optical display forces all the light beams from the visual source to converge at the center of the pupil. This point is the center of the optical system of the eye, and the light beams are supposed to be thin enough not to be affected by the eye lens. Thus, the thin light beam directly hits the retina and forms a minimal and sharp point of light.

Enforcing the Maxwellian view on an optical display can be achieved in several ways. The two main ones are: (1) using a Spatial Light Modulator (SLM) to generate a grid of thin parallel beams that converges toward the eye pupil using a lens system and (2) using a Micro-Electro-Mechanical System (MEMS) scanning mirror to reflect a laser beam toward a reflecting free-form mirror which lastly focuses the light in the eye pupil (see Fig. 3). While the first approach is mainly used for rendering only virtual images [7], the second one is more suitable for AR purposes since the free-form mirror can be semi-transparent to make both real and virtual elements visible. Several other Maxwellian-based configurations and more devices layout have also been proposed in the literature, and, for more details, the reader is referred to Lin et al.'s survey [26].

Thanks to their design based on the principle of Maxwellian view, RPDs can render an image that is always sharp, regardless of the focal accommodation of the eye. Besides, recent technical improvements related to MEMS mirrors allow the emergence of more compact devices that could be used for AR headsets. These devices, therefore, have the potential to replace the current OST HMDs if their specificities allow them a better depth perception for users.

4 USER EXPERIMENT

We conducted a user experiment to evaluate the egocentric distance perception in RPDs compared with OST HMDs. Participants had to perform a blind-reaching task; they had to align their finger with a target without seeing their hand. Two groups were involved in this experiment. The first group was wearing a RPD while the other group used an OST HMD. In each group, participants had to perform the blind reaching task with both virtual and real targets in order to compare their egocentric distance perception in AR with real perception. In the following, the apparatus, the protocol, as well as the experimental design are presented.

²<https://www.qdlaser.com/en/applications/eyewear/>

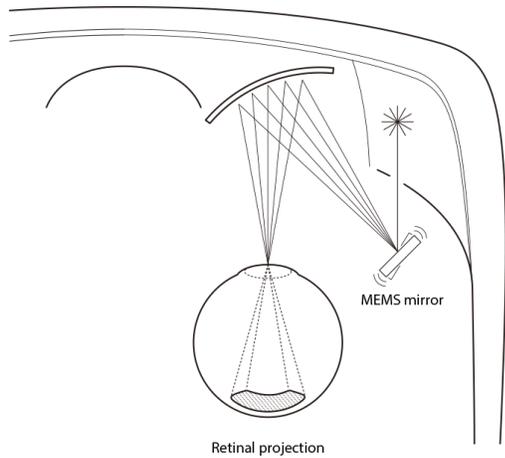


Figure 3: QD-Laser VISIRIUM® technology using a MEMS mirror and a free-form reflecting mirror ².

4.1 Participants

Twenty one right-handed participants took part in the study (18 males, 3 females, mean age: 24.0 (SD: 3.0), range: 21–34 years). The eye dominance was assessed using the method proposed by Durand [11] in which participants were asked to look at a distant object through a hole made with their two hands. Five participants had a dominant left eye and 16 had a dominant right eye. Participants' interpupillary distance (IPD) was measured using the technique detailed in [49] and ranged between 48 and 69 mm (mean: 59.5, SD: 5.76). Participants were students of the university or members of the laboratory and naive about the purpose of the experiment. They were rewarded 1,000 JPY for their participation. All participants had normal or corrected-to-normal vision, 5 (in the OST group) were wearing glasses and 2 (one in the RPD and one in the OST group) were wearing contact lenses. Nineteen out of 21 had already used a VR HMD before and 13 had already used an OST AR HMD.

This study was conducted in accordance with principles as stated in the declaration of Helsinki. The experiment was approved by the ethical committee of the Tokyo Institute of Technology .

4.2 Apparatus

Overall structure and tracking system. A structure was built (30 cm wide, 65 cm long and 18 cm high, see Fig. 4.a) and covered by white cardboard panels to prevent the use of visual cues from the background. A moving part carrying a magnet was placed beneath the "floor" cardboard panel. Another magnet, fixed to the base of the white pole, allowed the experimenter to place the target at a specific distance (see Fig. 5). The absolute position of the white pole was recorded using the reflective ball located at the top of the white target pole. The average reported placement error was 5 mm in depth which provides a sufficient precision for the placement since the following analysis is made using the real position. The participant was wearing a ring carrying a 5-sphere motion tracking constellation to measure the position of his fingertip (see Fig. 4.c). The motion tracking system was a 6-camera OptiTrack system. Under the experimental conditions, the precision of the system once calibrated was always below 0.1 mm. The headset used to display the virtual images (either a RPD or an OST HMD) was fixed on a chin rest. The height of the chair and the height of the chin rest were adjusted at the beginning of the experiment to maintain the same eye level for every participant (i.e. for each participant, the sphere was always visible in the center of their field of view). See Fig. 1 for a global view of the experimental setup.



Figure 4: (a) Overview of the experimental system, here with the RPD. The red cross is used only for the calibration and removed during the experiment itself. (b) The real pole with the orange sphere used as target with the reflective tracking marker on the top. (c) The 3D printed ring used to support 5 reflective tracking markers and provide the position of the fingertip.

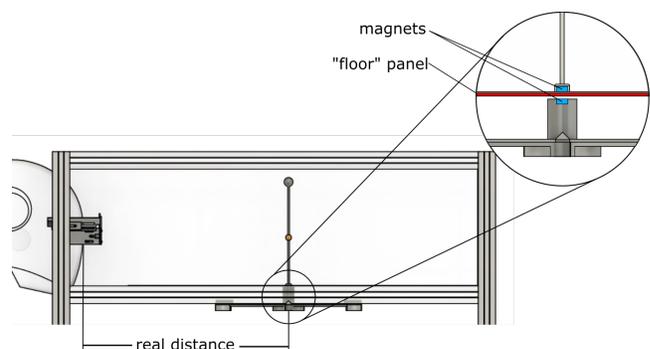


Figure 5: Diagram of the moving element under the structure which allows the experimenter to move the target using two magnets.



Figure 6: Close-up of the custom binocular headset with its two RPD modules mounted on an eye tracking device.

Real and virtual target. The real stimulus was an orange 3D-printed sphere with a 10 mm radius, placed on a 1.5 mm radius white pole. The center of the ball was at 85 mm height above the frame floor, so as to be at eye height (see Fig. 4). The white pole continues above the 3D-printed orange sphere to support an OptiTrack motion tracking ball marker, placed at 180 mm high (see Fig. 4). Given the characteristics of the experimental setup, the base of the pole and the motion tracking marker were out of the participant’s sight. The virtual stimulus was designed to reproduce the appearance of the real stimulus, including the pole. Colors, lightning and shading were carefully chosen to be as close as possible to the real target.

The true distance of the target, reported as *real distance* in the following was, for the real target, the distance measured by the tracking system and, for the virtual target, the distance defined in the virtual scene.

Rendering displays. One group of participants used the Epson Moverio BT-30C AR glasses³. The Moverio headset has a $22.8^\circ \times 12.8^\circ$ FoV with a resolution of 1280×720 pixels per eye. The other group of participants was using a custom mounted binocular RPD (see Fig. 6). This headset consisted in two QD Laser RETISSA Displays, one for each eye. Each display module had a resolution of 1024×600 pixels and an approximate horizontal FoV of 20° . These two modules have been assembled onto Tobii Pro Glasses 2⁴ and the whole unit can be worn directly on the head. To the best of our knowledge, this custom HMD is the first reported prototype of a binocular AR headset using RPD technology and capable of eye tracking. However, like the OST headset, this headset was rigidly mounted on the structure to ensure optimal alignment. In addition, eye tracking data was not collected during this experiment. The positioning of both RPD modules ensured that the stimuli was always in the FoV for both eyes. Moreover, each module has to project the image onto the observer’s retina and was manually positioned and rotated along every axis to ensure the best viewing conditions. This adjustment was made at the beginning of the experiment (during the installation process) for each participant. Note that participants kept wearing the AR device (the OST or the RPD) at all time during the experiment. This means that whenever a real object was showed to them, it was seen through the AR device.

We provided virtual renderings as close as possible for the two devices. While the fields of view were different, the size and distances were chosen in order to make the whole target visible with both devices and the calibration phase ensured that the final rendering was of the same size. Resolutions of the devices were very close so, given the characteristics of the devices, it ranged from 67 to 112px in OST and from 64 to 106px in the RPD. According to previous

³<https://epson.com/moverio-bt30c-usb-c-compatible-smart-glasses>

⁴<https://www.tobii.pro.com/>

studies, the effect of resolution is usually considered minor [8,20] while some significant underestimation effects were found in VR for very low resolution [40]. Since the resolutions of our two conditions are very close, we believe that the effect of resolution remains minor. Since the brightness of the OST device was adjustable, it was manually adjusted prior to the experiment to be as close as possible to that of the RPD. However, the differences of technology between the two devices make it impossible to provide the same brightness, contrast and transparency. Although brightness is recognized as an important aspect of distance perception [41,42], the differences in brightness between the two conditions remained marginal compared to experiences in previous studies.

The AR scene was built using Unity3D v2019.1.8f1. The spheres and their support were displayed using the standard Unity shader, they were lit with a single frontal directional light and without texture. This provided a coherent ambient and specular lightning with no visible shadow (since the light was frontal).

Calibration procedure. The calibration process was divided into several steps. Before the experiment, the FoV of the AR device was assessed during an iterative calibration process on an optical bench. Since both AR displays are always the same for every participant, the FoV of each device remains the same for the whole study. The other parameters of the calibration (rotation and translation) are adjusted for each eye separately by the participant. This was achieved by aligning, in each eye display, a virtual cross with a real calibration cross drawn on the “back” panel of the setup (see Fig. 4).

At the end of this phase, the calibration was checked by visually assessing the alignment between a real and a virtual object placed at the same location. The calibration cross was removed after the calibration procedure and the participant was able to see the target as depicted in Fig. 4.b.

4.3 Task and Procedure

Participants started by reading and filling out a form containing written instructions about the experiment and giving their consent. After some verbal explanations, they carried out the calibration procedure as described above. Then, participants performed one trial using a real target and one trial using a virtual target to get used to the experimental procedure. Finally, participants performed eight blocks of 8 trials. Each successive block used a different type of target. Four blocks were performed with a real target and four with a virtual target. In each group of participants, half of them started with the virtual target and the other half started with the real target. Participants were allowed to have a break after the four first blocks.

Each trial was performed as follows: the target was displayed in front of the participant at a certain depth. If the target was virtual, it was directly displayed by the AR device. If the target was real, the experimenter moved the part under the frame to place the magnet at the chosen distance and then placed the real target over it at the correct location thanks to the magnetic guidance (see Figure 5). The participant was then asked, without seeing his hand, to place his fingertip on the side of the setup, so as to align it with the target (see Fig. 1). They were free to use either hand but all our participants were right-handed and thus used their right hand. The target remained visible during this step. When the participant considered the alignment as correct, s/he reported it to the experimenter and the depth answer was recorded. Finally, the virtual target disappeared or the real target was removed. The experiment was not time-constrained and the participant was able to observe the target for as long as necessary. On average, one trial lasted 8.3 seconds.

The total average time per participant was 45 min, including instructions, calibration, experiment, breaks, post-questionnaire and debriefing. Participants worn the AR displays for ~30 min.

4.4 Experimental Design

We used a mixed-model design with the following conditions:

- **C1**: The distance between the participant’s eye and the target. This distance was pseudo-randomly picked between 30 and 50 cm. For a given block of eight trials, the total range was divided into eight segments and one distance was randomly picked inside each segment.
- **C2**: The type of the object: real or virtual.
- **C3**: The device used: RPD or OST HMD. Each participant performed the experiment with only one device.

The distance between the participant and the target (**C1**) was chosen to be in the near distance field and easily reachable to ensure the precision of the reporting technique. The closest distance was limited by the specificity of the AR devices used. The size of the target and the dimensions of all the surrounding elements were chosen to ensure that the target was always fully visible by both eyes and was the only element which could be seen by the participant during the experiment.

In summary, participants were presented with 64 trials: 8 distances (**C1**) \times 2 target types (**C2**) \times 4 repetitions. For each block, the order of **C1** was randomized and the first type of object (**C2**) was counterbalanced among participants. The participants were divided in two groups according to the AR device they were using (**C3**). The only dependent variable was the reported distance (**D1**). Eventually participants rated the difficulty of the task after achieving all trials. This experimental design is summarized in Table 1.

4.5 Research Hypotheses

The key element of this experimental study is the egocentric distance perception of virtual elements, compared to real element distance perception. Given the literature detailed in Sect. 2, consistent results assess that, while using an OST HMD, egocentric distances of virtual objects are usually overestimated when compared with real objects’ perception [42, 45].

In this experiment, two types of rendering are compared. Their specificities rely on the accommodation cue they are providing. OST HMDs provide a fixed but incorrect accommodation cue which is suspected to be partially responsible for the depth overestimation in AR. On the other hand, RPDs have no focal plane since the image is always sharp and then the accommodation cue is lacking. However, since the eye has to accommodate at a certain distance, the accommodation demand of the eye is expected to be driven by vergence, through the accommodation-vergence reflex. As such, the accuracy of distance estimation of virtual objects is expected to be better with RPDs than with OST HMDs.

However, since this accommodation cue is lacking with RPDs the overall depth perception performance may be reduced. Indeed, the focal plane of OST HMDs is incorrect but fixed, which usually drive the depth overestimation by a fixed amount [45]. However, the lack of focal plane for RPDs could confuse the viewer, leading to a distance more difficult to evaluate. Distance estimation is then expected to be less precise with RPDs, i.e. with a larger spreading of the results.

Considering this forewords, our main research hypotheses were the following:

- **H1**: Egocentric distances are overestimated for virtual targets compared with real targets when seen with an OST HMD.
- **H2**: Egocentric distance estimation in AR is more accurate (i.e. smaller bias) with RPDs than with OST devices.
- **H3**: The precision of egocentric distance estimation is better (i.e. smaller variability) with OST devices compared to RPDs.

4.6 Statistical Analysis

The reported distance **D1** was analysed using a multiple linear regression analysis. The model included the independent variables **C1**, **C2** and **C3** as fixed effect and the *participant* as a random factor.

Table 1: Summary of Independent and Dependent Variables.

Independent Variables			
	<i>observers</i>	21	(random variable)
C1	<i>distance</i>	8	pseudo-random, from 30 to 50 cm
C2	<i>object type</i>	2	virtual or real object (within subject)
C3	<i>device</i>	2	OST HMD or RPD (between subject)
	<i>repetition</i>	4	
Dependent Variables			
D1	<i>reported distance</i>	Projected distance between the eye and the fingertip of the participant along the depth axis	

Since **C1** is continuous, a multiple linear regression approach allows to evaluate the depth estimation as a continuous variable, along another continuous variable (the real distance) and to highlight the differences across **C2** and **C3** which are categorical. Contrary to an ANOVA, a Linear Mixed-Effects (LME) analysis is able to consider **C1** as a continuous variable and then provides slopes and intercepts for each condition. This approach provides more information about the evolution of depth perception along the distances considered [34, 45].

For each participant and condition a linear regression was also computed. This supplementary element provides additional information at an individual level and also allows to evaluate the precision for each participant. Usually, the Sum of Squared Estimate of Errors (SSE) is used to represents the part of the dispersion not explained by the model. However, the SSE is dependent to the number of data points in each group. Here, not all groups have the same number of participants and therefore the same number of data points. So, for virtual and real targets for each participant, the Mean Squared Error (MSE) of the residuals is used instead to evaluate the precision of the participants. This result was analysed using a two-way ANOVA with the *device* (**C3**) considered as a between-subject factor and the *object type* (**C2**) of the target as a within-subject factor. Finally, the results of the Likert-scale estimation of the difficulty of the task was analysed. A Wilcoxon-Mann-Whitney test was used since the results did not pass the requirement for a parametric test.

5 RESULTS

Overall, distances were underestimated by 20.3% on average. This finding is coherent with results reported in the literature for depth estimation with the chosen answering protocol [45].

Effects of object type and device. When analyzing the results, the LME model showed a significant main effect for *object type* of the target (**C2**) on the reported distance ($F_{(1,1314)} = 186.56, p < 0.0001$). On average, distances of virtual objects were overestimated by 9.3%. The overestimation of virtual objects at near-field distances is a well documented effect and one of the motivation of the present study. Here, this effect is in the same magnitude (though close to the upper bound) as those reported in previous studies. As expected, the real distance **C1** had a significant effect on the reported distance ($F_{(1,1314)} = 1064.93, p < 0.0001$). The *device* itself (**C3**) was found not significant.

The interaction effect *device* \times *object type* was found significant ($F_{(1,1314)} = 96.25, p < 0.0001$). This implies that the overestimation of the virtual target position, compared to the real case, is different according to the device used. Finally, the interaction effect *object type* \times *position* was also found significant ($F_{(1,1314)} = 26.05, p < 0.0001$) which shows that the overestimation of distances of virtual objects varies according to the position of the object, i.e. the slope of the linear regression is different if the object is real or virtual. The other interaction effects were non-significant.

To summarize, when comparing virtual and real objects, virtual objects are perceived 5.9 cm farther at 30 cm and 2.7 cm farther at

50 cm when seen through OST devices, while this bias is of only 2.4 cm at 30 cm and -1.2 cm at 50 cm with a RPD device. For each group, the global regressions are presented in Fig. 7A. and the individual regressions in Fig. 7C.

Precision and Consistency. To evaluate the effect of the device on the precision of the observers reported distance, we consider the quadratic means of the residuals of the linear regression (see Sect. 4.6). These results are plotted in Fig. 7B. While the *object type* significantly affects the precision of the results ($F_{(1,19)} = 25.08, p < 0.0001$), the device itself has no effect ($F_{(1,19)} = 0.89, p = 0.45$). The interaction term was also not significant. This result shows that the distance evaluation of virtual objects is less precise than the evaluation of real objects. Since the depth cues provided by the HMDs are not as precise and coherent as the ones provided by the real world, the depth estimation is usually altered, which leads to this decrease in precision. However, the type of the *device* does not affect the precision of the evaluation. This result shows that RPDs' specific way of rendering virtual objects does not significantly affect the precision of the depth estimation. In particular, the lack of accommodation cues does not provide a more confounding depth estimation compared to usual OST renderings.

Moreover, participants were asked to evaluate the difficulty of the task, using a Likert-scale ranging from 1 to 7. There was no significant difference ($W = 39.5, p = 0.28$) between the results of participants who performed the task on the RPD ($M = 4.30, SD = 1.49$) and those who used the OST device ($M = 3.64, SD = 1.36$). Then, while each technology has its specific rendering, no difference of task difficulty is reported by users when performing a depth estimation task.

To summarize, results showed no difference in distance estimation difficulty between RPDs and OST devices, whether in terms of user perception or objective results.

6 DISCUSSION

First, our results show a global underestimation of distances for every condition, when compared with the veridical distance. Our setup is very similar to that of Singh et al. [43] which also reported an underestimation of distance on a reaching task but with a better accuracy. Moreover, this result is coherent with the effort-based theory of distance perception. Indeed, the resting position of the arm is closer to the body of the participants than to the field where the objects were displayed. As such, participants had to extend their arm to point at the object and, since no visual corrective feedback was provided, the effort needed to perform the task may induce the observed underestimation [3, 50]. Interestingly, we did not find any difference between the RPD and OST groups regarding real distance estimation. While the two groups were seeing real objects through specific headsets, with different fields of view, bulkiness and in-built lenses, this did not led to a bias for the real object perception. Their answers remain underestimated by around 24% compared with the veridical distance, whichever was the device worn. This is a strong evidence that the overall underestimation observed in this study is conveyed by the nature of the reaching task and not by the device itself. In addition, the consistency of answers among groups, for real objects, is also an argument in favour of the use of this implementation of the reaching task for highlighting biases caused by virtual objects.

The results of our study show a significant overestimation of virtual estimates in OST devices, compared to real estimates. These results validate our hypothesis **H1**. However, this overestimation bias is significantly reduced for virtual images seen with RPDs. From an average 15% overestimation of virtual objects seen with the OST HMD compared to real estimation, this bias is reduced to only 3% with the RPD. These results are also supported by individual data and confirm our hypothesis **H2**. Previous research underlined the importance of the accommodation cue for distance estimation in

OST displays [42]. In our study, the focal distance of the Moverio BT-30C was 2.5 m. Since the evaluated distances were under 50 cm, the overestimation of virtual objects with the OST device may be explained by its bigger accommodation distance. It is still unclear if this effect is a direct consequence of the VAC or a specific bias which allows the accommodation cue to influence distance estimation.

However, on the other hand, we did not observe this overestimation bias with RPDs. Since these devices do not provide any accommodation cue, the usual bias inherent in OST devices seems to be drastically reduced. When the observers were performing the estimation task, two scenarios can occur: (1) the accommodation cue can be perceived as missing or (2) the accommodation is still driven by the vergence, and as such, perceived as always correct. While we cannot assess strongly for one option or the other given our results, participants never reported any discomfort regarding accommodation while wearing the RPD, during the experiment or as a comment in the following questionnaire. Thus, there is no obvious discomfort with accommodation when participants use RPDs, but it is unclear whether this is the main reason why they performed better for the estimation task.

Finally, the results of the precision and consistency of the answers, as well as the perceived difficulty of the task shed an interesting light on these results. Indeed, results showed no difference in distance estimation difficulty between RPDs and OST devices, which rejects our hypothesis **H3**. Thus, the VAC should not occur when using a RPD since the image is always sharp and then does not conflict with vergence. But, the VAC is known to induce fatigue, discomfort and overall making the distance estimation task more difficult [16]. So, if the VAC was responsible for the overestimation reported for OST devices, we should report a decrease in precision, with more spreading of the data, for the OST group (compared to the RPD), which we did not. Those results suggest that the overestimation reported for OST devices is more likely to rely on a specific bias induced by accommodation, rather than being an effect of the Vergence-Accommodation Conflict.

On the other hand, the lack of an accommodation cue with RPDs was expected to make the task more difficult for the observers. However, it turns out that participants not only gave more accurate answers, but also as precise ones. These results suggest that the absence of the accommodation depth cue is not specifically impacting in this case and that to have no accommodation cue at all is better than an incorrect accommodation cue.

6.1 Limitations and Future Work

As in many distance perception studies in AR, the answering protocol might be partly responsible for the overall underestimation of the perceived distances. Reaching tasks are generally used to report perceived distances as they are not biased by any visual feedback but suffer from proprioception-based biases since no correction is provided. Other tasks without visual feedback such as verbal report or indirect pointing could be used to confirm the particularity of distance perception with RPDs. However, in our case, absolute perception is not important as the bias is also affecting real answers. Here, the key point is the difference in perception between real and virtual objects.

On the other hand, matching tasks could also be considered in a similar setup. Those tasks are different from the task used in our study since they provide visual feedback during distance estimation. If the participant has to align the designated target with another object, say, a cursor, the nature of this cursor is fundamental. Indeed, if the cursor has the same type (real or virtual) as the target, then the potential bias is the same and the alignment should be accurate. However, when the target and the cursor are of a different type, the depth cues provided by the two kinds of rendering would interfere, leading to potential unusual biases. Moreover, in the specific case of OST devices vs. RPDs, we already underlined the major influence

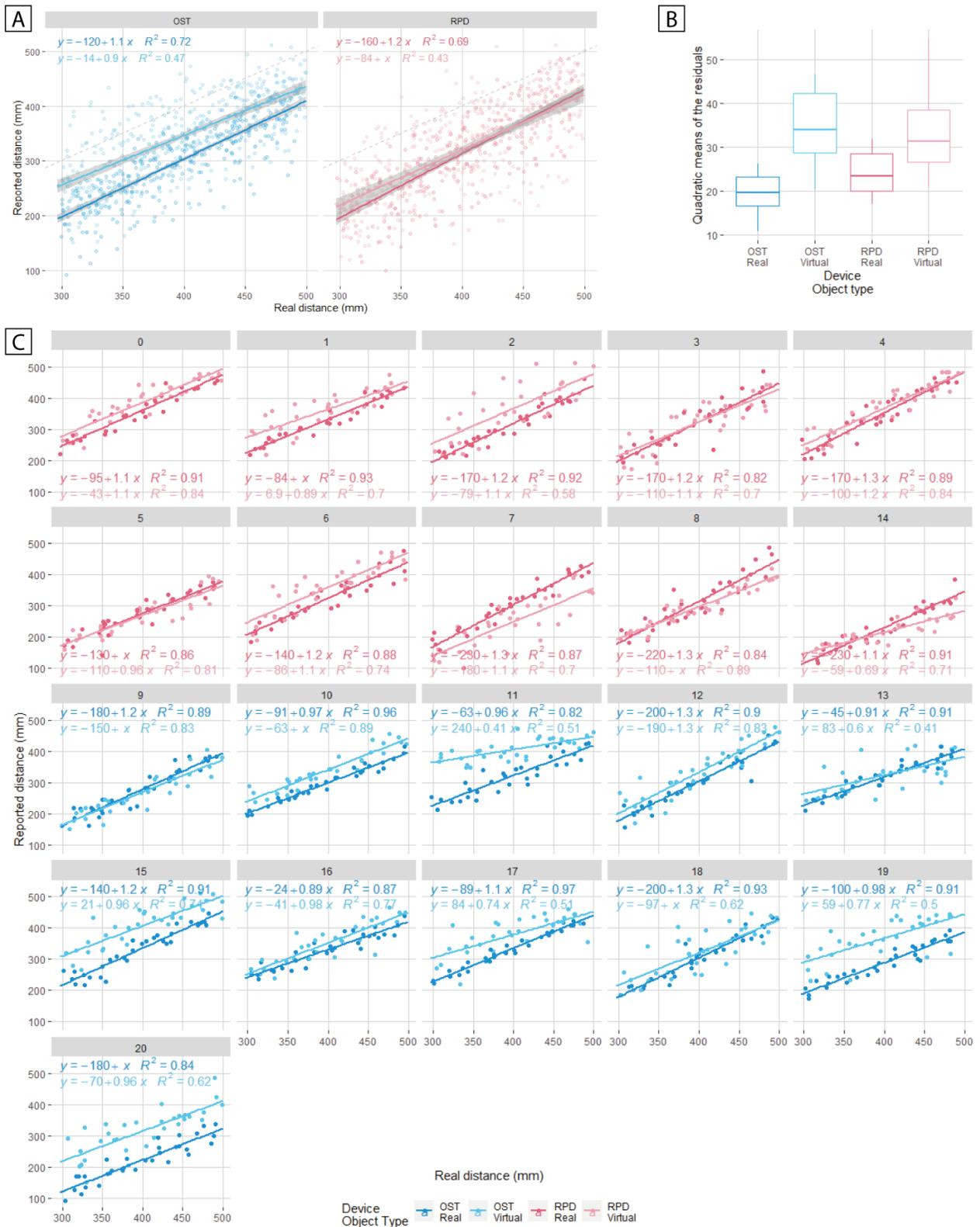


Figure 7: A) Reported distance vs. real distance for each object type and device, with linear regression for each group. B) Quadratic means of the residuals of the linear regression for each object type for each participant, aggregated by object type and device used. C) Reported distance vs. real distance for each participant, with linear regression for both object type. The *reported distance* is the projected distance between the eye and the fingertip of the participant along the depth axis and the *real distance* is the distance between the eye of the participants and the stimulus (measured by the tracking system if the target is real or placed by the rendering device if virtual). The dashed line represents the 1:1 mapping, corresponding to the veridical performance. Each dot represents one trial.

of the accommodation cue. Here, the different nature of the target and the cursor might be even more impacting for distance perception since accommodation is a very strong cue for ordering [33].

We have just discussed the implications of adding a cursor in the scene. However, it would be interesting to evaluate the effect of several objects added to the virtual scene on the observers' depth perception. Then, rather than a cursor, one or several supplementary targets could be added to the experiment. These new targets could be organized in several layouts: two side-by-side or one behind the other, or even in a regular grid or some other random pattern. Those layouts would allow evaluating the potential impact of references, occlusion, ordering or relative distances. While the literature provides an extensive corpus around egocentric distance perception in AR (see Sect. 2), the specificities of multi-targets distance evaluation and their interaction still require more evaluation. The specific case of accommodation is especially interesting for RPDs. Indeed, when comparing two objects' depths, occlusions are a very strong cue. When using OST devices, this cue is not available and the VAC would prevent the observer to adjust the accommodation in a natural way. With RPDs, the increment of vergence from one object to another could be associated with an increment in depth of the accommodation without inducing a conflict. Retinal projection could then be an even more accurate technology for depth perception when it comes to inter-objects estimations.

Other characteristics of the experiment could be modified to further explore the use of RPDs in visual perception. First, the near distance field was chosen in this experiment to highlight the specificities of accommodation in RPDs but other distance fields could be considered. It is possible that for greater distances the effect is less important or even disappears because the accommodation effect decreases with distance. Second, since the evaluation of distance in real or mixed environments has been reported to be dependent on the shape, size and rendering of the virtual object [12, 35, 36], these parameters should also be evaluated for RPDs. Finally, the effect of other usual anchoring cues such as projected shadows should also be evaluated for this kind of displays.

In our setup, every part of the installation was fixed, from the target to the head of the participant itself (supported by the chin-rest). However, this configuration does not represent the common use case of AR environments. Usually, the observer is free to move his/her head and small movements of the head can help evaluating distances. The observed objects can also be moving, thus providing more depth cues such as motion parallax. Then, a richer and more ecological AR environment would provide a specific blending of many more cues which could interact with each others and the evaluation of the specific interaction of retinal projection in such an environment is a very exciting open question.

One last specificity of RPDs to mention is their use for people suffering from eye disorders. Since RPDs project directly on the retina with a laser beam, the resulting image is always sharp, even for people suffering from eye accommodation disabilities such as myopia or hypermetropia. Then, virtual images rendered with RPDs are even sharper than real images. As such, while improving the quality of the image seen, it is still unclear if retinal projected rendering would significantly enhance distance perception for visually impaired people or if this kind of rendering, drastically different from their usual viewing, will disturb their perception of distance.

7 CONCLUSION

In this paper, we evaluated the performance in depth estimation of Retinal Projection Displays for AR when compared with Optical See-Through displays on a perceptual basis. During the experiment, participants had to designate the perceived position of virtual or real objects, using either an OST device or a RPD.

We found that participants' distance estimation of virtual objects were closer to that of real objects when using RPDs than with OST

devices. While OST devices created an overestimation bias for virtual objects in egocentric distance estimation of 15%, the bias induced by RPDs was only of 3%. Besides, participants' precision was also not affected by the technology of the device. Moreover, the depth estimation task was not perceived as more difficult when performed with the RPDs than when performed with an OST HMD. As a consequence, RPDs appear to be a good alternative to usual OST HMDs, in particular whenever a good precision in distance perception is required. Indeed, they provide a significantly better accuracy for a depth estimation task without any significant drawback in precision or perceived difficulty.

As such, our results shed the first light on Retinal Projection Displays' benefits in terms of user's perception in Augmented Reality, suggesting that RPD is a promising technology for AR applications in which an accurate distance perception is required.

ACKNOWLEDGMENTS

This work was supported by JST PRESTO Grant Number JP-MJPR17J2 and JSPS KAKENHI Grant Number JP17H04692, Japan.

REFERENCES

- [1] K. Akşit, W. Lopes, J. Kim, P. Shirley, and D. Luebke. Near-eye varifocal augmented reality display using see-through screens. In *ACM Transactions on Graphics*, vol. 36, pp. 1–13. Association for Computing Machinery, 11 2017. doi: 10.1145/3130800.3130892
- [2] R. Azuma. A Survey of Augmented Reality, 1997. doi: 10.1162/pres.1997.6.4.355
- [3] G. Berkeley. *An essay toward a new theory of vision*. Century psychology series. 1709. doi: 10.1037/11304-009
- [4] L. E. Buck, M. K. Young, and B. Bodenheimer. A Comparison of Distance Estimation in HMD-Based Virtual Environments with Different HMD-Based Conditions. *ACM Transactions on Applied Perception*, 15(3):1–15, 7 2018. doi: 10.1145/3196885
- [5] S. H. Creem-Regehr, J. K. Stefanucci, W. B. Thompson, N. Nash, and M. McCardell. Egocentric distance perception in the Oculus Rift (DK2). In *SAP 2015: ACM SIGGRAPH Symposium on Applied Perception*, pp. 47–50. Association for Computing Machinery, Inc, 2015. doi: 10.1145/2804408.2804422
- [6] J. E. Cutting. How the eye measures reality and virtual reality. *Behavior Research Methods, Instruments, and Computers*, 29(1):27–36, 1997. doi: 10.3758/BF03200563
- [7] D. S. Dewald, A. T. Evans, N. Welch, A. Gross, and G. Hill. 8-1: Invited Paper : The Avegant Glyph: Optical Design Considerations and Approach to Near-eye Display. *SID Symposium Digest of Technical Papers*, 47(1):69–71, 5 2016. doi: 10.1002/sdtp.10609
- [8] A. Dey, G. Jarvis, C. Sandor, and G. Reitmayr. Tablet versus phone: Depth perception in handheld augmented reality. In *ISMAR 2012 - 11th IEEE International Symposium on Mixed and Augmented Reality 2012, Science and Technology Papers*, pp. 187–196. IEEE, 11 2012. doi: 10.1109/ISMAR.2012.6402556
- [9] C. Diaz, M. Walker, D. A. Szafrin, and D. Szafrin. Designing for Depth Perceptions in Augmented Reality. In *2017 IEEE International Symposium on Mixed and Augmented Reality*, 2017. doi: 10.1109/ISMAR.2017.28
- [10] D. Dunn, C. Tippetts, K. Torell, P. Kellnhofer, K. Akşit, P. Didyk, K. Myszkowski, D. Luebke, and H. Fuchs. Wide Field of View Varifocal Near-Eye Display Using See-Through Deformable Membrane Mirrors. *IEEE Transactions on Visualization and Computer Graphics*, 23(4):1275–1284, 4 2017. doi: 10.1109/TVCG.2017.2657058
- [11] A. C. Durand and G. M. Gould. A method of determining ocular dominance. *Journal of the American Medical Association*, 55(5):369–370, 7 1910. doi: 10.1001/jama.1910.04330050007004
- [12] R. G. Eggleston, W. P. Janson, and K. A. Aldrich. Virtual reality system effects on size-distance judgments in a virtual environment. In *Proceedings - Virtual Reality Annual International Symposium*, pp. 139–146. IEEE, 1996. doi: 10.1109/vrais.1996.490521

- [13] E. F. Fincham and J. Walton. The reciprocal actions of accommodation and convergence. *The Journal of physiology*, 137(3):488–508, 1957. doi: 10.1113/jphysiol.1957.sp005829
- [14] Y. Gao, E. Peillard, J. M. Normand, G. Moreau, Y. Liu, and Y. Wang. Influence of virtual objects' shadows and lighting coherence on distance perception in optical see-through augmented reality. *Journal of the Society for Information Display*, 28(2):117–135, 2 2020. doi: 10.1002/jsid.832
- [15] T. Hamasaki and Y. Itoh. Varifocal Occlusion for Optical See-Through Head-Mounted Displays using a Slide Occlusion Mask. *IEEE Transactions on Visualization and Computer Graphics*, 25(5):1961–1969, 5 2019. doi: 10.1109/TVCG.2019.2899249
- [16] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks. Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision*, 8(3):33, 3 2008. doi: 10.1167/8.3.33
- [17] I. P. Howard. *Seeing in depth, Vol. 1: Basic mechanisms*. University of Toronto Press, Toronto, ON, Canada, 2002.
- [18] X. Hu and H. Hua. High-resolution optical see-through multi-focal-plane head-mounted display using freeform optics. *Optics Express*, 22(11):13896, 6 2014. doi: 10.1364/oe.22.013896
- [19] H. Hua and B. Javidi. A 3D integral imaging optical see-through head-mounted display. *Optics Express*, 22(11):13484, 6 2014. doi: 10.1364/oe.22.013484
- [20] K.-M. Jaeae-Aro and L. Kjeldahl. Effects of image resolution on depth perception in stereo and nonstereo images. In S. S. Fisher, J. O. Merritt, and M. T. Bolas, eds., *Stereoscopic Displays and Virtual Reality Systems IV*, vol. 3012, p. 319. SPIE, 5 1997. doi: 10.1117/12.274474
- [21] B. Julesz. *Foundations of cyclopean perception*. U. Chicago Press, Oxford, England, 1971.
- [22] J. W. Kelly, L. A. Cherep, and Z. D. Siegel. Perceived Space in the HTC Vive. *ACM Transactions on Applied Perception*, 15(1):1–16, 7 2017. doi: 10.1145/3106155
- [23] H. J. Kim, S. K. Lee, M. L. Piao, N. Kim, and J. H. Park. Three-dimensional holographic head mounted display using holographic optical element. In *2015 IEEE International Conference on Consumer Electronics (ICCE)*, pp. 132–133. Institute of Electrical and Electronics Engineers Inc., 3 2015. doi: 10.1109/ICCE.2015.7066351
- [24] E. Kruijff, J. E. Swan, and S. Feiner. Perceptual issues in augmented reality revisited. In *2010 IEEE International Symposium on Mixed and Augmented Reality*, pp. 3–12. IEEE, 10 2010. doi: 10.1109/ISMAR.2010.5643530
- [25] S. A. Kuhl, W. B. Thompson, and S. H. Creem-Regehr. Minification influences spatial judgments in virtual environments. *APGV '06: Proceedings of the 3rd symposium on Applied perception in graphics and visualization*, 1(212):15, 2006. doi: 10.1145/1140491.1140494
- [26] J. Lin, D. Cheng, C. Yao, and Y. Wang. Retinal projection head-mounted display. *Frontiers of Optoelectronics*, 10(1):1–8, 3 2017. doi: 10.1007/s12200-016-0662-8
- [27] S. Liu, D. Cheng, and H. Hua. An optical see-through head mounted display with addressable focal planes. In *2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, pp. 33–42, 2008. doi: 10.1109/ISMAR.2008.4637321
- [28] M. A. Livingston, Z. Ai, and J. W. Decker. A user study towards understanding stereo perception in head-worn augmented reality displays. *2009 8th IEEE International Symposium on Mixed and Augmented Reality*, pp. 53–56, 2009. doi: 10.1109/ISMAR.2009.5336496
- [29] M. A. Livingston, Z. Ai, J. E. Swan, and H. S. Smallman. Indoor vs. Outdoor Depth Perception for Mobile Augmented Reality. In *2009 IEEE Virtual Reality Conference*, pp. 55–62, 2009. doi: 10.1109/VR.2009.4810999
- [30] A. Maimone, A. Georgiou, and J. S. Kollin. Holographic near-eye displays for virtual and augmented reality. In *ACM Transactions on Graphics*, vol. 36, pp. 1–16. Association for Computing Machinery, 7 2017. doi: 10.1145/3072959.3073624
- [31] T. G. Martens and K. N. Ogle. Observations on accommodative convergence; especially its nonlinear relationships. *American journal of ophthalmology*, 47(1 Part 2):455–62, 1 1959.
- [32] J. W. McCandless, S. R. Ellis, and B. D. Adelstein. Localization of a Time-Delayed, Monocular Virtual Object Superimposed on a Real Environment. *Presence: Teleoperators and Virtual Environments*, 9(1):15–24, 2 2000. doi: 10.1162/105474600566583
- [33] M. Mon-Williams and J. R. Tresilian. Ordinal depth information from accommodation? *Ergonomics*, 43(3):391–404, 2000. doi: 10.1080/001401300184486
- [34] E. J. Pedhazur. *Multiple Regression in Behavioral Research.*, vol. 70. 1975. doi: 10.2307/2285468
- [35] L. Phillips, B. Ries, V. Interrante, M. Kaeding, and L. Anderson. Distance perception in NPR immersive virtual environments, revisited. *APGV '09: Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization*, p. 11, 2009. doi: 10.1145/1620993.1620996
- [36] R. S. Renner, B. M. Velichkovsky, and J. R. Helmer. The perception of egocentric distances in virtual environments - A review. *ACM Computing Surveys*, 46(2):1–40, 11 2013. doi: 10.1145/2543581.2543590
- [37] J. P. Rolland, W. Gibson, and D. Arieli. Towards Quantifying Depth and Size Perception in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, 4(1):24–49, 1 1995. doi: 10.1162/pres.1995.4.1.24
- [38] J. P. Rolland, M. W. Krueger, and A. Goon. Multifocal planes head-mounted displays. *Applied Optics*, 39(19):3209, 7 2000. doi: 10.1364/ao.39.003209
- [39] C. S. Rosales, G. Pointon, H. Adams, J. Stefanucci, S. Creem-Regehr, W. B. Thompson, and B. Bodenheimer. Distance Judgments to On- and Off-Ground Objects in Augmented Reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 237–243, 2019. doi: 10.1109/VR.2019.8798095
- [40] J. Ryu, N. Hashimoto, and M. Sato. Influence of resolution degradation on distance estimation in virtual space displaying static and dynamic image. In *Proceedings - 2005 International Conference on Cyberworlds, CW 2005*, vol. 2005, pp. 43–50, 2005. doi: 10.1109/CW.2005.54
- [41] S. Schmidt, G. Bruder, F. S. ICAT-EGVE, and U. 2017. Moving Towards Consistent Depth Perception in Stereoscopic Projection-based Augmented Reality. *Icat-Egve*, pp. 161–168, 2017.
- [42] G. Singh, S. R. Ellis, and J. E. Swan. The Effect of Focal Distance, Age, and Brightness on Near-Field Augmented Reality Depth Matching. *IEEE Transactions on Visualization and Computer Graphics*, 26(2):1385–139, 2018. doi: 10.1109/TVCG.2018.2869729
- [43] G. Singh, J. E. Swan, A. Jones, and S. R. Ellis. Depth judgments by reaching and matching in near-field augmented reality. *2012 IEEE Virtual Reality Workshops (VRW)*, pp. 165–166, 2012. doi: 10.1109/VR.2012.6180933
- [44] J. E. Swan, A. Jones, E. Kolstad, M. A. Livingston, and H. S. Smallman. Egocentric depth judgments in optical, see-through augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 13(3):429–442, 2007. doi: 10.1109/TVCG.2007.1035
- [45] J. E. Swan, G. Singh, and S. R. Ellis. Matching and Reaching Depth Judgments with Real and Augmented Reality Targets. *IEEE Transactions on Visualization and Computer Graphics*, 21(11):1289–1298, 2015. doi: 10.1109/TVCG.2015.2459895
- [46] H. A. Swenson. The relative influence of accommodation and convergence in the judgment of distance. *Journal of General Psychology*, 7(2):360–380, 10 1932. doi: 10.1080/00221309.1932.9918473
- [47] J. Vishton and P. Cutting. Perceiving Layout and Knowing Distances: The Integration, Relative Potency, and Contextual Use of Different Information about Depth. *Perception of Space and Motion*, 22(5):69–117, 1995. doi: 10.1016/B978-012240530-3/50005-5
- [48] R. H. Webb, G. W. Hughes, and O. Pomerantzeff. Flying spot TV ophthalmoscope. *Applied Optics*, 19(17):2991, 9 1980. doi: 10.1364/ao.19.002991
- [49] P. Willemsen, A. A. Gooch, W. B. Thompson, and S. H. Creem-Regehr. Effects of stereo viewing conditions on distance perception in virtual environments. *Presence: Teleoperators and Virtual Environments*, 17(1):91–101, 2008. doi: 10.1162/pres.17.1.91
- [50] J. K. Witt, D. R. Proffitt, and W. Epstein. Perceiving distance: a role of effort and intent. *Perception*, 33(5):577–90, 5 2004. doi: 10.1068/p5090