HAL
open science

Influence of virtual objects' shadows and lighting coherence on distance perception in optical see-through augmented reality

Yuan Gao, Etienne Peillard, Jean-Marie Normand, Guillaume Moreau, Yue Liu, Yongtian Wang

To cite this version:

Yuan Gao, Etienne Peillard, Jean-Marie Normand, Guillaume Moreau, Yue Liu, et al.. Influence of virtual objects' shadows and lighting coherence on distance perception in optical see-through augmented reality. Journal of the Society for Information Display, 2020, 28 (2), pp.117-135. 10.1002/jsid.832 . hal-02200576

HAL Id: hal-02200576

https://hal.science/hal-02200576

Submitted on 16 Dec 2019

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.

Influence of Virtual Objects' Shadows and Lighting Coherence on Distance Perception in Optical See-Through Augmented Reality

Yuan Gao[1,2], Etienne Peillard[2], Jean-Marie Normand[2], Guillaume Moreau[2], Yue Liu[1,3], Yongtian Wang[1,3]

Correspondence
Yue Liu, Beijing Engineering Research Center of Mixed Reality and Advanced Display, School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China; AICFVE of Beijing Film Academy, 4, Xitucheng Rd, Haidian, Beijing, China.
Email: liuyue@bit.edu.cn
Yuan Gao, Beijing Engineering Research Center of Mixed Reality and Advanced Display, School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China; Ecole Centrale de Nantes, AAU UMR CNRS1563.
Email: gaoyuan1988.ok@163.com

Abstract

This paper focuses on how virtual objects' shadows as well as differences in alignment between virtual and real lighting influence distance perception in Optical See-Through (OST) Augmented Reality (AR). Four hypotheses are proposed: (H1) Participants underestimate distances in OST AR; (H2) Virtual objects' shadows improve distance judgment accuracy in OST AR; (H3) Shadows with different realism levels have different influence on distance perception in OST AR; (H4) Different levels of lighting misalignment between real and virtual lights have different influence on distance perception in OST AR scenes. Two experiments were designed with an OST Head Mounted Display (HMD), the Microsoft HoloLens. Participants had to match the position of a virtual object displayed in the OST-HMD with a real target. Distance judgment accuracy were recorded under the different shadows and lighting conditions. The results validate hypotheses H2 and H4 but surprisingly showed no impact of the shape of virtual shadows on distance judgment accuracy thus rejecting hypothesis H3. Regarding hypothesis H1, we detected a trend towards underestimation, given the high variance of the data, more experiments are needed to confirm this result. In addition, the study also reveals that errors on perceived distances and completion time

[1] Beijing Engineering Research Center of Mixed Reality and Advanced Display, School of Optics and Photonics, Beijing Institute of Technology, Beijing, China
[2] Ecole Centrale de Nantes, AAU UMR CNRS1563-INRIA Hybrid
[3] AICFVE of Beijing Film Academy, Beijing, China

of trials increase along with targets' distance.

Keywords: Augmented Reality, Optical See-through HMD, Distance perception, Shadows, Lighting coherence

1. Introduction

Augmented Reality (AR) is an important and rapidly growing technology which has been widely used in many areas such as education, medical science, industry, and military. The three main types of AR displays are: Video See-Through (VST), Optical See-Through (OST) and Spatial Augmented Reality (SAR) [1]. OST Head Mounted Displays (HMDs) such as the Microsoft HoloLens[4] or the Magic Leap One[5] are becoming more popular thanks to recent developments in optical as well as tracking technologies. These devices allow for a natural AR integration of real and virtual objects since they do not require users to see the world through a camera but directly with their own eyes. AR HMDs also offer a natural way of interaction by leaving users use their hands freely, unlike phones or tablets. While the most important issue related to AR concerns the registration of virtual objects in the real scene, some questions remain regarding how users perceive those virtual objects in comparison with real ones. This is a crucial issue regarding perception and interaction in AR.

In this paper, we study how virtual objects' shadows as well as differences in alignment between virtual and real lighting can influence distance estimation in OST AR. In particular, we conduct two experiments in an OST-HMD (the Microsoft HoloLens) where participants had to match the position of a virtual object displayed in OST AR with real targets. The first experiment aimed at testing the influence of virtual shadows while the second experiment's goal is to study whether a misalignment between real and virtual lights could impact distance estimation in the

4 https://www.microsoft.com/en-us/hololens
5 https://www.magicleap.com/magic-leap-one

same OST-HMD AR environment.

The remainder of this paper is organized as follows: Sect. 2 presents the related work on distance estimation both in Virtual Reality (VR) and AR with a focus on the influence of shadows and lighting, Sect. 3 defines the problem as well as our research hypotheses. Our experimental design, common to both experiments is presented in Sect. 4. Experiments on the influence of virtual objects' shadows and lighting placement are detailed respectively in Sect. 5 and Sect. 6. Finally our results are discussed in Sect.7 before drawing a conclusion in Sect. 8.

2. Related work

Over the past decades, depth perception has been widely studied in VR [2, 3, 4, 5], and while some research has been carried out in AR (e.g. [6]), this issue was, comparatively to VR, less investigated.

Underestimation of egocentric distances in VR has been a puzzling problem for the community for several years [7, 8]. However, with the development of recent commodity level VR hardware, in particular new HMDs, this issues has been reinvestigated and previous findings about distance underestimation do not seem to be as strong as they used to. Indeed, Kelly et. al [9] found that while distance underestimation in the HTC Vive is lower than when using older displays, it still remains. More recently, Buck et.al [10], examined eight different VR HMDs to evaluate how much users underestimate distance. Their results show that new commodity-level HMDs provide more accurate estimates of distance than prior generations of HMDs, but this question remains open since their findings are slightly different from Kelly et al. [9]'s ones regarding the particular case of the HTC Vive.

Some studies focus on the factors which influence distance perception in VR. Li et al. [11] found that field of view (FOV) plays an important role in distance perception in VR, their

experiment result shows that a smaller FOV caused a more serious underestimation of distance judgment in VR. Jones et al. [12] found that applying constant white light in an observer's far periphery lead to more accurate distance judgments using blind walking in a Virtual Environment (VE). In addition, it was also shown the users overestimate distances in the real world after they interacted with a VE for 18 trials [13]. This result was confirmed by Jones et al. [14]'s research. Their experiment reveals that participants tend to walk roughly 5% farther when performing blind walking in the real world after accomplishing the tasks in VR. The phenomenon was explained as the mismatched physical and visual movement in a synthetic environment can cause errors on returning to a natural environment.

As for research on AR, results of different studies are not uniform. Rolland et al. [7] found that the depth of virtual objects lying within 0.8-1.2m was overestimated in AR. While, in their follow-up experiment [15], results show that there is no consistent bias in depth judgment. However, Swan et al. [16] found an interesting phenomenon: there is a switch in users' bias from underestimation to overestimation in AR OST-HMDs when distance to virtual objects is ≥ 23m. Livingston et al. [17] found that there is an overall pattern of underestimation in indoor environments, whereas participants overestimated depth in outdoor environments through OST-HMDs in AR. Jones et al. [18] conducted an experiment to measure egocentric depth perception in VR, AR as well as in reality, and they found that participants underestimate depth in the VR condition while no underestimation was found in the AR condition.

In order to improve depth accuracy in VR and AR, numerous experiments were conducted to identify factors causing depth perception errors. It was found that one of the most fundamental factors is the breakdown between convergence and accommodation in VR and AR display systems [19]. When users view virtual objects through a binocular AR display, their eyes accommodate on the 2D screen where the augmented information is in focus, but at the same time, the eyes need to

converge on the rendered depth of the real object to fuse the stereoscopic pair [20]. When viewing real objects, accommodation distance and convergence depth are the same. Swan et al. [21] found that observers can accurately match the distance of a real target, but when viewing an AR target, they overestimated the distance by 0.5cm to 4.0cm. An explanation for this phenomenon is that collimating optics cause eyes' vergence angle to rotate outward by a constant angular amount. Singh et al.'s experiment [22] confirmed these results and showed that the bias of vergence angle bias is constant within distances of 33.3cm to 50cm when the subjects match virtual targets in AR. They also found that presenting AR objects at the AR display's focal distance results in more accurate depth matches. Obviously, it is hard to eliminate the influence of this accommodation/convergence issue since it is caused by most AR displays hardware.

Some researchers also focused on studying factors related to the display of virtual objects, which are easier to control, compared with the hardware of AR display systems. For example, Singh et al. [23] found that the presence of a highly salient occluding surface has a complicated effect that disrupts the linear relation between depth judging distance and distance errors within near field distance in OST-HMDs. Swan et al. [16] designed an experiment to study the effects of upper versus lower visual field locations, the number of repetitions of the task as well as X-Ray vision (i.e. displaying virtual objects in AR environment when located behind opaque surfaces) in OST-HMDs. Results show that distance was overestimated by 10 percent in the upper visual field condition compared with the lower visual field condition within the first three repetitions. An equation was established to show that observers had an additional 8cm-error in the non-X-Ray condition compared with the X-Ray condition. They also found that the accuracy of distance judgment increases along with the number of repetitions. Messing and Durgin [24] found that lowering the horizon line in VR HMDs lead to an overestimation in distance judgments. Also known to all, the size of objects is a depth cue in that distant objects looks smaller than close

objects. Some papers studied the depth cue of relative size of objects. Diaz et al. [6] found a significant linear relationship between distance error and objects' size which reveals that participants' depth perception improved as virtual objects got bigger in AR.

In addition to the above factors, shadow is another important depth cue for the human's visual system. Mamassian et al. [25] found that cast shadows are helpful to the recovery of spatial arrangement, especially when the shadow is in motion. Hubona et al. [26] designed an experiment to verify the effectiveness of computer rendered shadows, and results show that the use of object shadows enhances the accuracy of object positioning in a VR-like environment. As for shadows in AR, they allow for a connection between virtual objects and the real world as well as help virtual objects look more realistic. Diaz et al. [6] conducted an experiment to prove that the presence of shadows can improve the accuracy of depth perception in AR. They used two types of shadows in their research: (i) "drop shadows" and (ii) "ray-traced" cast shadows. Neither solutions correspond to realistic shadows since "drop shadows" correspond to a vertical projection of the virtual object on the floor while the so-called "ray-traced" cast shadows correspond to the display of a elliptic shadow at a position computed by a ray-cast between the light source and the floor plane. Hence, the shape of the shadows were not realistic in that they were always elliptic and do not change according to the real shape of the objects. Shadows' opacity was modified according to a relatively simple approximation based on the distance between lights sources and the furthest point of the shadow. Moreover, it should be noted that the use of "drop shadows" changes the experimental task since participants do not evaluate egocentric distance to a virtual object but rather to the virtual shadow.

In fact, with the development of the rendering capabilities in AR, it is now possible to use different kinds of real-time shadows in AR scenes. Indeed, shadows can range from being very realistic to very crude. Of course, the more realistic they look like, the more expansive they are to

compute. This can still be an issue for AR HMDs which do not always benefit from a lot of computing power.

Based on those observations, a question naturally arises regarding the proper choice of AR shadows under different conditions. Do different kinds of shadows influence differently human perception in AR? To the best of our knowledge, research carried out to answer this question is very limited. One of the related work is that of Sugano et al. [27], who studied the influence of lighting and shadows in AR in a subjective task where subjects had to evaluate the "virtual presence of virtual objects in an AR scene". Results of subjective questionnaires showed that virtual objects' shadows enhance their "virtual object presence" by providing a strong connection between them and the real world. In the same paper, the authors used dropped shadows in spatial arrangement tasks (depth ordering and height ordering virtual objects).

Related to the notion of virtual shadows in AR, the alignment of the simulated lighting model in AR with the physical lighting in the real world is also an issue that needs to be considered. Indeed, it can prove difficult to control and perform this alignment in a very precise way especially in uncontrolled real environments. Does a misalignment can influence human's perception in AR? Here again, to the best of our knowledge, there has been limited research on lightings misalignment between AR scenes and the real world. Still using the same AR VST-HMD setup, Sugano et al. [27] compared two lights conditions in a spatial arrangement task: (i) the virtual light position is same as the real light; (ii) the virtual light positioned "15 degrees higher from the real one", and results show no difference between the two conditions. It should be noted that in both conditions authors used "dropped shadows" and not realistic ones. Moreover, in condition (ii) the difference between both lights seems limited and is difficult to interpret. Finally, the experimental setup consisted of an AR scene displayed in near-field (on a table in front of the participant) and the real light used is a table lamp positioned next to the participant.

According to these observations, questions about the influence of realistic virtual shadows as well as alignment between real and virtual lights on distance perception in OST AR remain.

3. Research Hypotheses

It can be seen that perceptual issues have been relatively left aside in OST-HMD AR, mainly due to the lack of available hardware. As a consequence, the influence of virtual objects shadows as well as lighting (which can be particularly difficult to accurately reproduce in an AR environment) on distance perception remains mainly unresolved. Based on previous research, we propose the four following hypotheses:

H1 Participants underestimate egocentric distances in OST AR.

H2 Virtual objects shadows improve distance perception accuracy in OST AR.

H3 Shadows with different realism levels have different influences on distance perception in OST AR.

H4 Lighting misalignment between the real and the virtual environments negatively impacts distance perception in OST AR.

The purpose of the present study is to explore issues related to shadows in AR and their influence on depth perception. We decided to focus on an AR scene displayed in the medium-field distance, which, according to Cuttings' research [28] corresponds to the 1.5m-30m range and covers a wide scope of AR applications.

First, we explore the accuracy of distance evaluation in an OST AR scene as well as the influence of virtual shadows on distance perception. In a second step, we study the impact of lights' misalignment on depth perception in OST AR. Two experiments based on an OST-HMD were designed to study these issues.

Participants were required to match a virtual object with a real line on the floor in each trial. Distance judgment accuracy for each trial was recorded and analyzed.

4. Experimental Design

Two experiments were designed to verify our hypotheses. Since both experiments have very similar experimental procedures, tasks and apparatus, we decided to regroup their description in the following.

4.1. Apparatus

Both experiments were conducted on a Microsoft HoloLens OST-HMD. The HoloLens is an untethered HMD with a high angular resolution and offers an inside-out global sensor fusion system controlled by a fully custom onboard GPU [29] as well as precise head tracking and 3D mapping. In details, the HoloLens has a $30^{\circ} \times 17.5^{\circ}$ FOV with a resolution of 1268×720 pixels per eye. It contains an inertial measurement unit (IMU) and four sensors (a depth camera, a RGB camera, a four-microphone array, and an ambient light sensor) used to map the 3D environment in real-time. The HoloLens also supports voice input, basic gesture recognition, and head tracking. In addition, the HoloLens displays are fixed at an optical distance approximately 2.0m away from the user, which means that users have to accommodate near 2.0m to maintain a clear image. In other words, placing virtual augmentations around this distance can minimize the vergence-accommodation conflict. We selected the HoloLens mainly due to the quality and robustness of its tracking. In addition, a Bluetooth mouse was adopted as an input mechanism for our task (see Sect. 4.2). The AR scenes were built using Unity3D v2017.3.1f1 and the HoloLens Toolkit[6].

4.2. Experimental Environment and Tasks

[6] https://github.com/Microsoft/MixedRealityToolkit-Unity/releases?after=2017.4.0.0-rc

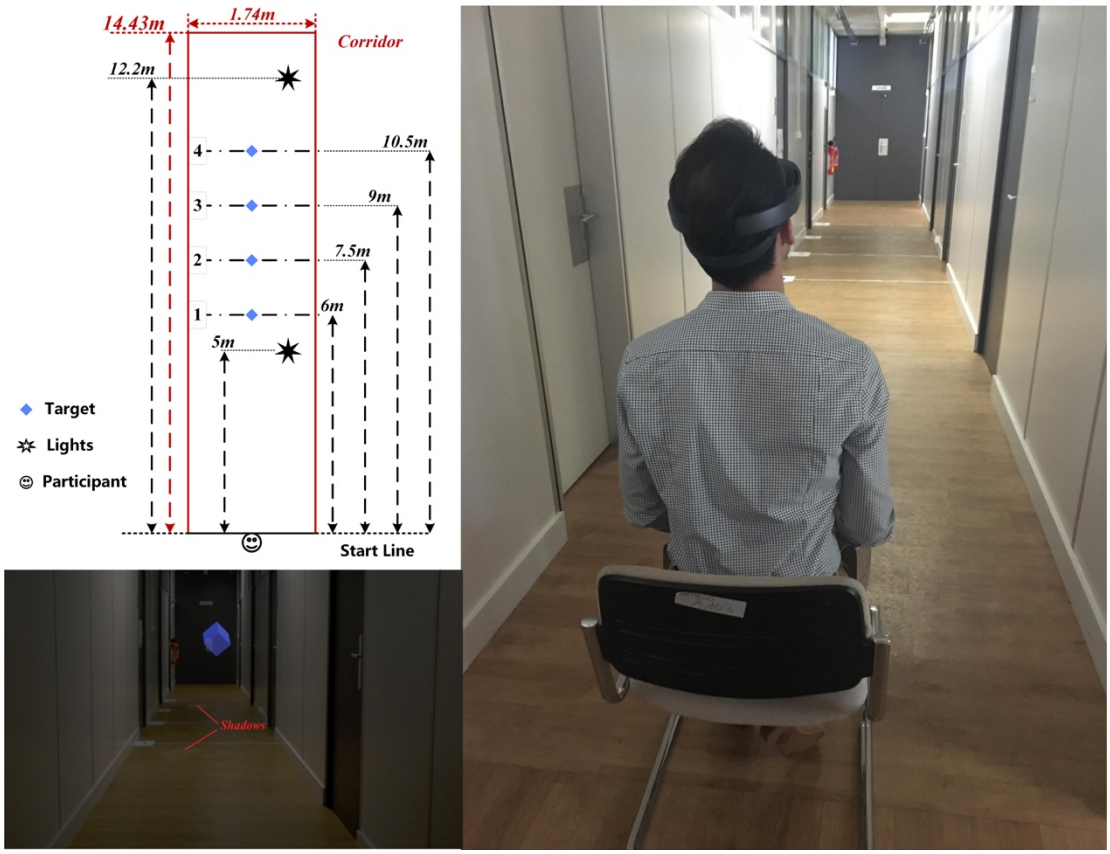

Figure 1: Top left: Top view of the position of the different targets and two real lights. Bottom left: View from the HoloLens. Right: Overall view of a participant in the corridor

The experiments were conducted in a corridor which is 14.43m long, 1.74m wide and 3.15m high. The corridor is an enclosed space and there were two lights on the ceiling for lighting, each light consisting of four fluorescent tubes. A virtual cube floated in the air at a height of 1.25m from the ground (see Fig. 1) and rotated at a constant speed of 12°/s to enhance participants' stereoscopic sensation. In order to facilitate the task the cube is oriented such that one of its vertex points to the floor (see Fig. 1). Participants saw the virtual cube through the HoloLens. Due to the HoloLens' limited FOV, it was necessary to guarantee that subjects had a low visual horizon so that they could see virtual shadows on the floor. In order to do so, subjects were seated on a chair located on a reference line at the end of the corridor (see Fig. 1). In addition, four white lines were drawn on the floor (at 6m, 7.5m, 9m and 10.5m from the subjects position) to serve as references for four different targets distances. Four A4 (210 × 297mm) pieces of paper with numbers from 1 to 4 printed on

them served as a way to recognize the targets (named as Target 1, Target 2, Target 3, Target 4 from nearest to furthest from the subjects position), see Fig. 1. The two real lights were located at 5m and 12.20m from the subjects' position and at 3.15m high.

According to Cutting's research [28], humans' perceptual space is divided into three distinct regions which are termed as near-field, medium-field, and far-field. The near-field refers to the distance within 1.5 meters, a little longer than arms' length, where people can perceive depth accurately. The medium field is between 1.5 meters and 30 meters, people can talk with each other within this distance and depth perception for stationary observers becomes underestimated within this field. Far-field refers to distances over 30 meters, and it was found that depth perception becomes increasingly underestimated as distance increases. In the present study, we selected four target distances within the medium-field perceptual space, i.e. between 6m and 10.5m. Three main factors were considered when determining this range for our targets: (i) the shadows of the virtual objects had to be visible by the participants, (ii) virtual objects should be displayed at a distance allowing natural interaction and (iii) virtual objects should be within a comfort zone of the HoloLens.

Regarding (i), as shown in Fig.1, the two real lights were located at 5m and 12.20m from the subjects' position at a height of 3.15m. If the virtual cube was too close to the participants, its shadows could be displayed behind them. On the other hand, if the virtual cube was too far away from the participants, its shadow would be invisible (i.e. too far away and too small). Preliminary tests allowed us to determine that the target distance range of 6m to 10.5m could guarantee shadows of the virtual cube to be displayed on the ground within a visible area. This range also corresponds to a common distance range for human daily interaction which is highly suitable for most AR applications (cf. criterion (ii)). Finally, although the distance range of 6m to 10.5m

exceeds the optimal distance (1.25m to 5m) for placing virtual objects in the HoloLens, this distance range is still within a relative comfort zone for HoloLens, and will not be uncomfortable[7] for users (criterion (iii)).

Each experiment consisted of a fixed number of trials. Participants saw a virtual cube floating in the air with or without a shadow. They could move the cube forward and backward along the axis of the corridor by scrolling the wheel of the Bluetooth mouse. The minimal step of the wheel induced a 5cm translation of the cube. Participants had to match the lowest vertex of the cube with a target line on the ground specified in each trial (see Fig. 1, bottom left). Target lines were announced vocally to participants directly using the HoloLens built-in speakers at the beginning of each trial. Participants needed to confirm the target line to make sure they picked the correct one. Participants confirmed the final location of the cube by clicking on the mouse's wheel. We recorded the final distance of the cube for each trial and loaded the next trial automatically. In order to guarantee that participants understood the procedure of the experiment correctly, each of them went through a pre-experiment practice session which included 10 trials of the same task.

4.3. Experimental Procedure

Both experiments were within-subject studies and followed the same procedure. Each experiment lasted for 20-25 minutes. The procedure of the experiment is depicted in Fig. 2 and consisted of five steps: (1) Screening; (2) Introduction and subject informed consent letter; (3) Calibration; (4) Practice session; (5) Formal experiment. Screening consisted of making sure that participants had normal or corrected to normal vision as well as normal color vision and normal stereoscopic vision. In order to do so, all the participants performed visual function tests including the visual acuity test, the stereo acuity test and the Ishihara test [30].

[7] https://docs.microsoft.com/en-us/windows/mixed-reality/comfort

In addition, the subjects experience with VR and AR was also recorded through a subjective questionnaire before the experiment. After that, the experimenter introduced the experiment content to the subjects, without mentioning the goals, and gave them the Subject Informed Consent Letter. This letter tells subjects that the experiment design and procedure were in accordance with the Declaration of Helsinki. Subjects were also informed that they could leave the experiment at any time or if they felt uncomfortable without the need of getting permission from the experimenter nor to justify themselves. Finally, they signed the letter.

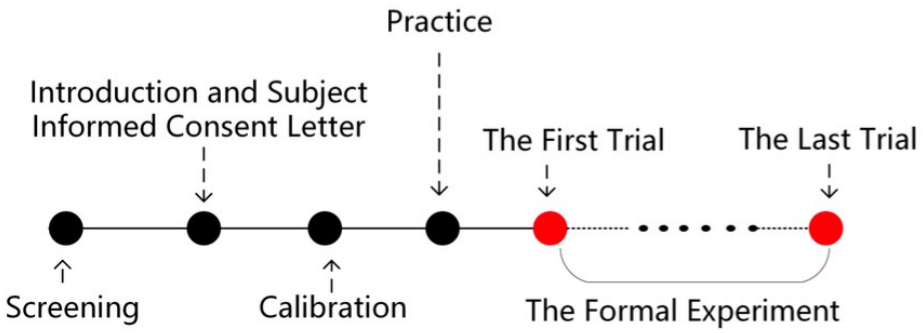

Figure 2: Experimental Procedure.

Participants sat on a chair and were helped to put on the HoloLens. A calibration procedure was carried out to guarantee that the trajectory of the virtual cube matched the axis of the corridor. The calibration procedure is as follows: participants saw three dots arranged in a line in the HoloLens' FOV. They had to align the dots with the middle axis of the corridor and make sure this line was perpendicular to the corridor's start line (see Fig. 1, top left). After the calibration, the practice session, consisting of 10 trials and during which the experimenter explained the task, loaded automatically. Once the practice session was over, the formal experiment started. The order of the trials was randomized for each participant. In both experiments, the task was identical to that of the practice session which is to match the position of the virtual cube with a specified target on the floor. At the beginning of each trial, a vocal announcement in the HoloLens built-in speakers specified the target.

Objective measures were recorded: signed distance errors, absolute distance errors and completion time. The signed distance error represents the distance between the center of the final position of the virtual cube (aligned with the lowest vertex) and the target along the Z-axis (i.e. the axis of the corridor). A signed distance error > 0, means that it is an egocentric underestimation (i.e. the cube is behind the target line). A signed distance error < 0, means that it is an egocentric overestimation (i.e. the cube is in front of the target line). The absolute distance error is the absolute value of the signed distance error.

5. Experiment 1: Studying the influence of the shape of virtual objects' shadows on distance perception in OST AR

The main purpose of the first experiment is to verify: **H1** "Participants underestimate egocentric distances in OST AR", **H2** "Virtual objects' shadows improve distance perception accuracy in OST AR", and **H3** "Shadows with different realism levels have different influences on distance perception in OST AR".

A matching task under different conditions was designed to verify these hypotheses. For more details on the task, please refer to Sect. 4.2. Participants accomplished the tasks under the conditions of presence or absence of virtual objects' shadows in the AR scene. In addition, when virtual shadows were present, three different levels of shadow realism were studied. Hypotheses **H1**, **H2** and **H3** were then assessed by comparing distance errors under the aforementioned conditions.

5.1. Experimental Conditions

The conditions that we want to manipulate in this experiment were (i) the presence (Shadow ON) or absence (Shadow OFF) of virtual objects' shadows in the AR scene and the impact of the level of realism of virtual objects' shadows. This second condition was decomposed into three sub-conditions, namely:

- “Round” shadows (see Fig. 3 top right): they have an incorrect shape and have sharp crispy edges. In the following, we refer to them as *round shadows*.
- “Hard-edge” shadows (see Fig. 3 middle right): they have a correct shape but have sharp crispy edges. In the following, we refer to them as *hard-edge shadows*.
- “Soft-edge” shadows (see Fig. 3 bottom right): the more realistic shadows which shape is correct and which fade off at the edges to represent a penumbra area. In the following, we refer to them as *soft-edge shadows*.

When the shadows are ON, we display two shadows (see Fig. 3) on the ground in a coherent way: the ground plane being detected during the HoloLens’ calibration procedure; the virtual shadows lie on a virtual plane that exactly corresponds to the real ground plane.

To implement our shadows we rely on Unity3D’s internal shadow implementation using the “hard shadows” and “soft shadows” options as well as the “Ultra” quality settings for our *hard-edge* and *soft-edge* shadows respectively. While Unity’s “hard shadows” are generated using the *very high* resolution parameter, the “soft shadows” are produced in *low resolution*. This creates a blurrier shadow and increases the difference between the two shadow renderings. Regarding the “round shadows” we used the same settings as for the “soft-edge shadows” for which we used a sphere to cast the shadows instead of the cube. The sphere has a diameter equals to that of the cube’s side. Participants still see the virtual cube, the sphere was only used as the geometry to cast shadows. More information about Unity3D’s shadows is available online[8].

We therefore had 4 shadows conditions: (i) Shadows OFF: no shadows were displayed in the AR scene; (ii) “Round” Shadows ON; (iii) “Hard-Edge” Shadows ON; (iv) “Soft-Edge” Shadows ON.

[8] https://docs.unity3d.com/Manual/ShadowOverview.html

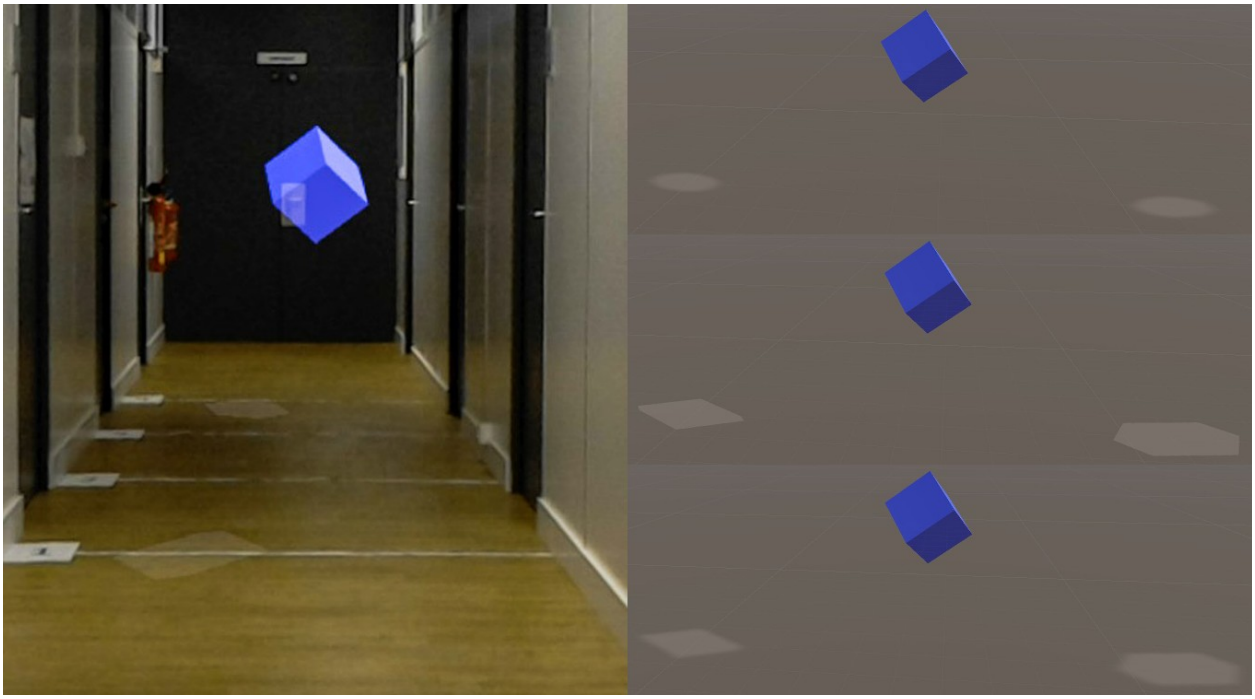

Figure 3: Left: Participant's viewpoint within the HoloLens in the "hard-shadow" condition. Right: The different shadow conditions, from top to bottom: "Round", "Hard-Edge" and "Soft-Edge" shadows.

The shadows that would have been cast by the two lights of the room are computed in real-time in Unity3D and are consistent with the size and positions of the real lights in the corridor thus the shape, size and location of the virtual shadows vary according to the position of the virtual cube in the corridor.

Regarding virtual shadows and OST-HMDs, one should take into account the following limitation: due to the transparency of the screens, it is impossible to display black in OST-HMDs. Two options were then possible regarding the display of virtual shadows in our AR scene: (i) artificially increase the brightness of the whole AR scene, except for the shadows or (ii) increase the brightness of the shadows only.

Both options present advantages and drawbacks. Option (i) would allow for more realistic shadows since they would appear darker than option (ii). Nevertheless, given the HoloLens' relatively small FOV, option (i) would give rise to a brighter window within the user's perspective. Indeed, the screens (i.e. where it is possible to render in the HoloLens) would seem brighter than

the rest of the user's view, e.g. the peripheral view. Option (ii) on the contrary would keep a coherent brightness between the screens and the "non renderable" part of the HoloLens' lenses. As a consequence, we chose option (ii), that is to display our virtual shadows in a slightly brighter way.

The experiment was carried out at four different target distances: 6m, 7.5m, 9m and 10.5m. Moreover, the largest distance used in our studies (10.5m) guarantees that participants were still within a relative comfort zone for the HoloLens[9].

960 trials (twenty observers× four types of shadow ×four distances ×three repetitions) were conducted in the experiment.

5.2. Participants

A total of 22 participants passed the screening step and took part in this experiment (11 males and 11 females). They were between 21 and 31 years of age (M = 24.91, SD = 2.72). Results of the subjective questionnaires show that 7 participants had no experience with VR or AR, 12 participants had little experience (used VR or AR 1-5 times), and 3 participants were experienced users (> 5 times). Each participant completed 48 trials.

5.3. Data Analysis and Results

In order to conduct our analysis, and based on our measurements, we used as dependent variable the relative error of egocentric distance estimation. This is computed as follows:

$$E_{\mathrm{relEgo}} = (d_{\mathrm{cube}} - d_{\mathrm{target}})/d_{\mathrm{target}} \quad (1)$$

where d_{cube} is the distance of the cube estimated by the participant and d_{target} is the real distance of the target.

We removed 5 participants from the initial dataset: 2 of them misunderstood instructions and

9 https://docs.microsoft.com/en-us/windows/mixed-reality/comfort

thus did not conduct the tasks in a right way, another 2 subjects had abnormally high standard deviations (more than 3 times the standard deviations of other subjects). The last one was removed because of its large number of outliers (12 among 48 trials). Data from the remaining 17 participants were kept for our statistical analysis, using IBM SPSS Statistics 24.0 with a 0.05 significance level.

We first conducted Shapiro-Wilk tests to assess data normality for different levels of factors, which includes four types of shadows (OFF, Round, Hard-edge, Soft-edge) and four distances (6m, 7.5m, 9m, 10.5m). The results show that the data does not distributed normally: OFF (W = 0.970; $p < 0.001$), Round ($W = 0.981$; $p = 0.008$), Hard-edge ($W = 0.984$; $p = 0.022$), Soft-edge ($W = 0.976$; $p = 0.002$); 6m ($W = 0.924$; $p < 0.001$), 7.5m ($W = 0.945$; $p < 0.001$), 9m ($W = 0.960$; $p < 0.001$), 10.5m ($W = 0.975$; $p = 0.001$).

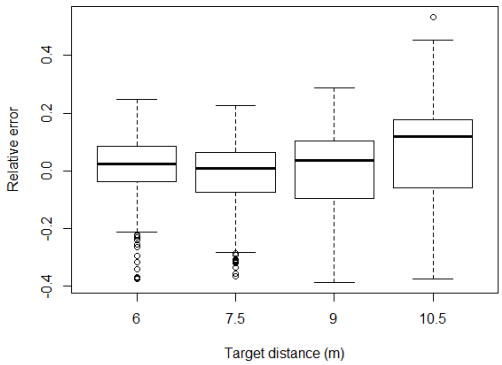

Figure 4: Experiment 1. Relative errors in distance judgments, averaged for all participants grouped per distance condition.

Thus, we decided to conduct non-parametric tests on the dataset. First, a Kruskal-Wallis H Test was applied. Results showed, as expected, a significant effect of distance ($H = 52.631$; $p < 0.001$) confirming that judgment errors increase along with the distance between the cube and participants, see Fig. 4. We followed by computing post-hoc tests: all pairwise comparisons with adjusted p-values using the Bonferroni correction for multiple tests. Significant effects were found between (i) 6m and 10.5m ($Z = -5.227$; $p < 0.001$), (ii) 7.5m and 10.5m ($Z = -6.951$; $p < 0.001$), (iii) 9m and 10.5m ($Z = -4.442$; $p < 0.001$).

As for the effects of shadows, unexpected to us, the Kruskal-Wallis H test showed a non-significant effect of shadow type ($H = 7.446$; $p = 0.059$), see Fig. 5. This was unexpected since the presence of shadows has been reported to increase accuracy of distance judgments in AR, see e.g. [6]. However, since the Kruskall-Wallis' p value is close to the significance level used in our analysis, we computed Mann-Whitney U Tests to compare data of each different shadow pair.

Results showed that there is significant difference between the Shadow OFF condition and other conditions (i) OFF and Round-edge ($Z = -2.064$; $p = 0.039$), (ii) OFF and Hard-edge ($Z = -2.435$; $p = 0.015$), (iii) OFF and Soft-edge ($Z = -2.106$; $p = 0.035$), see Fig. 5. Moreover, results showed a non-significant effect of the three Shadow ON conditions.

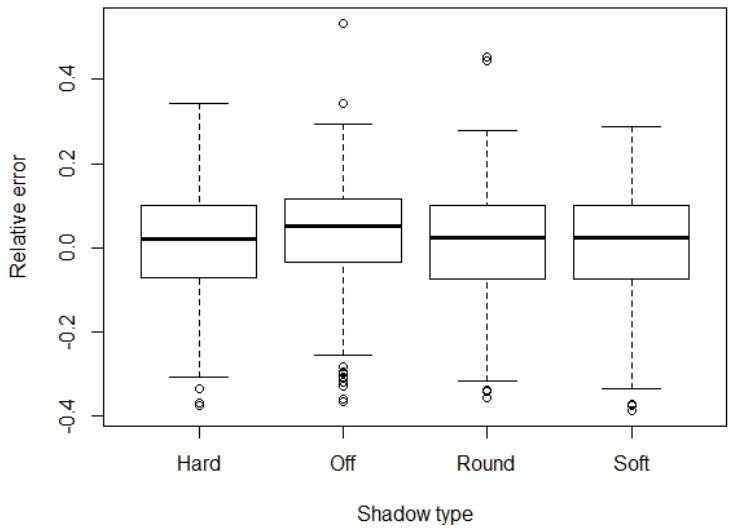

Figure 5: Experiment 1. Relative errors in distance judgments, averaged for all participants grouped per shadow type.

In addition, the completion time (i.e. the time it took for each participant to complete the task) has been analyzed for each trial. First, Shapiro-Wilk tests were applied to assess data normality for different levels of factors. Results show that the data is not normally distributed for any of the factors: OFF ($W = 0.841$; $p < 0.001$), Round ($W = 0.752$; $p < 0.001$), Hard-edge ($W = 0.840$; $p < 0.001$), Soft-edge ($W = 0.869$; $p < 0.001$); 6m ($W = 0.905$; $p < 0.001$), 7.5m ($W = 0.851$; $p < 0.001$), 9m ($W = 0.840$; $p < 0.001$), 10.5m ($W = 0.770$; $p < 0.001$).

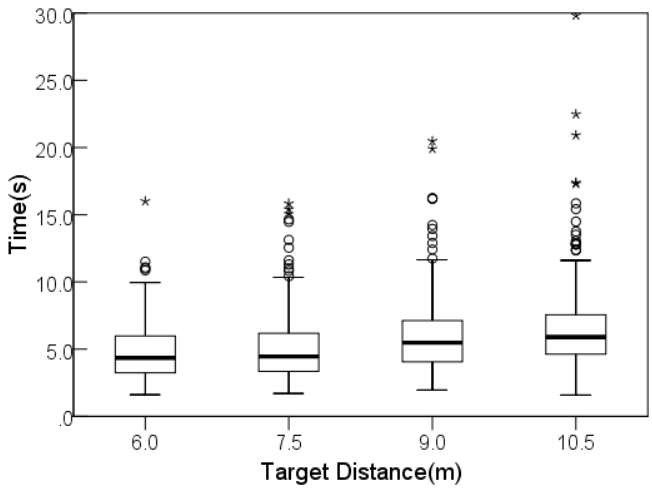

Figure 6: Experiment 1. Completion time in distance judgments, averaged for all participants grouped per target distance.

Then, a non-parametric Kruskal-Wallis H Test was applied to the data of completion time. Results show a significant effect of distance ($H = 67.520$; $p < 0.001$) proving that completion time increases along with the distance between the cube and participants, see Fig. 6. We followed by computing post-hoc tests: all pairwise comparisons with adjusted p-values using the Bonferroni correction for multiple tests. Significant effects were found between (i) 6m and 9m ($Z = -4.488$; $p < 0.001$), (ii) 6m and 10.5m ($Z = -7.078$; $p < 0.001$), (iii) 7.5m and 9m ($Z = -3.886$; $p = 0.001$), (iv) 7.5m and 10.5m ($Z = -6.515$; $p < 0.001$).

As for the effects of shadows, the Kruskal-Wallis H test showed a non-significant effect of shadow type ($H = 1.904$; $p = 0.593$), see Fig. 7. Afterwards, we computed Mann-Whitney U Tests to compare data of each different shadow pair. However, no significant effect was found in the results.

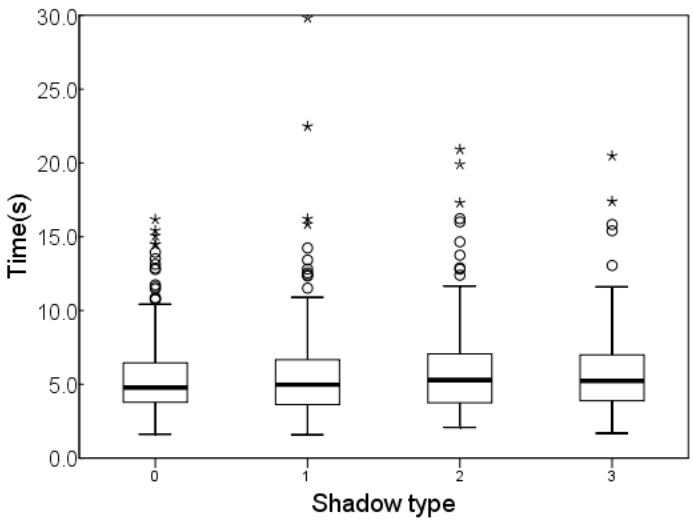

Figure 7: Experiment 1. Completion time in distance judgments, averaged for all participants grouped per shadow type.

From the results, it can be seen that there is a trend towards underestimation but the high variance levels of our results could not confirm this bias. In order to further analyze this effect, we studied results for each participant individually.

Participants' individual data were observed to confirm the stability of the above findings. As shown in Fig. 8, 9 out of 17 subjects have positive relative errors, 5 out of 17 subjects have negative relative errors, and 3 out of 17 subjects have both positive and negative relative errors. In other words, the number of subjects who underestimate distances in AR accounts for the largest proportion in experiment 1. This result is consistent with the general trend of the data.

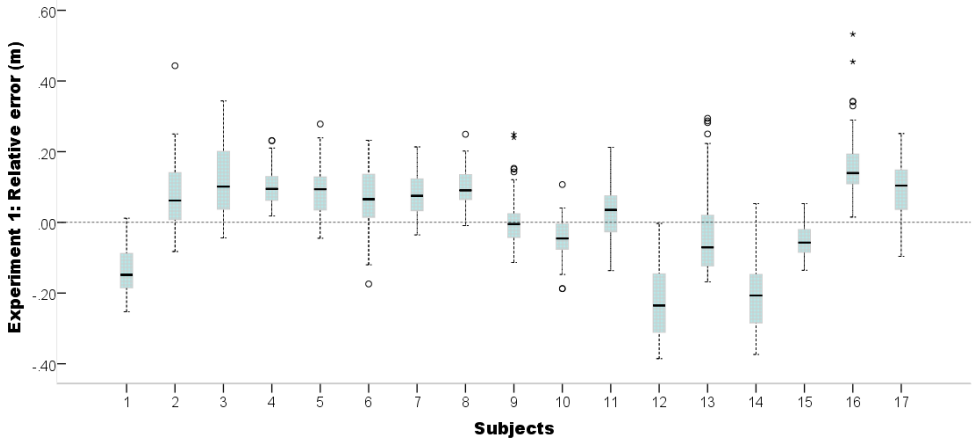

Figure 8: Experiment 1. Individual relative errors in distance judgments.

In addition, the mean relative error of each subject under different shadow types was computed. Results show that 11 out of the 17 subjects have higher distances judging accuracy under at least one of the shadows ON conditions compare with the shadows OFF condition (See Tab. 1).

Table. 1 Individual mean relative errors in distance judgments under different shadow types.

Shadows/Subject ID	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
Off	*-.115*	*.076*	*.128*	*.099*	*.107*	*.086*	*.073*	*.114*	*.014*
Round	*-.143*	*.100*	*.115*	*.108*	*.118*	*.035*	*.075*	*.091*	*.008*
Hard	*-.144*	*.063*	*.123*	*.091*	*.072*	*.060*	*.080*	*.084*	*.011*
Soft	*-.144*	*.060*	*.118*	*.108*	*.075*	*.076*	*.081*	*.095*	*-.012*
Shadows/Subject ID	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*-*
Off	*-.030*	*.049*	*-.225*	*.164*	*-.177*	*-.058*	*.199*	*.091*	*-*
Round	*-.028*	*.014*	*-.227*	*-.085*	*-.208*	*-.055*	*.147*	*.093*	*-*
Hard	*-.062*	*.019*	*-.220*	*-.087*	*-.226*	*-.036*	*.156*	*.076*	*-*
Soft	*-.057*	*.027*	*-.235*	*-.086*	*-.211*	*-.054*	*.156*	*.112*	*-*

On the other hand, the mean relative error of each subject under different targets' distances was computed. Results show that 14 out of 17 subjects have the highest mean relative errors under

distance 10.5m (See Tab. 2).

Table. 2 Individual mean relative errors in distance judgments under different targets' distances.

Distances/Subject ID	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
6m	*-.184*	*-.002*	*.039*	*.091*	*.052*	*.017*	*.087*	*.123*	*-.004*
7.5m	*-.195*	*.049*	*.055*	*.058*	*.078*	*.007*	*.014*	*.059*	*-.039*
9m	*-.113*	*.083*	*.142*	*.079*	*.098*	*.067*	*.056*	*.061*	*-.002*
10.5m	*-.054*	*.170*	*.249*	*.178*	*.144*	*.165*	*.152*	*.142*	*.066*
Distances/Subject ID	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*-*
6m	*-.014*	*.073*	*-.219*	*.019*	*-.216*	*-.044*	*.133*	*.032*	*-*
7.5m	*-.065*	*.007*	*-.233*	*-.031*	*-.223*	*-.064*	*.106*	*.067*	*-*
9m	*-.027*	*-.016*	*-.239*	*-.031*	*-.212*	*-.074*	*.123*	*.114*	*-*
10.5m	*-.072*	*.045*	*-.216*	*-.050*	*-.171*	*-.022*	*.298*	*.159*	*-*

As for the results of the completion time, we also computed the individual data for each subject. Due to our statistical analysis only found significant distances' effect on completion time, we computed the mean completion time for each subject under different targets' distances (See Tab. 3). The result shows 14 out of 17 subjects have longer mean completion time under the distance of 9m and 10.5m than under the distance of 6m and 7.5m. This result is coherent with the global statistical analysis.

Table. 3 Individual mean completion time in distance judgments under different targets' distances.

Distances/Subject ID	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
6m	*4.184*	*4.901*	*5.012*	*4.472*	*5.568*	*3.604*	*6.043*	*5.062*	*5.634*
7.5m	*5.299*	*5.319*	*6.373*	*4.768*	*6.433*	*3.762*	*6.465*	*4.854*	*4.620*
9m	*5.300*	*6.176*	*5.467*	*5.922*	*6.857*	*4.182*	*8.763*	*5.794*	*6.148*
10.5m	*6.099*	*5.599*	*6.455*	*6.461*	*7.126*	*5.036*	*8.668*	*6.966*	*5.610*
Distances/Subject ID	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*-*

6m	*3.504*	*3.534*	*6.894*	*3.423*	*4.242*	*8.830*	*3.982*	*2.807*	*-*
7.5m	*3.176*	*3.125*	*7.970*	*3.854*	*4.028*	*11.473*	*4.193*	*2.557*	*-*
9m	*4.624*	*4.685*	*8.540*	*3.857*	*3.797*	*13.706*	*5.510*	*3.497*	*-*
10.5m	*4.227*	*5.177*	*10.292*	*7.202*	*4.425*	*15.265*	*6.338*	*4.113*	*-*

5.4. Discussion

As mentioned in Sect. 2, many previous studies showed that shadows are very important for human's visual perception, and that they represent an important depth cue [25, 26]. While it has been showed that shadows can help people confirm objects' position in a virtual scene [31, 32], research on shadows' influence on depth perception in AR is very limited, even more so in OST AR. The only relevant study we know of is that of Diaz et al. [6].

While our results verified the results from [6]. We believe that our experiment design was more rigorous, which makes the conclusion more convincing. Indeed, while our shadows were designed on purpose to be as close as possible to reality, [6] relied on either "dropped shadows" or "ray-traced shadows" (see Sect. 2) which could have impacted their results. Moreover, our experimental design, unlike [6], purposely reduced the number independent variables in order to (i) avoid overfitting (i.e. being able to find statistical evidence even in presence of noise) and (ii) respect as possible the principle of parsimony.

The present experiment aimed at investigating the influence of different levels of realism shadows on distance perception in OST AR. First of all, we wanted to test whether the presence or absence of virtual objects' shadows would improve distance perception accuracy in OST AR (hypothesis **H2**). Our results show that while distance perception is better in the shadows ON condition, no significant difference was found between the three shadows conditions. Participants' individual data were also observed to confirm this phenomenon. Results show that for most

subjects (11 out of 17), their individual data were coherent with this result. We believe it is a very interesting result, since it would tend to show that shadow type does not to play an important role on distance judgment in OST-HMDs, disproving our hypothesis **H3**. This is a meaningful result since application developers could use "cheap" non-realistic shadows, like our round shadows, instead of costly ones, especially since most OST-HMDs use standalone computers with limited CPUs/GPUs.

Finally, regarding hypothesis **H1** (participants underestimate distances in OST AR), we can confirm a trend towards underestimation but high variance levels of our results could not confirm this bias. The relative errors for each individual were also observed, and we found that most of the participants (9 out of 17) underestimate distance in AR, and 3 out of 17 subjects have both positive and negative relative errors. In addition, our results also confirm that errors on perceived distances increase along with the distance between the virtual object and participants.

As for the completion time of each trial, although the initial position of the virtual cube is random, the statistical analysis reveals that completion time of trials increases significantly along with the distance, however, the shadow type has no significant effect on the completion time of trials. It could be inferred that the participants need to spend more time on matching task when the targets' distances were farther away.

6. Experiment 2: Studying the influence of lighting misalignment on distance perception in OST AR

This second experiment aims at assessing hypotheses **H1**: "Participants underestimate egocentric distances in OST AR" and **H4**: "Lighting misalignment between the real and the virtual environments impact negatively depth perception in OST AR". Similar to experiment 1, details about the apparatus and the task have been given in Sect. 4.2.

6.1. Experimental conditions

In order to select the type of virtual shadow used in this experiment, we used results from the first experiment. Since we did not want to induce a bias in depth perception, we chose to rely on a realistic type of shadow, namely the *hard-edge* shadow. Thus, all of the trials in this experiment were performed with the shadow ON condition and using the *hard-edge shadow* setting, see Fig. 3, middle right.

We selected three virtual lights' positions, see Fig. 9: (i) *L0m*: coherent lighting conditions where the positions of virtual lights coincide with the real lights' positions, (ii) *L+3m*: virtual lights were positioned 3 meters behind the real lights' positions along the Z-axis, and (iii) *L-3m*: virtual lights were positioned 3 meters before the real lights' positions along the Z-axis.

In addition, we also conducted the experiment under four different target distances, as in the first experiment, namely 6m, 7.5m, 9m and 10.5m. All the other settings of the trials as well as the task are similar with those of the first experiment. A total of 756 trials (twenty-one observers × three lights' positions × four distances three repetitions) were conducted in this experiment.

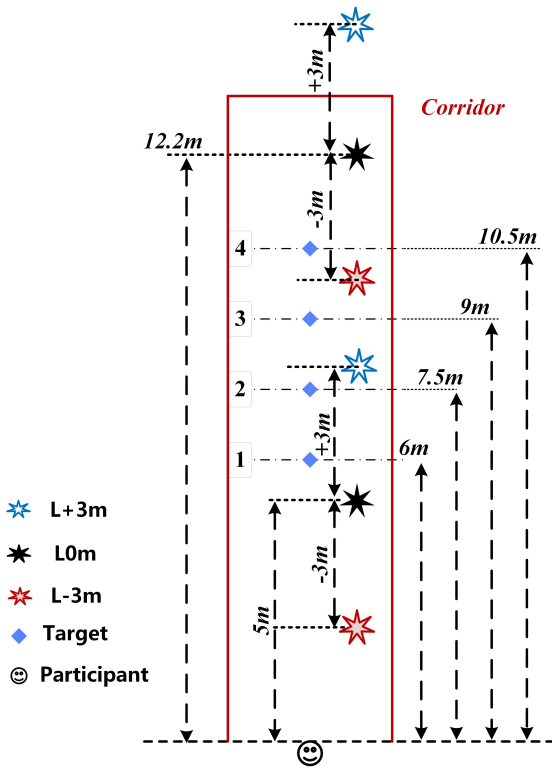

Figure 9: Experiment 2. Lighting conditions.

6.2. *Participants*

A total of 21 participants (all different from experiment 1) passed the screening step and took part in this experiment (11 males and 10 females). Participants were between 21 and 29 years of age (M = 25.9, SD = 3.19). Results of the subjective questionnaires showed that 11 participants had no previous experience with VR or AR, 9 participants had little experience (1-5 times) with VR or AR, and 1 participant had extensive experience (> 5 times) with VR or AR. Each participant completed a total of 36 trials.

6.3. *Data Analysis and Results*

Like experiment 1, the dependent variable of our statistical analysis is the relative error between the cube and the floor reference target.

From the initial dataset, 3 participants were removed: 2 because of a standard deviation more than 5 times higher than the standard deviation of the other participants and one because of an

abnormal mean (being more than three times the standard deviation smaller than the global mean). Here again the analysis was performed using IBM SPSS Statistics 24.0 with a 0.05 significance level.

The final data of different light position levels and distance levels was processed by a Shapiro-Wilk test respectively. The results show that some levels of data were not distributed normally: L0 ($W = 0.978$, $p = 0.002$), 7.5m ($W = 0.978$, $p = 0.011$). Thus, non-parametric tests were applied to all of the data.

As in the previous experiment, we first computed a Kruskal-Wallis H Test. Results showed a significant effect of distance ($H = 22.610$; $p < 0.001$) and lights' position ($H = 15.961$; $p < 0.001$).

We thus followed with post-hoc tests: all pairwise comparisons with adjusted p values using the Bonferroni correction for multiple tests. Results found significant differences between (i) 6m and 10.5m ($Z = -4.277$, $p < 0.001$), (ii) 7.5m and 10.5m ($Z = -3.873$, $p = 0.001$). Other distances groups were not significantly different.

As for lights' positions: pairwise comparisons revealed significant difference between groups (i) L+3m and L-3m ($W = -3.947$, $p < 0.001$) and (ii) L+3m and L0m ($W = -2.511$, $p = 0.036$). Note that no significant difference was found between the L0m and L-3m groups.

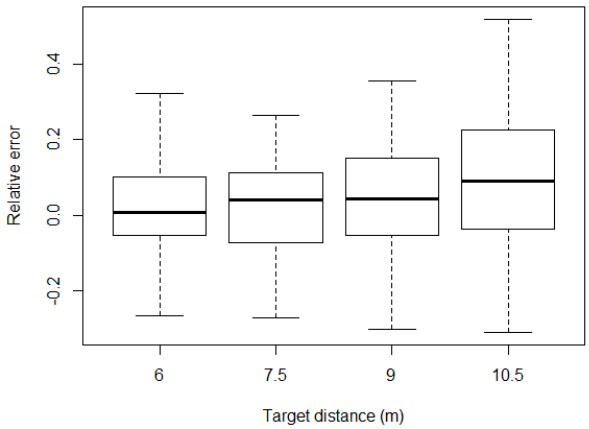

Figure 10: Experiment 2. Relative errors in distance judgments, averaged for all participants grouped per distance condition.

Fig. 10 presents the mean relative errors for all participants under four distance conditions. It can be observed that the mean relative error increases along with the distance of the virtual object.

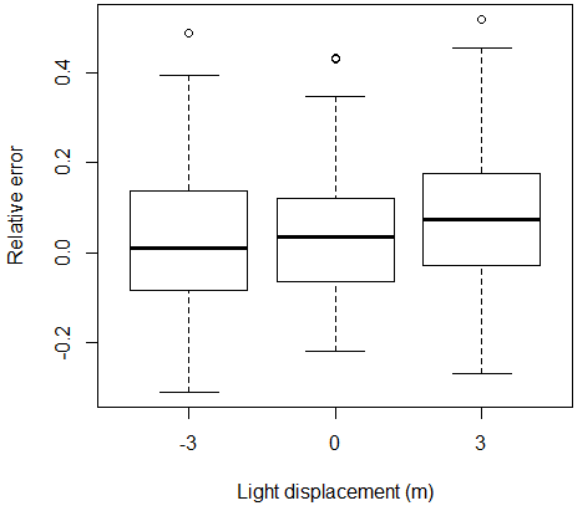

Figure 11: Experiment 2. Relative errors in distance judgments, averaged for all participants grouped per lighting placement.

Regarding hypothesis **H1**, our results are not conclusive. There seems to be a tendency toward a slight underestimation (*mean* = 0.046m), but the high variance levels do not allow us to conclude on a clear distance underestimation in OST AR no matter the lighting misalignment condition.

Fig. 11 presents the mean relative errors for all participants in distance judgments grouped per lighting placement. The mean relative error has the highest value under the condition L+3m. The condition L0m has a mean relative error slightly higher than that compared with the condition L-3m, however, the condition L0m also corresponds to a smaller standard deviation.

In the same way, the completion time for each trial was analyzed. Results of the Shapiro-Wilk test show that data were not distributed normally: L-3m ($W = 0.902$, $p < 0.001$), L0m ($W = 0.827$, $p < 0.001$), L+3m ($W = 0.879$, $p < 0.001$), 6m ($W = 0.838$, $p < 0.001$), 7.5m ($W = 0.795$, $p < 0.001$), 9m ($W = 0.874$, $p < 0.001$), 10.5m ($W = 0.909$, $p < 0.001$). Thus, a Kruskal-Wallis H Test was applied to all of the data. Results showed a significant effect of distance ($H = 41.523$; $p < 0.001$), see Fig. 12, and non-significant effect of lights' position ($H = 0.468$; $p = 0.791$), see Fig. 13.

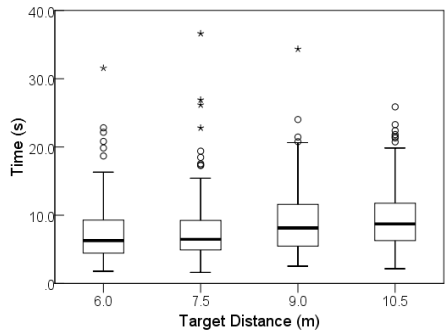

Figure 12: Experiment 2. Completion time in distance judgments, averaged for all participants grouped per target distance.

As post-hoc tests for the effect of distance we computed all pairwise comparisons with adjusted *p* values using the Bonferroni correction for multiple tests. Results found significant differences between (i) 6m and 9m ($Z = -3.902$, $p = 0.001$), (ii) 6m and 10.5m ($Z = -5.618$, $p < 0.001$), (iii) 7.5m and 9m ($Z = -3.091$, $p = 0.002$), (iv) 7.5m and 10.5m ($Z = -4.807$, $p < 0.001$). Other distances groups were not significantly different.

Then, the Mann-Whitney U Test was used to check the differences between different levels of lights position, however no significant result was found.

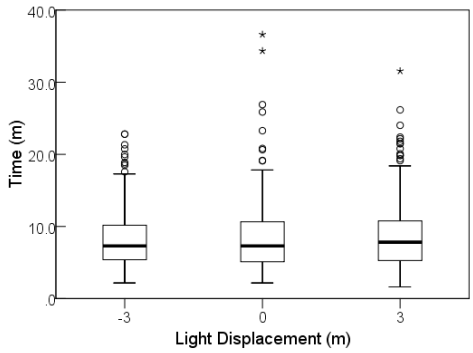

Figure 13: Experiment 2. Completion time in distance judgments, averaged for all participants grouped per lighting placement.

Similar to experiment 1, we computed the participants' individual data to confirm the stability of our findings. As shown in Fig. 14, 8 subjects have positive relative errors, 4 subjects have negative relative errors, and 6 of them have both positive and negative relative errors. In other words, the number of subjects who underestimate distances in AR accounts for the largest

proportion in experiment 2. This result is consistent with the general trend of the data.

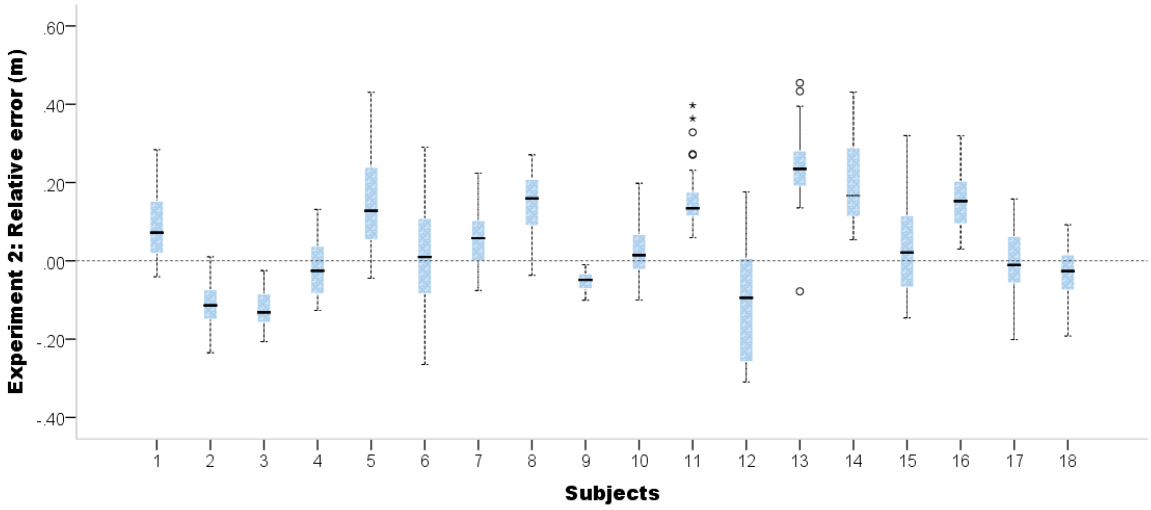

Figure 14: Experiment 2. Subjects individual relative errors.

In order to reinforce the findings about the lights placements' effects on distance relative errors, we computed for each subject the mean relative error under different light placements (see Tab. 4). As mentioned above, results show that L+3m corresponds to a significant higher relative error than L0m, while L-3m corresponds to smaller relative error than L0m. From Tab. 4, we can see that for 9 out of the 18 subjects, results were coherent with the general trend. In addition, three subjects (Subject 3, Subject 13, and Subject 16) show the highest distance judging accuracy under the condition L0m. A Shapiro-Wilk test was adopted to assess data normality for relative error of these three subjects under different lights placements. The results show that the data of Subject 3 and Subject 13 distributed normally, while the data of Subject 16 under L-3m condition does not distributed normally: L-3m (W =0.847 ; p =0.034). A paired sample t-test was applied to data of Subject 3 and Subject 13. The result shows that there is no significant difference between L-3m and L0m for these two subjects. Subject 13 exhibit a significant higher relative error under L+3m condition compared to the L0m condition (Subject 13: t_{11}=2.366, p=0.037). In addition, A Mann-Whitney U was adopted to analyze relative errors of Subject 16. The result shows that L+3m

exhibits significant higher relative errors than L0m (Z=2.021, p=0.043). This means that most of the subjects' (12 out of 18) individual data were coherent with the general result.

Table 4. Individual mean relative errors under different lights placements.

Lights/Subject ID	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
L-3m	.092	-.113	-.144	-.092	.187	.024	.016	.137	-.049
L0m	.067	-.117	-.121	.023	.145	-.006	.065	.119	-.052
L+3m	.105	-.110	-.098	.019	.127	.044	.093	.182	-.057
Lights/Subject ID	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*
L-3m	-.009	.131	-.257	.219	.171	-.090	.140	.022	-.046
L0m	.006	.157	-.098	.239	.186	.039	.169	-.019	-.039
L+3m	.064	.188	.043	.250	.219	.155	.153	-.020	-.008

Moreover, we computed the mean relative error under different distances for each subject (see Tab. 5). The result shows that 13 out of 18 subjects have the highest mean relative error under the distance of 10.5m. This means most of the subjects' data coherent with the general trends.

Table 5. Individual mean relative errors under different targets' distance.

Distances/Subject ID	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
6m	*.002*	*-.133*	*-.130*	*-.003*	*.036*	*-.048*	*-.002*	*.118*	*-.034*
7.5m	*.078*	*-.137*	*-.144*	*-.004*	*.131*	*-.049*	*.056*	*.125*	*-.065*
9m	*.091*	*-.108*	*-.115*	*-.017*	*.143*	*.022*	*.072*	*.174*	*-.057*
10.5m	*.181*	*-.076*	*-.095*	*-.041*	*.302*	*.158*	*.106*	*.168*	*-.054*
Distances/Subject ID	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*
6m	*.026*	*.113*	*-.059*	*.200*	*.160*	*.102*	*.105*	*-.075*	*-.046*
7.5m	*.010*	*.122*	*-.098*	*.185*	*.138*	*.047*	*.097*	*-.066*	*-.048*
9m	*.015*	*.144*	*-.155*	*.261*	*.210*	*.016*	*.172*	*.017*	*-.022*
10.5m	*.030*	*.256*	*-.104*	*.298*	*.260*	*-.027*	*.241*	*.102*	*-.009*

Similar with experiment 1, we also computed the individual mean completion time for each

subject. Due to our statistical analysis only found significant distances' effect on completion time, we computed the mean completion time for each subject under different targets' distances (See Tab. 6). The result shows 14 out of 18 subjects have longer mean completion time under the distance of 9m and 10.5m than the distance of 6m and 7.5m. This result is also coherent with the global statistical analysis.

Table. 6 Individual mean completion time in distance judgments under different targets' distances.

Distances /Subject ID	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
6m	*6.162*	*5.864*	*7.287*	*9.205*	*8.295*	*8.125*	*11.711*	*3.999*	*9.144*
7.5m	*6.041*	*6.003*	*9.850*	*10.107*	*8.363*	*14.086*	*12.263*	*5.212*	*10.173*
9m	*7.172*	*7.269*	*11.585*	*7.186*	*11.133*	*14.807*	*13.424*	*5.212*	*13.605*
10.5m	*10.173*	*7.971*	*14.900*	*8.130*	*10.209*	*13.583*	*14.317*	*6.365*	*11.725*
Distances /Subject ID	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*
6m	*13.283*	*10.545*	*5.801*	*5.346*	*8.803*	*8.925*	*3.507*	*3.567*	*5.382*
7.5m	*10.02*	*11.133*	*3.730*	*5.788*	*7.000*	*6.989*	*4.113*	*3.992*	*6.041*
9m	*10.514*	*12.934*	*3.935*	*6.802*	*10.015*	*9.296*	*4.679*	*6.344*	*7.649*
10.5m	*11.168*	*13.30*	*5.317*	*9.612*	*9.193*	*8.931*	*5.843*	*8.477*	*8.144*

6.4. Discussion

The present paper is the first study to investigate the influence of the lighting misalignment on distance perception quantitatively in OST AR. As in the first experiment, we found a distance underestimation trend in AR OST-HMDs. However, high variance levels prevent us from

confirming it statistically.

As expected, we found a significant influence of lighting coherence's on distance perception in OST AR. The condition L+3m has a significantly higher distance judgment error than L0m. Although the mean relative error of L-3m is slightly lower than L0m, we did not find any significant difference between these two conditions. In addition, since L-3m leads to a better distance estimation accuracy than L+3m, this could imply that placing the virtual lights closer to the participants might be better than placing them farther away. Individual data were also observed to confirm these results, and we found that 12 out of the 18 subjects' individual data were coherent with the general result.

Obviously, the direct consequence of changing virtual lights' locations is the modification of the virtual shadows' positions. In condition L+3m, the virtual lights were farther away from the participants and as a consequence virtual shadows were closer to them. Since our participants wore an OST-HMD, this created a perceptual conflict. Indeed, in OST AR, participants still see the real lights. Thus participants had to estimate distance of the virtual cube based on the real lights' positions and the virtual shadows (that were closer to them than they should have been). This could lead participants to believe that objects were closer to them and thus to underestimate the egocentric distance to the virtual cube. In other words, the lights' location L+3m might have increased the degree of distance underestimation in OST AR.

In a similar, but opposite way, condition L-3m should have lessened underestimation of distance perception in OST AR. However, results showed a non-significant lower relative error in condition L-3m compared to L0m (the same relative error is significant between L-3m and L+3m). This means that we did not notice a cancelling, let alone a reduction of egocentric distance estimation in OST AR when moving the virtual lights closer to the participant than in reality.

As for the distances' influence on the relative errors, we found some significant results that

distance perception at 10.5m is significantly worse than the condition of 6m and 7.5m. We can also observe a global trend that the mean relative error increases with the distance between the cube and participants. This trend is consistent with the results of the first experiment. Individual data for each subject in Tab. 4 shows that 13 out of 18 subjects have the highest mean relative error under the distance of 10.5m which means that most of the subjects coherent with the global trend.

Moreover, our results reveal that trials' completion time is significantly longer under farther targets' distances (9m and 10.5m) than under nearer targets' distances (6m and 7.5m) , while lighting placement has no significant effect on it. The participants' individual data was also consistent with the global trends.

7. General discussion

Our first hypothesis was, following some previously reported results, that people underestimate egocentric distances in OST AR. While we found a trend towards underestimation in both experiments, we cannot confirm it statistically due to the high variance levels of our data. This may be due to the task used in our experiments. Indeed, while our matching task (participants had to match the position of a virtual cube with that of a real target) is easier to evaluate and gives rise to very precise results, it may not be the best for egocentric distance estimation. In order to obtain more conclusive results, one should probably rely on other tasks such as verbal report, bisection tasks, blind walking, pointing or even throwing tasks. Nevertheless, it has been shown that tasks can have a strong effect on distance evaluation [21] and verbal report is known to be less accurate. In order to confirm the phenomenon of underestimation in the present study, we computed individual data for each subject in both of the two experiment, and we found that the number of subjects who underestimate distances in AR accounts for the largest proportion in both

of the two experiments.

In addition, our research verified that virtual objects shadows improve depth judgment accuracy in OST AR (**H2**), and that different levels of realism for virtual shadows have different influences on distance perception in OST AR (**H3**). Regarding **H2** and **H3**, while our results confirmed that the presence of shadows improve the accuracy of distance perception in OST AR (Mann-Whitney U Test for **H2** showed that the shadows OFF condition has a significant higher relative error than all the other shadows ON conditions). We found an interesting result in that there is no significant difference between the three types of virtual shadows used, rejecting **H3**. Hence, the results show that different realism levels of virtual shadows have no significant influence on the distance perception in OST AR. This is interesting because it implies that in OST AR application developers could rely on "cheap" unrealistic shadows instead of realistic costly virtual shadows. This is especially important since most OST AR HMDs use standalone computers with limited CPUs/GPUs. The study also reveals that errors on perceived distances increase along with the distance between the cube and participants.

Regarding hypothesis **H4**, namely that lighting misalignment between the real and the virtual environments negatively impact depth perception in OST AR, we only found a significant result in condition L+3m where virtual lights were farther away (+3m) from the participants than real lights. In that case, distance perception was significantly underestimated by participants. A possible explanation to this phenomenon is that by moving away virtual lights, the virtual shadows come closer to the participant thus impacting distance estimation and leading to an underestimation. Surprisingly, the opposite effect did not happen in condition L-3m where virtual lights were closer than real lights. One could wonder why the same amount of modification in positions of virtual lights in two opposite directions lead to different effects on egocentric distance perception. A possible reason is that underestimation is a complex phenomenon caused by

comprehensive factors in AR on general which have been proved in some previous research. Thus, moving the lights location closer cannot relieve the underestimation significantly. This result is intriguing since it could also mean that OST AR environments have a higher tolerance to lighting misalignment when the virtual lights are closer to the participants than when they are farther away from them.

In addition to the above hypotheses, a global trend was observed in both of the two experiments that the mean relative error increases with the distance between the cube and participants. The results found the relative errors under the target's distance of 10.5m are significant higher than those under other target's distances (6m, 7.5m, 9m) in experiment 1, while the distance of 10.5m corresponds to significant higher relative errors than the distance of 6m and 7.5m in experiment 2. The result also reveals that completion time of trials increases along with the distance between the virtual cube and the participants, while the shadow type and lights placement have no significant effect on completion time.

8. Conclusion and future work

In this paper, we investigated the influence of virtual objects shadows and lighting coherence on distance perception in OST AR. Our results can be summarized as follows.

We observed the previously reported trend towards underestimation of distance perception in OST AR, however, given the high variance levels of our data, more experiments are needed to confirm this result. Moreover, our matching task may not be the best to evaluate egocentric distance perception in OST AR and open-loop judgmental tasks could be used (e.g. blind walking, pointing, etc.).

Virtual shadows can improve distance perception accuracy in OST AR. Interestingly, the level of realism of virtual objects' shadows does not have an influence on distance perception in OST AR. It is thus sufficient to use "cheap" and easy to compute shadows instead of expensive realistic

ones.

Lighting misalignment between real and virtual environments negatively impacts distance perception in OST AR. Interestingly, putting the virtual lights closer than the real ones seems to lead to better results than when putting them farther away.

Relative errors on perceived distances as well as completion time for each trial tend to correspond to higher values under the condition of farther targets' distances.

With the recent advances in commercial AR HMDs technologies, we believe it is of tremendous importance to address potential AR-related perceptual biases. By exploring the influence of virtual objects shadows and lighting coherence in OST AR, our results contribute to better understand the complex phenomenon of how people perceive egocentric distances with virtual objects.

In order to further study this phenomenon, we envision different future leads such as: studying the influence of how shadows are rendered in OST AR (e.g. increasing the bright- ness of the rest of the scene to make them more noticeable); comparing the impact of using "drop shadows" vs. realistic shadows on distance perception accuracy (this was not directly compared in [6] since they used drop shadows and "ray-traced" shadows in two separate experiments); using an open-loop judgment task (e.g. blind walking, pointing, virtual throwing, etc.), that may give better results for egocentric distance estimations but is more complex to use in a controlled experiment; studying the influence of shadows on the intrinsic characteristics of virtual objects (e.g. does the size of the shadow influence the perceived size, weight, etc. of the virtual object). Moreover, as for the experimental design in future studies, a comparative experiment could be added to explore the distance perception of matching the real objects with the targets through watching AR HMDs. This may provide us some comparable data, and help us isolating the factors leading to egocentric distance errors in OST AR.

Results from such experiments could help define some guidelines when designing OST AR applications which may become more and more accessible in a near future.

Reference:

[1] Bimber O, Raskar R. Spatial Augmented Reality: Merging Real and Virtual Worlds. Natick, MA, USA: A. K. Peters, Ltd.; 2005. DOI: 10.1201/b10624

[2] Renner RS, Velichkovsky BM, Helmert JR. The Perception of Egocentric Distances in Virtual Environments - A Review. ACM Comput Surv. 2013 Dec; 46(2):23:1–23:40. DOI: 10.1145/2543581.

[3] Naceri A, Chellali R, Hoinville T. Depth Perception Within Peripersonal Space Using Head-Mounted Display. Presence: Teleoper Virtual Environ. 2011; 20(3):254–272. DOI: 10.1162/PRES_a_00048

[4] Interrante V, Ries B, Lindquist J, Kaeding M, Anderson L. Elucidating Factors that Can Facilitate Veridical Spatial Perception in Immersive Virtual Environments. Presence: Teleoper Virtual Environ. 2008; 17(2):176–198. DOI: 10.1162/pres.17.2.176

[5] Interrante V, Kaeding M, Ries B, Anderson L. Correlations Between Physiological Response, Gait, Personality, and Presence in Immersive Virtual Environments. Presence: Teleoper Virtual Environ. 2014;21(2):119–141.

[6] Diaz C, Walker M, Albers Szafir D, Szafir D. Designing for Depth Perceptions in Augmented Reality. In: Proceedings of the 2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR); 2017. p. 111–122. DOI: 10.1109/ISMAR.2017.28

[7] Rolland JP, Gibson W, Ariely D. Towards Quantifying Depth and Size Perception in Virtual Environments. Presence: Teleoper Virtual Environ. 1995 Jan; 4(1):24–49. DOI: 10.1162/pres.1995.4.1.

[8] Willemsen P, Colton MB, Creem-Regehr SH, Thompson WB. The Effects of Head-mounted Display Mechanical Properties and Field of View on Distance Judgments in Virtual Environments. ACM Trans Appl Percept. 2009 Mar; 6(2):8:1-8:14. DOI: 10.1145/1498700.1498702

[9] Kelly JW, Cherep LA, Siegel ZD. Perceived Space in the HTC Vive. ACM Trans Appl Percept. 2017 Jul;15(1):2:1–2:16. DOI:10.1145/310615

[10] Buck LE, Young MK, Bodenheimer B. A Comparison of Distance Estimation in HMD-Based Virtual Environments with Different HMD-Based Conditions. ACM Transactions on Applied Perception (TAP). 2018;15(3):21. DOI: 10.1145/3196885

[11] Li B, Nordman A, Walker J, Kuhl SA. The Effects of Artificially Reduced Field of View and Peripheral Frame Stimulation on Distance Judgments in HMDs. In: Proceedings of the ACM Symposium on Applied Perception. SAP ’16. New York, NY, USA: ACM; 2016. p. 53–56. DOI: 10.1145/2931002.2931013

[12] J Adam J, J Edward S, Mark B. Peripheral stimulation and its effect on perceived spatial scale in virtual

environments. IEEE Trans Vis Comput Graph. 2013;19(4):701–710.

[13] Waller D, Richardson AR. Correcting distance estimates by interacting with immersive virtual environments: Effects of task and available sensory information. J Exp Psychol Appl. 2008;14(1):61–72. DOI: 10.1037/1076-898X.14.1.61

[14] Jones JA, Krum DM, Bolas MT. Vertical Field-of-View Extension and Walking Characteristics in Head-Worn Virtual Environments. ACM Trans Appl Percept. 2016 Oct; 14(2):9:1–9:17. DOI: 10.1145/2983631

[15] Rolland JP, Meyer C, Arthur K, Rinalducci E. Method of Adjustments Versus Method of Constant Stimuli in the Quantification of Accuracy and Precision of Rendered Depth in Head-mounted Displays. Presence: Teleoper Virtual Environ. 2002 Dec;11(6):610–625. DOI: 10.1162/105474602321050730

[16] Swan II JE, Jones A, Kolstad E, Livingston MA, Smallman HS. Egocentric Depth Judgments in Optical, See-Through Augmented Reality. IEEE Transactions on Visualization and Computer Graphics. 2007 May;13(3):429–442. DOI: 10.1109/TVCG.2007.1035

[17] Livingston MA, Ai Z, Swan JE, Smallman HS. Indoor vs. Outdoor Depth Perception for Mobile Augmented Reality. In: Proceedings of the 2009 IEEE Virtual Reality Conference; 2009. p. 55–62. DOI: 10.1109/VR.2009.4810999

[18] Jones A, Swan JE, Singh G, Kolstad E. The Effects of Virtual Reality, Augmented Reality, and Motion Parallax on Egocentric Depth Perception. In: Proceedings of the 2008 IEEE Virtual Reality Conference. VR 2008; 2008. p. 267–268. DOI: 10.1109/VR.2008.4480794

[19] Ellis SR, Bucher UJ, Menges BM. The Relationship of Binocular Convergence and Errors in Judged Distance to Virtual Objects. IFAC Proceedings Volumes. 1995;28(15):253–257. 6th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design and Evaluation of Man-Machine Systems 1995, Cambridge, MA, USA, 27-29 June 1995. DOI: 10.1016/S1474-6670(17)45241-4

[20] Hua H, Javidi B. A 3D integral imaging optical see-through head-mounted display. Opt Express. 2014 Jun;22(11):13484–13491. DOI: 10.1364/OE.22.013484

[21] Swan JE, Singh G, Ellis SR. Matching and Reaching Depth Judgments with Real and Augmented Reality Targets. IEEE Transactions on Visualization and Computer Graphics. 2015 Nov;21(11):1289– 1298. DOI: 10.1109/TVCG.2015.2459895

[22] Singh G, Ellis SR, Swan II JE. The Effect of Focal Distance, Age, and Brightness on Near-Field Augmented Reality Depth Matching. IEEE Transactions on Visualization and Computer Graphics. 2018;p. 1–1. DOI: 10.1109/TVCG.2018.2869729

[23] Singh G, Swan JE II, Jones JA, Ellis SR. Depth Judgment Measures and Occluding Surfaces in Near-field Augmented Reality. In: Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization. APGV ’10. New York, NY, USA: ACM; 2010. p. 149–156. DOI: 10.1145/1836248.1836277

[24] Messing R, Durgin FH. Distance Perception and the Visual Horizon in Head-Mounted Displays. ACM Trans Appl Percept. 2005 Jul;2(3):234–250. DOI: 10.1145/1077399.1077403

[25] Mamassian P, Knill DC, Kersten D. The perception of cast shadows. Trends in Cognitive Sciences. 1998;2(8):288–295. DOI: 10.1016/S1364-6613(98)01204-2

[26] Hubona GS, Wheeler PN, Shirah GW, Brandt M. The Relative Contributions of Stereo, Lighting, and Background Scenes in Promoting 3D Depth Visualization. ACM Trans Comput-Hum Interact. 1999 Sep;6(3):214–242. DOI: 10.1145/329693.329695

[27] Sugano N, Kato H, Tachibana K. The effects of shadow representation of virtual objects in augmented reality. In: Proceedings of the Second IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR); 2003. p. 76–83. DOI: 10.1109/ISMAR.2003.1240690

[28] Cutting, J. E. (2003). Reconceiving perceptual space. In H. Hecht, R. Schwartz, & M. Atherton (Eds.), Looking into pictures: An interdisciplinary approach to pictorial space (pp. 215-238). Cambridge, MA, US: MIT Press.

[29] Kress BC, Cummings WJ. 11-1: Invited Paper: Towards the Ultimate Mixed Reality Experience: HoloLens Display Architecture Choices[C]// Sid Symposium Digest of Technical Papers. 2017:127-131. DOI: 10.1002/SDTP.11586.

[30] Hardy LH, Rand G, Rittler MC. The Ishihara test as a means of detecting and analyzing color vision. Journal of General Psychology. 1947;36(2):79. DOI: 10.1080/00221309.1947.9918108

[31] Yonas A, Goldsmith LT, Hallstrom JL. Development of Sensitivity to Information Provided by Cast Shadows in Pictures. Perception. 1978;7(3):333–341. DOI: 10.1068/p070333

[32] Thompson W, Fleming R, Creem-Regehr S, Stefanucci JK. Visual Perception from a Computer Graphics Perspective. 1st ed. Natick, MA, USA: A. K. Peters, Ltd.; 2011.