Technical Section

Shadow induction on optical see-through head-mounted displays

Sei Ikeda [a,b,∗], Yuto Kimura [b], Shinnosuke Manabe [b], Asako Kimura [b], Fumihisa Shibata [b]

[a] *Osaka University, Machikaneyama 1–3, Toyonaka, Osaka 560–8531 Japan*
[b] *Ritsumeikan University, Nojihigashi 1-1-1, Kusatsu, 525–8577 Shiga, Japan*

ARTICLE INFO

ABSTRACT

In this paper, we present two effects of illusion-based virtual shadows on a real surface for optical see-through head-mounted displays without devices that selectively reduce or intercept light rays from the real scene. The virtual shadows are based on the simultaneous brightness-contrast illusion in which a central region surrounded by a lighter one is perceived as darker than one without such a lighter region. To display a virtual shadow, a virtual object with a real image texture is placed around the shadow region as the lighter region to provide a sense of a darker area. We conducted two kinds of experiments to validate the feasibility of our method. The first experiment showed that users tend to perceive the virtual shadow part as darker than the surrounding area with the same physical luminance. The second experiment provided evidence that there are cases in which the virtual shadows can improve the reality of virtual objects.

© 2020 Elsevier Ltd. All rights reserved.

1. Introduction

Optical see-through head-mounted displays (OST-HMDs) are the most promising displays for augmented reality (AR) in the sense that the optical element is already thinned to the same level as ordinary glasses [1]. As far as we know, there is no example of the same level of thinning that has yet been achieved with video see-through HMDs. However, such AR glasses have no prospect at all for the means to attenuate each light ray from the real scene (Section 2.1). The inability to physically control the amount of the incident light reaching the retina means that the virtual objects occluding or shadowing real objects cannot be represented.

The aim of this paper is to show the possibility of relaxing one of the limitations by using visual effects. Specifically, we demonstrate the feasibility of our virtual shadow representation method using brightness contrast on OST-HMDs. Fig. 1 shows examples of the virtual shadow representation. In this method, a bright virtual object (hereinafter referred to as *the shadow inducer*) is placed around the area to be shadowed (named *the shadow region*) in order to induce the brightness contrast. The shadow inducer amplifies the luminance of the real surface so that the boundary between the shadow region and the shadow inducer becomes a strong edge. The outer edge of the shadow inducer is weakened gradually to make it more difficult to detect the luminance amplification.

In the experiments, we show the feasibility of our method from the following two perspectives. The first experiment was designed based on a psychophysical aspect. We have demonstrated that the shadow areas are perceived darker than the original texture. The second experiment was designed from another perspective concerned with the quality of the virtual shadows represented by the shadow inducers. We demonstrated that there are cases where the shadow inducers can improve the perception of a material of virtual objects. We chose transparent objects as virtual objects which are required to render caustics. The results show that it is possible to not only add a virtual darker region than the real environment, but also combine brighter regions with darker regions. Through the above two experiments, it is shown that our method is applicable for both opaque and transparent virtual objects.

2. Related work

2.1. Occlusion-Capable OST-HMDs

The only means to manipulate pixel-wise transmittance of display devices is to use spatial light modulators (SLMs), which are electrically programmable devices to modulate light pixel by pixel such as in liquid crystal displays (LCDs), digital micromirror devices (DMDs) and liquid crystal on silicon (LCOS) displays. HMDs with SLMs can be potentially applicable for the representation of occlusions of real objects and virtual shadows on real surfaces. To our best knowledge, the earliest occlusion capable HMDs using LCDs were ELMO-1 through ELMO-4, whose sizes were larger than human faces, developed by Kiyokawa et al. [2,3]. To avoid blurring

(a) A virtual plastic bottle on a real stone floor.　　　(a) A virtual wine glass on a real newspaper.　　　(a) A virtual glass dish on a real Moroccan tiled floor.
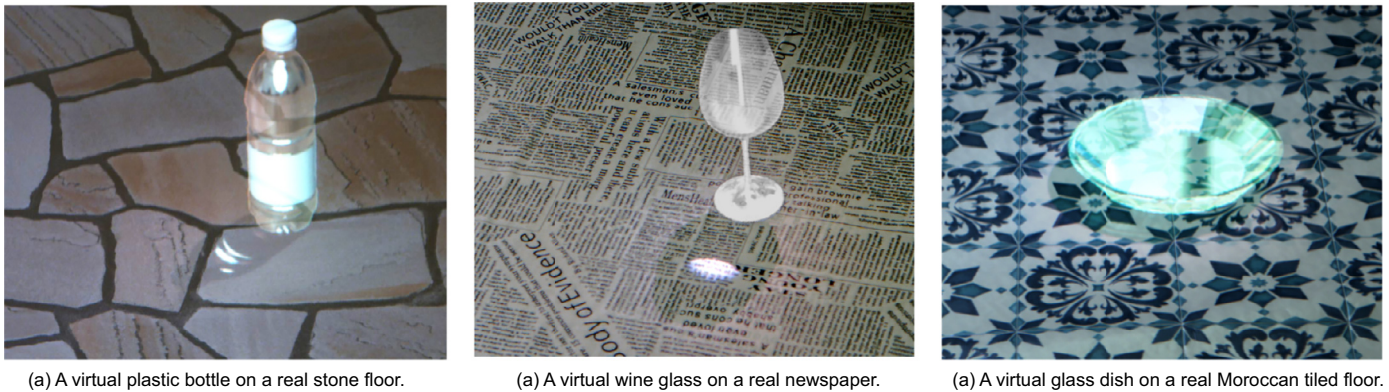
Fig. 1. Examples of augmented transparent objects with virtual shadows. All the images were captured through an OST-HMD (Microsoft HoloLens). Each object at the center of each image was a virtual object placed on a real planar surface where there were not any shadows originally. To make observers perceive the shadow as darker than the real surface, a brighter virtual object (i.e., shadow inducer) with the same texture as the real surface was overlaid. (See Appendix for the acquisition method of all the real scene images containing virtual objects throughout this paper.).

of occlusion regions, the HMDs were designed to secure the focal length of the display by arranging multiple optical elements in a complicated manner.

Since then, several technologies using DMDs [4] and LCOS [5–7] have been proposed to solve the lack of transparency in LCDs. However, they initially had not yet reached a level to discuss miniaturization and safety. OST-HMDs were dramatically downsized after the introduction of computational photography and wave optics approaches [8–10].

Itoh et al. [11] have developed a method that can make the thinnest OST-HMDs with occlusion capability at present. They compensated the blur of the occlusion regions by overlaying real images to the display without taking the approach of keeping the optical path long. Just as in their approach, our method also augments real image textures taken from a user-perspective camera. As described above, the thinning of optical elements in OST-HMDs is certainly progressing. However, the use of such SLMs means that the user's view may be lost in the event of a malfunction or failure. In this paper, we assume AR glasses without such SLMs.

2.2. Brightness induction

Our method is based on brightness induction. A large number of studies on brightness induction have been done in the field of human vision and neuroscience. The purpose of our research is neither to investigate the conditions under which such brightness induction occurs itself nor to elucidate the human vision system. Here, we pick up and introduce some research strongly relevant to our method or experiment.

Brightness induction is a general term for the phenomena that the perceived brightness of a region is modulated by the surrounding pattern. By making the surrounding area brighter or darker, it is possible to make the perceived brightness appear darker or lighter, respectively, than the original luminance [12]. This effect has been confirmed not only in experimental environments but also in real environments [13,14]. The target region should be close to the surrounding one [15–17], and smaller than the surrounding region [18]. The strength of this effect depends on the pattern of the surrounding region [17,19,20].

2.3. Shadow representation using brightness illusions

In spatial AR, researchers have proposed an approach to represent virtual shadows casting on real surfaces based on decreasing the relative brightness of the shadow region by brightening the

whole field of view (FOV) [21,22]. The main common point to this work is the use of the perceptual phenomenon of brightness.

Similar concepts has been patented [23,24]. They proposed similar methods for representing shadows by projecting a spotlight-like pattern around a virtual shadow region in spatial augmented reality in the same way as the method used in the current paper. Unlike the work just described [21,22], a gradation was added according to the distance from the shadow area instead of lightening the whole environment. Their patent documents, however, do not provide any experimental evidence or any parameters that were confirmed to be effective. [25] have shown similar effects, and then Kimura et al. [26] and Manabe et al. [27] have exhibited real-time demonstrations of similar shadow representations using OST-HMDs. In contrast to the patents, the goal of the current work is to explain these experiments in detail and to show the evidence of the effects of our shadow inducers.

3. Design of the shadow inducers

This section first describes the photometric model for inserting a shadow inducer. We clarify the conditions under which shadow inducers can be synthesized using real images of a user-perspective camera. Then, we also explain its implementation method using the model.

3.1. Photometric model around the boundary

The shadow inducers that create the illusion of virtual shadows do not simulate the physical phenomena themselves. With that being said, there is a risk that the reality decreases if the area around the shadow region is recklessly brightened. Therefore, we avoid a decrease in reality as much as possible by locally reproducing the optical phenomenon.

Specifically, a shadow inducer is presented by rendering as if the intensity of the existing light source was partially high. This operation is not difficult if we know or can estimate the three-dimensional geometry, the surface reflectance [28], and the lighting conditions (i.e., the intensity and distribution of the light sources) [29] of the real scene. As a frequently encountered condition in practical use, we introduce the following assumptions. First, the scene geometry is a plane. Second, the light distributions can be approximated by a single point light source and ambient light. The position and strength of the light source are known. The last assumption is about the shadowing model described as follows.
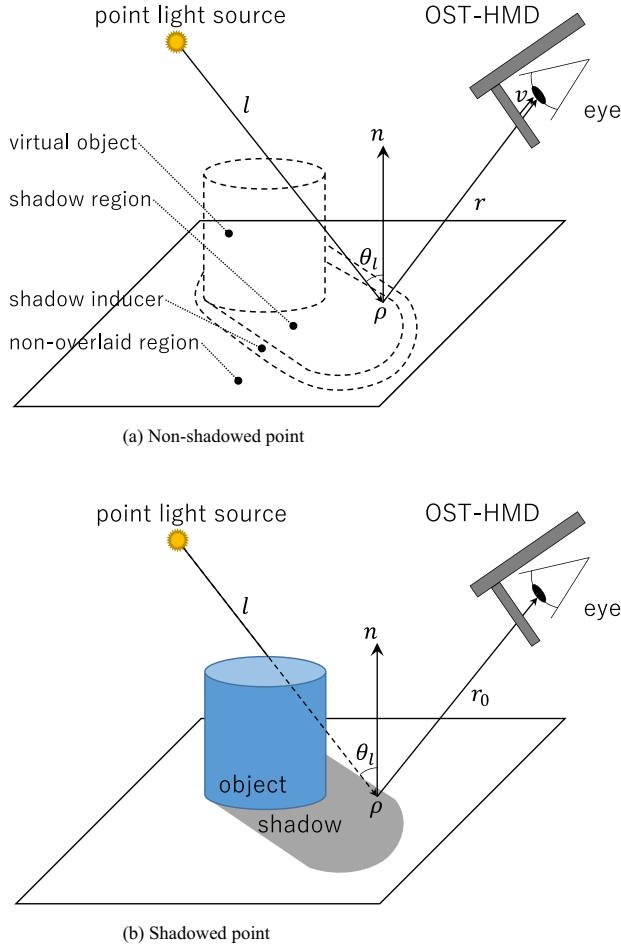
(a) Non-shadowed point

(b) Shadowed point

Fig. 2. Shadowing models. The luminance amplification coefficient at the boundary point between the shadow inducer and shadow region can be calculated as a ratio of the luminance values of shadowed (a) and non-shadowed (b) points.

We adopted a shadowing model where the strength of reflection r is proportional to the strength of the point light source l. This assumption is followed by most reflection models, including the bidirectional reflectance distribution function (BRDF) model. Fig. 2 (a) shows the illumination of a non-shadowed point. In this situation, the radiance r of the reflection is represented by using the irradiance l from the point light source and the ambient light term a as follows:

$$r = f(\omega_l, \omega_r)\, l + a, \tag{1}$$

where $f(\omega_l, \omega_r)$ is the BRDF of the light vector ω_l and the eye vector ω_r. In general, the ambient light is treated as a constant term. Considering that the single light source illuminates the atmosphere and surrounding environment and then provides illuminance that does not depend on the viewpoint or the light source position, the ambient factor can be considered to be proportional to the radiance l and reflection coefficient $\rho = \int f(\omega_l, \omega_r)\, n \cdot \omega_l\, d\omega_l$ as expressed by $a = \rho k l$, where k is its constant of proportion and n is the surface normal.

For a shadowed point, a direct illumination from the point light source does not exist, as shown in Fig. 2 (b). In this case, the reflected light r_0 is equal to only the ambient term:

$$r_0 = a. \tag{2}$$

In our method, the luminance around the shadow is locally amplified with the ratio α_0 of the luminance of the shadow region to

the non-shadow one as follows:

$$\alpha_0 = \frac{r}{r_0} = \frac{f(\omega_l, \omega_r) + \rho k}{\rho k}. \tag{3}$$

As shown in Fig. 2 (a), the light reaching the retina is the sum of the light r from the real scene and the light v from the virtual object, $r + v$, as presented on the HMD. The luminance of the shadow inducer near the shadow boundary should be displayed so that $r + v = \alpha_0 r$. On the HMD, an image with the luminance of $v = (\alpha_0 - 1)r$ is displayed. According to this equation, we require the BRDF $f(\omega_l, \omega_r)$ and the proportion coefficient k as parameters of the real environment.

3.2. Practical implementation of the shadowing inducers

Here, we explain a more practical model, which requires additional assumptions. It is the method we actually used in the experiments described in Sections 4 and 5. The image acquired from a user-perspective camera is a record of the radiance from the real scene without a shadow. In order to use this image for other viewpoints, we can additionally assume the Lambertian reflection, $f(\omega_l, \omega_r) = \rho_d/\pi \cos\theta_l$, with a diffuse coefficient ρ_d for ease of implementation, where θ_l is the zenith angle of the light vector ω_l. We deal with this image as the texture of a shadow inducer. Specifically, the user perspective image is multiplied by the ratio α_0 defined as follows:

$$\alpha_0 = \frac{r}{r_0} = \frac{\rho_d/\pi \cos\theta_l + k\rho_d}{k\rho_d} = \frac{\cos\theta_l + k\pi}{k\pi}. \tag{4}$$

We saw that the number of the parameters can be reduced by applying the above practical assumptions. The remaining parameters, i.e., the light source direction θ_l and the ambient coefficient k, must be estimated by sensing the real environment or should be determined empirically. For the experiments in this paper, we manually set the 3-D position of the point light source and computed the shadow regions by the traditional shadow mapping algorithm [30]. The coefficient k was empirically chosen. We tuned both parameters while looking at the rendering results.

Finally, we can explain how the luminance of the shadow inducer at the boundary was set. The amplification coefficient α of luminance at a point P is determined by the following linear gradation function depending on the distance d from P to the closest shadow region point Q:

$$\alpha(d) = \alpha_0 \frac{D - d}{D},$$

where D is the distance from Q to the outer edge point of the shadow inducer on the extension of the line segment $\overline{QP}$.

Assuming the Lambertian surface, the environment can be treated as a static object. In each of the following experiments, we created a shadow inducer generated from the acquired image using the method described above, and placed the shadow inducer in Unity as a textured planar object.

4. Experiment 1: Perceived brightness

The results presented here demonstrate that users perceived the area of the virtual shadow as dark. This means that the shadow area was perceived relatively darker than not only the shadow inducer region but also than a non-overlaid region with the same luminance as the shadow region. The hypothesis of this experiment was that users would tend to perceive the shadow region to be darker than a non-overlaid region.
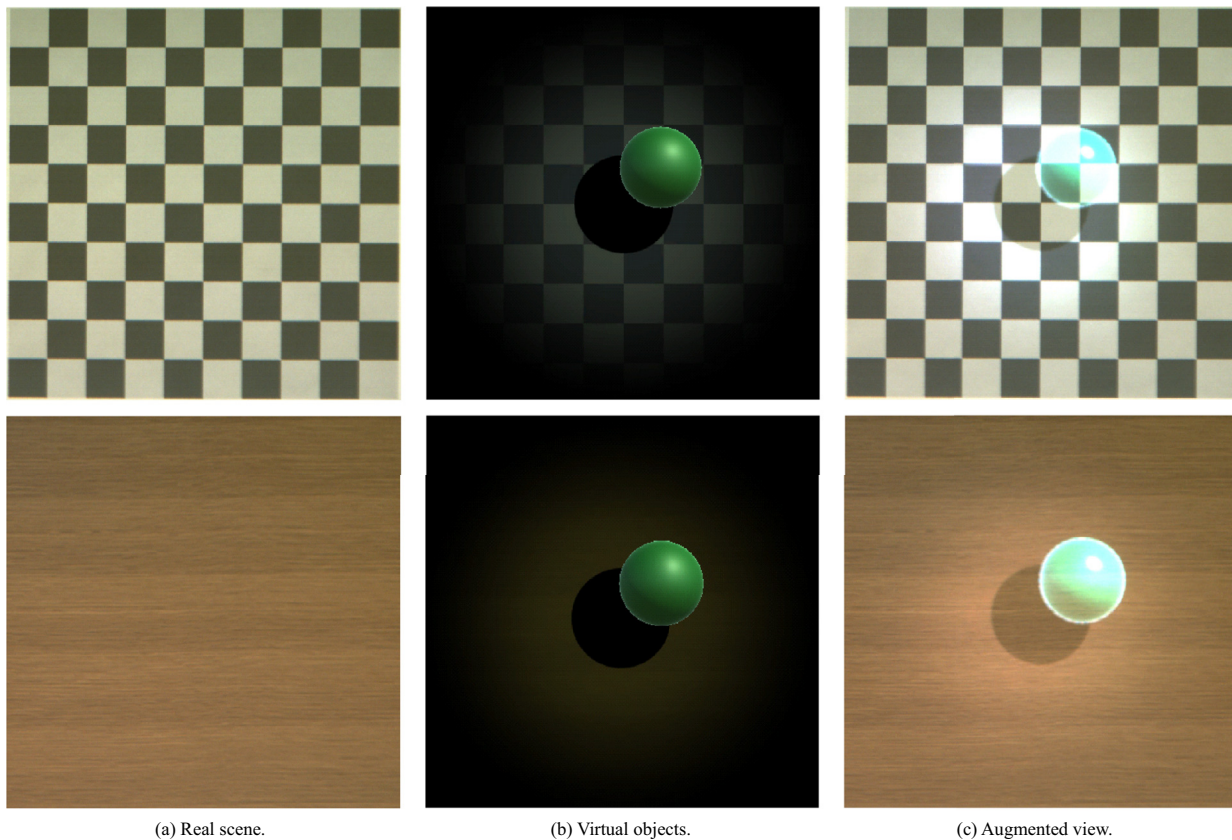
| (a) Real scene. | (b) Virtual objects. | (c) Augmented view. |

Fig. 3. Virtual shadows used in the experiment. The real scene (a) and the augmented view (c) were captured through the HoloLens without and with the virtual object images (b), respectively. Note that the shadow inducers in (c) may be affected by the color profile of the display or printer used to view them.

4.1. Visual stimuli

Our experimental design was inspired by "Checker Shadow Illusion," a computer-generated figure published on the Web by Adelson [31]. In this figure, there is a cylinder on a floor with a checkerboard pattern; a soft shadow of the cylinder is cast on the floor. In this illusion, observers compare two points *A* and *B*, which are placed at the center of gray and black checks, respectively. The gray check *A* is shadowed by the cylinder, but the black one is not. Most observers of this illusion answer that the gray check is brighter than the black one, when in fact, the two luminance values are actually the same. Our visual stimuli and procedure basically followed the manner of demonstrating the checker shadow illusion.

The major differences of our experiment from the checker shadow illusion were as follows:

- Three conditions were prepared.
- The scene containing a shadow was represented by the shadow inducer and presented through an OST-HMD.
- The observers did not directly compare the two points with the same luminance, but rather rated brightness magnitudes compared with a reference point.

In addition to these differences, we made some changes to the floor and the object according to the purpose of this experiment, as shown in Fig. 3. First, the object casting the shadow was not a cylinder but a sphere and was made so as to appear to float above the background, because the shadow was required to be small and visible from the observer's viewpoints. Second, the background grid was set larger than the original one and changed to
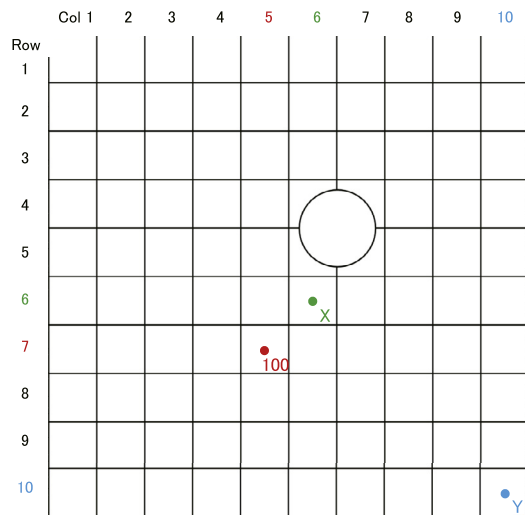
10×10 because the space for the shadow inducer was required in addition to the shadow and non-shadow regions.

We considered that some similar brightness induction was likely to occur even on OST-HMDs if the conditions were close to the checkerboard illusion. In contrast, such illusions would be less likely to occur if the conditions were far from the original one. Therefore, in order to set a stage that gradually varied from the original, we prepared the following three conditions:
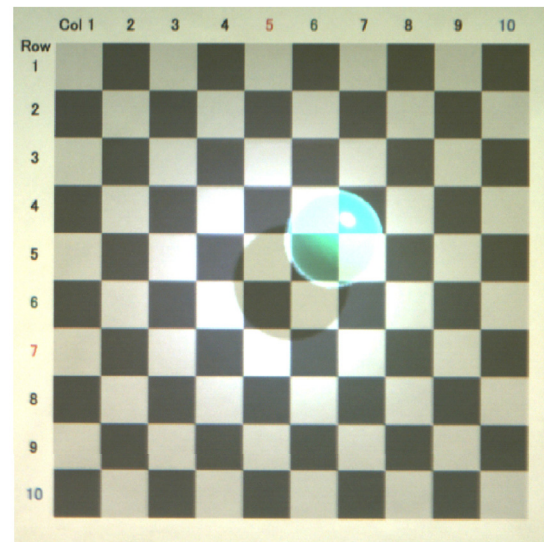
Condition 1: As a visual stimulus, we used a camera-captured image taken through the OST-HMD that showed a green sphere over a real checkerboard-patterned plane, as shown in the upper row of Fig. 3 (c). Each participant observed this image on their own individual desktop display. Therefore, the resolution, view angle, and monitor color settings were not controlled.

Condition 2: Each participant observed the same virtual objects and background as in Condition 1 but through the OST-HMD (and not on a desktop display) so that participant could slightly change the viewpoint from a standing point by moving their head left or right. The height of the physical checkerboard background was adjusted according to each participant's height.

Condition 3: We presented the same sphere over a wood grain surface as a visual stimulus in this condition, as shown in the lower row of Fig. 3 (c). As in Condition 2, each participant observed this through the OST-HMD from free viewpoints standing at the same position. Unlike the checkerboard pattern, it was more difficult to indicate a certain cell with this surface.

(a) Illustration for instruction.

(b) Actual view in evaluation.

Fig. 4. Illustration and actual view in Conditions 1 & 2. The participants were asked to observe the points in the illustration (a) before evaluating the brightness values in the actual view (b).

The initial geometric registration was performed manually. For each participant, the position of the virtual objects was refined by keyboard input before the instruction. While the participants were observing the target, we used frame-by-frame tracking on the HoloLens to maintain the initial precise registration. The photometric calibration was performed based on the linear relationships between the image intensity and illuminance of the camera, and between the output image intensity and luminance of the OST-HMD. The parameters for the linear gradation, α_0 and D, were set as 0.4 and 140 mm, respectively.

4.2. Participants & procedures

Condition 1. We had 23 unpaid participants (aged 21 to 26) recruited from a local university. All participants had normal or corrected to normal vision.

In this condition, each participant completed a questionnaire using Google Forms. After providing some basic information, each participant was then asked to carefully look at the illustration shown in Fig. 4 (a) and memorize the following three kinds of information:

- The position of the virtual object on the grid,
- The point with a red label 100 as a reference,
- Two points with a green label X and a blue label Y, whose brightness participants should also evaluate.

Lastly, each participant was asked to look at the augmented image shown in Fig. 4 (b) and rate the perceived brightness of X and Y, assuming the brightness of the reference was 100. Then, the participant asked to input two magnitudes of their perceived brightness.

Condition 2. In this condition, 23 participants (aged 21 to 26) who had normal or corrected to normal vision were recruited in the same way as the previous condition (20 of those participated in the experiment of Condition 1)[1].

In this condition, everything was explained to the participants verbally. We showed each participant the same illustration shown in Fig. 4 (a), and asked to memorize the positions of X, Y, and the reference in the same manner as Condition 1. After that, we asked each participant to stand in front of the checkerboard as shown in Fig. 5 and to rate the perceived brightness of X and Y.

Condition 3. For this condition, 23 participants (aged 21 to 26) who had normal or corrected to normal vision were recruited in the same way. The same 20 participants also were in this condition from the two earlier ones; two of the others participated previously only in Condition 2; the other one had participated only in Condition 1. The difference from Condition 2 is that we used a wooden surface as a background. Unlike Conditions 1 and 2 that used a checkerboard pattern, it was difficult to indicate the positions where the participants should look. We displayed triangle indicators, shown in Fig. 6 (b). Participants recognized the target points from the intersection of two straight lines connecting the paired triangles with the same color. Before presenting this visual stimulus, the participants were asked to look at an illustration shown in Fig. 6 (a) and memorize the positions in a similar manner as in Conditions 1 and 2. Evaluating the brightness and the target positions were performed in the same manner as in Conditions 1 and 2 (except for the differences in identifying the targets as previously described). Since there was almost no expected order effect in the preliminary experiments, the order of experiencing the three conditions was not counterbalanced. We gave time intervals of more than 24 hours between two different conditions so that order effects were unlikely to occur.

For quantitative evaluations, we used an average of the magnitude values and a dimming rate of the shadow region X to the non-overlaid one Y. Before the analyses, we applied the Shapiro-Wilk method [32]. The magnitude values of X and Y in all the conditions and the dimming rates of Condition 1 and 3 were not rejected ($p > .05$), only the dimming rates of Condition 2 were not normally distributed ($p < .05$). Therefore, in the comparison between X and Y, we applied the paired *t*-test. For the comparison of three dimming rates, we used the Steel-Dwass method [33,34], which is a non-parametric test assuming general distributions.
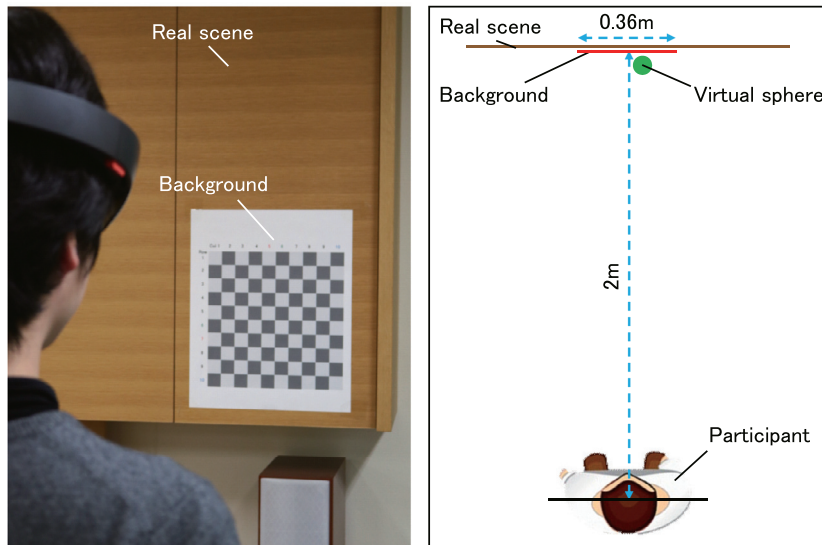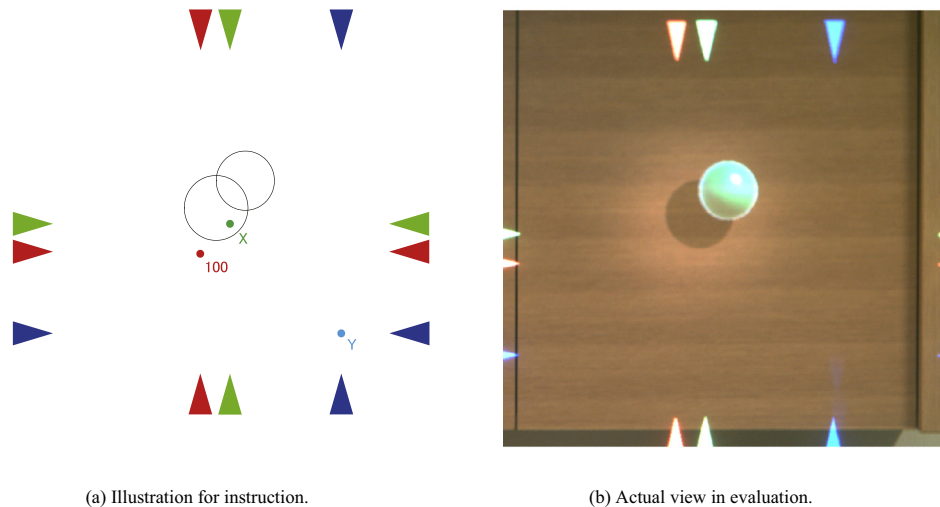
---

[1] Initially, we attempted to perform the three conditions of Experiment 1 with the same set of participants because of the insufficient number of participants available. After conducting Condition 1, three of them were unable to participate. Therefore, the additional participants were recruited in Condition 2 and 3.

Fig. 5. The experimental setup for Conditions 2 and 3. The checkerboard sheet was used only for Condition 2. The height of the sheet was adjusted according to the height of the participant's head.

(a) Illustration for instruction.

(b) Actual view in evaluation.

Fig. 6. Illustration and actual view in Condition 3. We used the triangle indicators to specify the positions of the evaluation points without affecting the brightness perception around them. Each participant was asked to observe and remember the evaluation points in the illustration (a) before looking at the actual view (b) for evaluation.

4.3. Results & discussion

Fig. 7 (a) shows the average magnitude values of the points X and Y in each condition. In every condition, the perceived brightness of the point X in the shadow region was significantly darker than that of the point Y in the non-overlaid region. Fig. 7 (b) shows the average dimming rate of the point X relative to the point Y in each condition. From this figure, the dimming rate in Condition 3 was significantly lower than that in both Conditions 1 and 2 although we did not find a significant difference between Conditions 1 and 2.

From the results in Fig. 7 (a), it can be said that users tend to perceive the shadow region inside the shadow inducer as darker than the non-overlaid region even though no virtual objects were overlaid on either region. According to the fact that the luminance of the non-overlaid region was the same as the original one, it is natural to consider that users tend to perceive the shadow region as darker than the original brightness of the real scene even on OST-HMDs. On the other hand, as in the case of the Craik-O'Brien-Cornsweet illusion, the brightness of the side with the gradually decreasing luminance can appear higher than the other side. Judging from the experimental results, it is not unthinkable that the reason for the difference in brightness between X and Y was because observers felt the non-overlaid region including the point Y was brighter than the original. If so, there must have been a difference in perceived brightness between the non-overlaid region and the peripheral FOV where no computer graphics can be presented. However, there were no data from our study that specifically addressed this. We may, therefore, reasonably conclude that users perceived the shadow region inside the shadow inducer as darker than the non-overlaid region.

One natural explanation for the results in Fig. 7 (b) may be that the dimming rate of a virtual shadow relative to the original depends on the background pattern. This is supported by previous work showing that the brightness contrast occurs more strongly when the surrounding inducer region has complex patterns [17,19,20] rather than homogeneous patterns. It is necessary to check the dimming rate in advance or to predict it with some visual model.
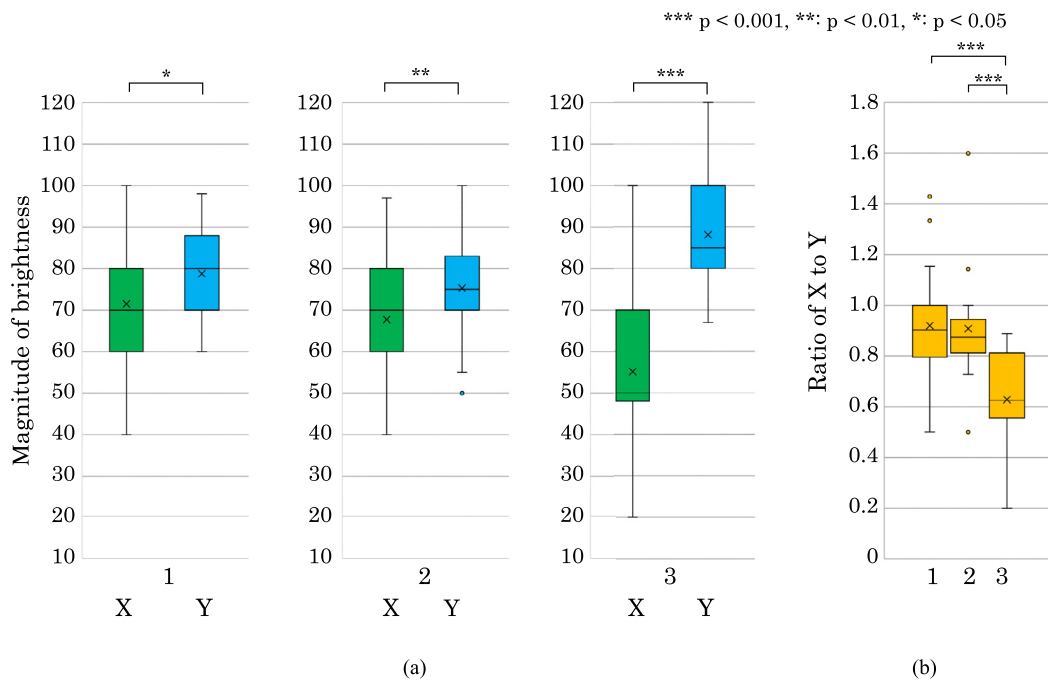
Fig. 7. Results of the experiment. (a) Average magnitudes of perceived brightness of the points *X* and *Y*. (b) Average dimming rates of perceived brightness. The bottom and top of the whiskers indicate the 5th and 95th percentiles, respectively. The bottom, middle, and top of the boxes represent the 1st quartile, the median, and the 3rd quartile, respectively.

5. Experiment 2: Reality

5.1. Overview

The purpose of this experiment was to confirm that the existence of a shadow created by the shadow inducer contributes the reality of a virtual transparent object. In addition to the dark area of the shadow, caustics also should be rendered to represent realistic shadows of transparent objects. Therefore, it was necessary to ensure that the factor that improves reality is not only the presence of caustics, but also shadows. In our experiment, the reality of the same virtual transparent object was compared under the following three conditions: (a) no shadow, (b) caustics only, and (c) caustics + shadow inducer.

5.2. Visual stimulus

The virtual transparent objects used in this experiment were a dense glass sphere and a diamond, shown in Figs. 8 and 9, respectively. The diameters of the sphere and diamond were both 80 mm. The refractive indexes of the glass and diamond were set to 1.51 and 2.42, respectively. The parameters for the shadow inducers were set as follows: $\alpha_0 = 0.3$ and $D = 140$ mm. The rendering methods of the body of the transparent object and the caustics part shown in Fig. 10 are described below.

Rendering of Transparent Objects. We used the Microsoft HoloLens as an OST-HMD and modeled the objects on Unity. A point light source was placed at $(x, y, z) = (0.0, 1.0, 1.0)$ [m] in the coordinate system whose origin was set at the round point of the virtual object, as shown in Fig 11. The specular reflection and refraction patterns were realized by the environment mapping of an omnidirectional image taken from an omnidirectional camera (Magicsee P3) placed at the center of the virtual objects. Both processes were performed by using a shader written in CG/HLSL in real time. Reflection and refraction were rendered by the cubemap and refract functions on Unity, respectively. We assumed that the light sources for the environment mapping were at infinity, and

both phenomena occurred only once on the virtual surface in order from the camera. The blending ratio of the reflection to the refraction was given as the Fresnel reflection coefficient calculated by Schlick's approximation.

Rendering of the Caustics. The caustics rendering was performed by the path-tracing method on Cycles Render of Blender, as shown in Fig. 10 (a). This rendering process was offline with the assumption of a static light source. The computation time of this rendering was approximately 20 minutes on a desktop PC (CPU: Intel Core i7-7700, GPU: Nvidia RTX 2070). Therefore, the participants could change their viewpoint but not move the objects or light source. The sampling number for the path-tracing was 40 rays per pixel, and the resolution of the texture was 1980 × 1980 pixels. Instead of measuring and using the physical luminance of the light source to generate the caustics, we subjectively adjusted the luminance so that the path-tracing results were full of texture on the HMD, because it was difficult to present the caustics with the real luminance due to the insufficient dynamic range of our HMD. The inducer was created based on the method described in Section 3. The real-world texture for the inducer was captured in advance with a separate DSLR camera (Canon EOS Kiss X8i) instead of the user-perspective camera.

5.3. Environment, procedure, & instructions

The design of the experimental environment is shown in Fig. 11. The lighting conditions were the same as in a typical office room. The distance between each participant and the virtual object was set as approximately 1,500 mm so that the virtual object appeared in the FOV of the HMD. We instructed each participant to stand at the designated position, and they were allowed to move their head freely within the natural range.

Each participant agreed with the purpose of this experiment and had a corrected visual acuity of 1.0 or more. The participants were 23 people in their 20s with no eye abnormalities. None had participated in Experiment 1. All the participants had experienced HoloLens.
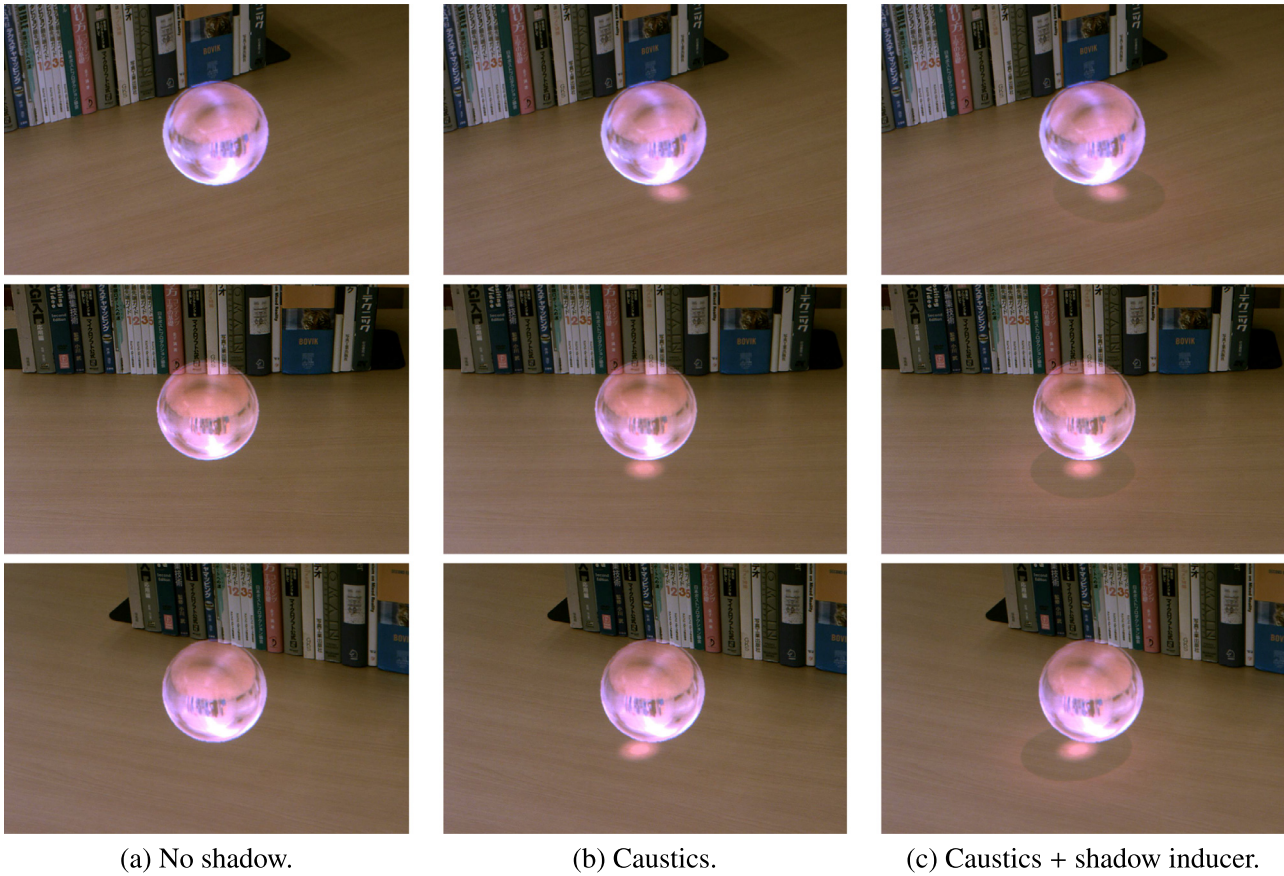
(a) No shadow.　　　(b) Caustics.　　　(c) Caustics + shadow inducer.

Fig. 8. Virtual glass sphere on a real wood grain face.

(a) No shadow.　　　(b) Caustics.　　　(c) Caustics + shadow inducer.

Fig. 9. Virtual diamond on real irregular stone tiles.

(a) Result of path-tracing.   (b) Caustics.   (c) Shadow inducer.   (d) Caustics + Shadow inducer.
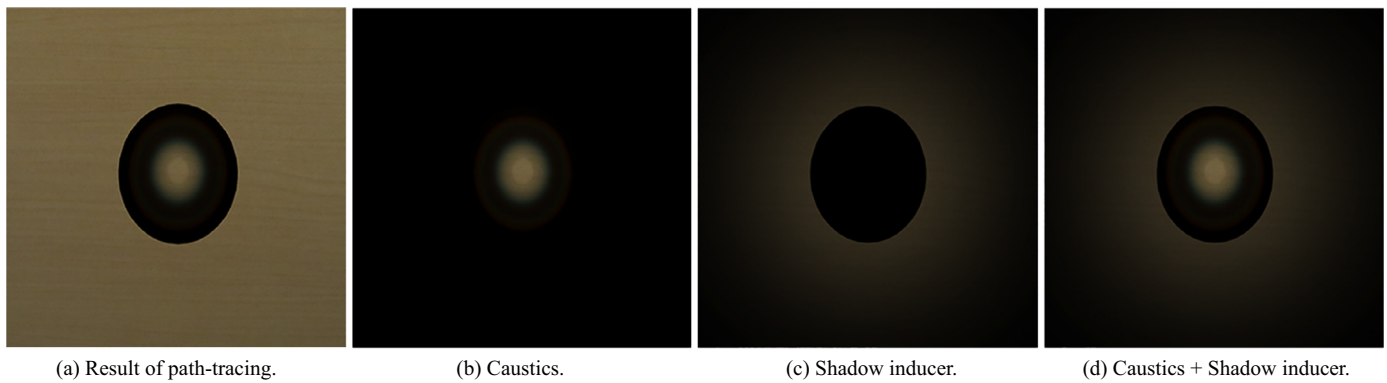
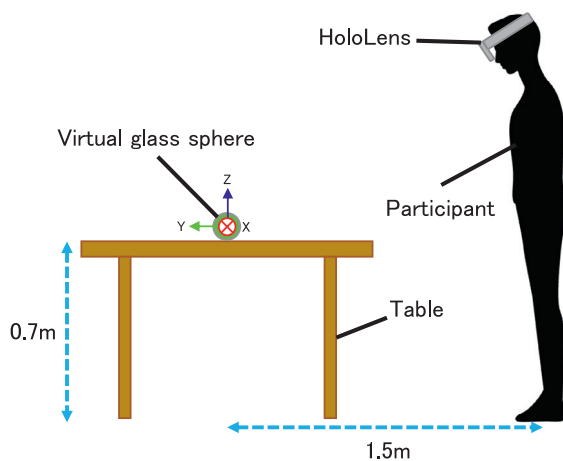**Fig. 10.** Rendering of the caustics and shadow inducer.

Fig. 11. Design of the experimental environment.

In order to inform the participants as to where the virtual objects appeared, we showed a white dummy cube at the same place in which the sphere/diamond would appear. After each participant found the cube, they compared every pair of the three conditions in randomized orders. They were allowed to move their heads freely to observe the virtual object, but were not allowed to walk or move from the designated position.

The procedure for comparing a pair of combinations was as follows. The first object appeared for 3.0 seconds, disappeared for 2.0 seconds, and then the second object appeared for 3.0 seconds. After observing each pair, each participant answered to the question: which object looked like a real dense glass sphere/diamond placed on a flat surface? The participant chose "the former" or "the latter" as the one that looked more realistic. Each participant compared six pairs in total, including reverse-ordered pairs to counterbalance any possible order effects.

In addition, we provided a small questionnaire to obtain open comments from the participants. We mapped the comparison results to the unified scale using Thurstone's paired comparison method (Case-V) and performed Mosteller's χ^2 test [35] of a significance level of 5%.

5.4. Results

Figs. 12 (a) and (b) show the results for the glass sphere and diamond, respectively. The horizontal axis indicates the psychological scale of the extent to which the virtual objects looked real. The significance of the results (a) and (b) were confirmed as ($\chi^2(1) = 4.10 \times 10^{-6}$, $p = .00$) and ($\chi^2(1) = 0.95 \times 10^{-6}$, $p = .00$), respec-

tively. Each graph illustrates the following findings: 1) that the virtual object with the caustics was more realistic than that without any shadow effects, and 2) that the virtual object with the caustics and shadow was more realistic than the others.

For the glass sphere, we obtained open comments from 16 out of 23 people. The comments by 13 of them were almost the same in that the object with both the caustics and shadow inducer appeared to be the most realistic. Two of the 16 answered that the shadow inducer looked unnaturally bright and did not look like that of a transparent object. One said that the caustics and shadow inducer did not look like a shadow that a transparent object creates.

For the diamond, 18 out of 23 people answered open comments. The comments by 14 of them were almost the same in that the object with both the caustics and shadow inducer appeared to be the most realistic. Fourteen participants stated that the virtual diamond with the shadow inducer looked like the most realistic one. Two answered that the shadow inducers were too conspicuous and did not look like a real shadow. One said he/she did not know how a diamond shadows on the ground. One selected the virtual object without the shadow every time because none of them looked a shadow.

5.5. Discussion

As shown in the results, we have confirmed shadow inducers have the effect of improving the reality of transparent objects. These results demonstrate there are at least a few cases where the shadow inducers can improve the reality of the virtual objects even though the objects must be rendered carefully photometrically. It is likely that the same effects can be achieved in more diverse environments and conditions, since we had not intentionally created unrealistic and rarely occurring experimental conditions. Kawabe's experiments showing that the presence of shadows improves the reality of transparent objects also supports our hypothesis [36]. Our results may potentially change the mind of people who think that realistic rendering is impossible in OST-HMDs without SLMs and OST-HMDs are only suitable for the applications using non-photorealistic virtual objects. This is not inconsistent with the fact that in the demonstration of our system in the international conference IEEE VR 2019[2], at least 90 out of approximately 130 participants commented that the reality was higher than they had imagined.

On the other hand, there were several comments that did not support our hypothesis, as mentioned above. Although the specific causes of these comments need to be further explored, there

---

[2] The 26th IEEE Conference on Virtual Reality and 3D User Interfaces, a top conference on virtual and augmented reality.
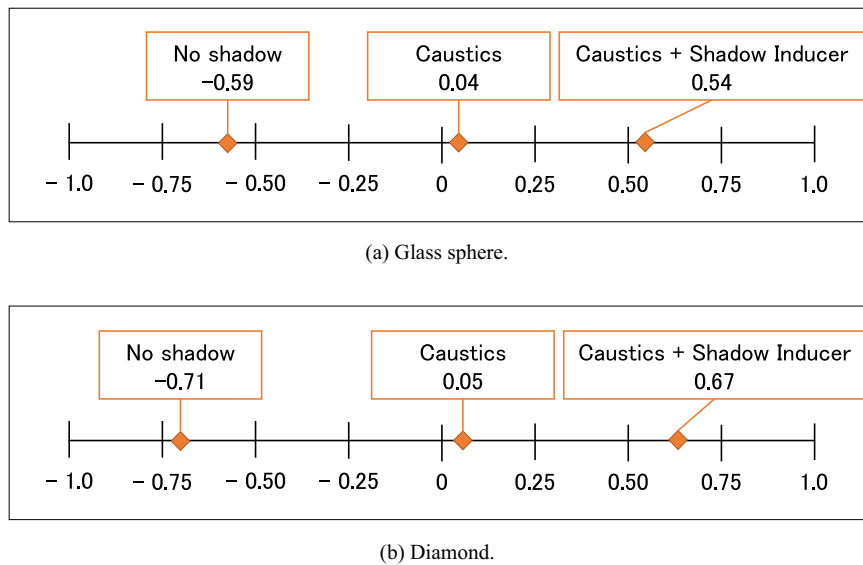
(a) Glass sphere.

(b) Diamond.

Fig. 12. Results of (a) glass sphere and (b) diamond.

are some possible explanations. First, the accuracy of tracking was insufficient, and the texture overlaid on the floor surface may have shifted in some cases. This time, we relied on the tracking function of the HoloLens. Empirically, registration errors of about 1 cm frequently occurred, and the tracking became stable even with these errors. When the participants or we found this kind of mis-registration, we reset and adjusted the world coordinate system by keyboard input. However, we cannot discard the possibility that some participants observed the virtual object without noticing small misalignment.

Second, the shadow edge might have been too sharp compared to the surrounding shadows, because the method assumed a point light source to generate the shadow region. This was also pointed out by the participants many times in the IEEE VR 2019 demo. The reason we used a point light source was to maximize the brightness contrast effect. If there are real shadows with a large-width gradation, sharp edges of virtual shadows obviously decrease the reality. Therefore, there is a trade-off between enhancing the brightness contrast and improving the reality. We need to investigate the relationship between the light source size and contrast effect.

Third, when the real background of the virtual object was very textured or bright as shown in Fig. 8 (middle row), the background image was prominent. It is possible that the effect of the reality improvement by the shadow decreases accordingly. Therefore, the virtual objects are required to be arranged so as not to do so. Extending to OST-HMD a rendering method based on the visibility of both the background and foreground textures may be a possible remedy [37].

In summary, we can conclude that presenting shadow inducers has the effect of improving the reality of virtual objects in various cases, even though the virtual objects require photo-realistic rendering. Further investigation is necessary to determine the effective range and conditions.

6. Limitations

Materials and Colors of the Ground We did not exhaustively investigate the kind of material and color of the ground to represent dark shadows without noticing the existing of the shadow inducer. From our experiences, we found that the shadow inducer works better on patterns with less regularity because the regularity makes it easier for users to predict which parts of the ground were brightened compared to the original brightness of the surface. This expectation is consistent with the comparison results of dimming rates in Fig. 7 (b).

Amplification by Shadow Inducer The conspicuousness of the shadow inducers in this paper is mainly a matter of our photography. We presented the virtual sphere in Fig. 9 to more than 10 people who had never experienced HoloLens, all of whom noticed the presence of shadows but did not notice the presence of the shadow inducer until they were briefed. However, there is so far no experimental evidence to show how much needs to be amplified or what is the optimal value of the parameter D for the shadow inducer. We believe that this problem can be solved by using a mathematical visual model of brightness induction because such visual models [38,39] can potentially predict the perceived darkness and gradient when given an arbitrary visual stimulus.

Boundary of Overlay FOV If the shadow inducer protrudes from the overlay FOV, an unnatural gap will appear at the edge of the overlay FOV. This gap can be thought to reduce the effect of shadow induction. This is an obvious weakness of our method if the overlay FOV is narrow. One possible solution is to make the boundaries of the overlay FOV unobtrusive by adjusting the gradient width D of the shadow inducer. However, this approach would undoubtedly reduce the effect of shadow induction.

Lighting Environment As described in Chapter 3, our method theoretically assumes that the illumination environment can be obtained and that the dominant light source can be approximated by a point light source. In our experiments, the position of the light source was given manually, and the environment map was obtained with an omnidirectional camera. Such a simple method may be useful in an environment where the light source is controlled, such as an exhibition space.

Capturing Real Environment Our method is based on the assumption that the real environment can be preliminary modeled to render the user-perspective images or the user-perspective images can be captured by cameras located near the positions of the eyes. However, it is difficult to capture images and overlay them in real time with a level of error that the user does not notice. In particular, there is currently no solution under the following conditions: the real surface near the shadow inducer has a specular reflection property, the real surface has complex shapes that are difficult to

represent with polygons, or the real environment including illuminations changes dynamically.

7. Conclusions

In this paper, we have presented the evidence of two kinds of effects caused by virtual shadows based on brightness inductions (in which a shadow perception is created by a brighter virtual object, the shadow inducer, placed around the shadow region). The region in a virtual shadow is perceived as darker than that in the non-overlaid outmost region even if both regions have almost the same luminance. Shadow inducers are capable of presenting weak shadows under the restriction that the shadow inducer keeps its inconspicuous appearance. There are multiple cases where shadow inducers improve the reality of virtual objects which are required to be rendered with multiple optical phenomena including shadows. Those facts show the feasibility of shadow inducers for optical see-through head-mounted displays without spatial light modulators.

Future work will include generation of the optimal shadow inducer based on predicting perceived images by a mathematical model of brightness induction [38,39]. By automatic generation of the optimal shadow inducers, it can be expected that the empirical parameters to render shadow inducers will be reduced and the effects of shadow induction will be maximized.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Sei Ikeda: Methodology, Investigation, Writing - review & editing. **Yuto Kimura:** Software, Data curation, Formal analysis, Visualization, Writing - original draft. **Shinnosuke Manabe:** Conceptualization, Software, Data curation, Formal analysis. **Asako Kimura:** Supervision, Validation. **Fumihisa Shibata:** Funding acquisition, Supervision, Project administration.

Acknowledgements

Apendix: Image Acquisition Method

A part of the images of the real scene including the virtual object shown in this paper were synthetic images generated from multiple real images with different exposures taken at the same viewpoint, because the dynamic range of our camera was insufficient to take augmented scenes. This section summarizes the detailed methods for image acquisition. For both experiments, all the real images containing virtual objects were taken through an OST-HMD (Microsoft HoloLens) with a digital video camera (Point Gray Research Flea 3). For Experiment 1, we carefully adjusted the brightness of the real scene and the virtual object and the shutter speed of the camera, then we succeeded in obtaining images of the virtual objects by single shots to avoid too much saturation. In

Table 1
ND filters used for image acquisitions.

Virtual object	Transmittance ratios of ND filter(s)
Glass Sphere	100%
Diamond	100%, 50%, 25%, 12.5%
Plastic bottle	50%, 25%,12.5%, 6.25%
Wine glass	100%, 50%, 25%,12.5%
Glass dish	50%,25%,12.5%, 6.25

this method, the virtual object must have the same level of brightness as the real scene and cannot be applied to scenes of transparent objects including high-brightness specular reflections and caustics. For Experiment 2, according to need, we obtained multiple images using neutral-density (ND) filters with different transmittance ratios and converted them to one merged image by Exposure Fusion [40], assuming a fixed camera position. Table 1 shows the transmittance ratios of the ND filters we used for each scene.

References

[1] Maimone A, Georgiou A, Kollin J. Holographic near-eye displays for virtual and augmented reality. ACM Trans on Graphics 2017;36(4):85:1–85:16.
[2] Kiyokawa K, Kurata Y, Ohno H. An optical see-through display for mutual occlusion with a real-time stereovision system. Computers & Graphics 2001;25(5):765–79.
[3] Kiyokawa K, Billinghurst M, Campbell B, Woods E. An occlusion capable optical see-through head mount display for supporting co-located collaboration. In: Proc. 2nd IEEE/ACM Int. Symp. on Mixed and Augmented Reality (ISMAR); 2003. p. 133–41.
[4] Uchida T, Sato K, Inokuchi S. An optical see-through mr display with digital micro-mirror device. Trans of the Virtual Reality Society of Japan 2002;7(2). (in Japanese)
[5] Cakmakci O, Ha Y, Rolland JP. A compact optical see-through head-worn display with occlusion support. In: 3rd IEEE and ACM Int. Symp. on Mixed and Augmented Reality (ISMAR); 2004. p. 16–25.
[6] Gao C, Lin Y, Hua H. Occlusion capable optical see-through head-mounted display using freeform optics. In: IEEE Int. Symp. on Mixed and Augmented Reality (ISMAR); 2012. p. 281–2.
[7] Gao C, Lin Y, Hua H. Optical see-through head-mounted display with occlusion capability. In: Proc. SPIE, 8735; 2013.
[8] Maimone A, Fuchs H. Computational augmented reality eyeglasses. In: Proc. IEEE Int. Symp. on Mixed and Augmented Reality (ISMAR); 2013. p. 29–38.
[9] Maimone A, Lanman D, Rathinavel K, Keller K, Luebke D, Fuchs H. Pinlight displays: wide field of view augmented reality eyeglasses using defocused point light sources. ACM Trans Graphics 2014;33(4):89:1–89:11.
[10] Yamaguchi Y, Takaki Y. See-through integral imaging display with background occlusion capability. Appl Opt 2016;55(3):A144–9.
[11] Itoh Y, Hamasaki T, Sugimoto M. Occlusion leak compensation for optical see-through displays using a single-layer transmissive spatial light modulator. IEEE Trans on Visualization and Computer Graphics 2017;23(11):2463–73.
[12] Reid RC, Shapley R. Brightness induction by local contrast and the spatial dependence of assimilation. Vision Res 1988;28(1):115–32.
[13] Gilchrist AL. Lightness contrast and failures of constancy: acommon explanation. Perception & Psychophysics 1988;43(5):415–24.
[14] Yang Z, Purves D. The statistical structure of natural light patterns determines perceived light intensity. Proc National Academy of Sciences of the United States of America 2004;101(23):8745–50.
[15] Jameson D, Hurvich LM. Complexities of perceived brightness. Science 1961;133(3447):174–9.
[16] Hong SW, Shevell SK. Brightness contrast and assimilation from patterned inducing backgrounds. Vision Res 2004;44(1):35–43.
[17] Bressan P, Actis-Grosso R. Simultaneous lightness contrast on plain and articulated surrounds. Perception 2006;35(4):445–52.
[18] Yund EW, Armington JC. Color and brightness contrast effects as a function of spatial variables. Vision Res 1975;15(8):917–29.
[19] Schirillo JA, Shevell SK. Brightness contrast from inhomogeneous surrounds. Vision Res 1996;36(12):1783–96.
[20] Sawayama M, Kimura E. Local computation of lightness on articulated surrounds. Iperception 2012;3(8):505–14.
[21] Bimber O, Frohlich B. Occlusion shadows: Using projected light to generate realistic occlusion effects for view-dependent optical see-through displays. In: Proc. Int. Symp. Mixed and Augmented Reality (ISMAR); 2002. p. 186–319.
[22] Bimber O, Grundhofer A, Wetzstein G, Knodel S. Consistent illumination within optical see-through augmented environments. In: Proc. 2nd IEEE/ACM Int. Symp. on Mixed and Augmented Reality (ISMAR); 2003. p. 198–207.
[23] Lee J.. Image projection apparatus, an image processing apparatus, image processing method, image processing program, and the use of the image data. 2015. JP5721210B2.
[24] Lamb M.. Display of shadows via see-through display. 2016. US Patent 9,311,751.

[25] Manabe S, Ikeda S, Kimura A, Shibata F. Casting virtual shadows based on brightness induction for optical see-through displays. In: Proc. IEEE Conf. on Virtual Reality and 3D User Interfaces (IEEE VR); 2018. p. 627–8.

[26] Kimura Y, Manabe S, Ikeda S, Kimura A, Shibata F. Can transparent virtual objects be represented realisticallyon OST-HMDs?. In: Proc. 26th IEEE Conf. on Virtual Reality and 3D User Interfaces (IEEE VR); 2019.

[27] Manabe S, Ikeda S, Kimura A, Shibata F. Shadow inducers: Inconspicuous highlights for casting virtual shadows on OST-HMDs. In: Proc. 26th IEEE Conf. on Virtual Reality and 3D User Interfaces (IEEE VR 2019); 2019.

[28] Kim K, Gu J, Tyree S, Molchanov P, Nieÿner M, Kautz J. A lightweight approach for on-the-fly reflectance estimation. In: Proc. IEEE Int. Conf. on Computer Vision (ICCV2017); 2017. p. 20–8.

[29] Gruber L, Langlotz T, Sen P, HÃ¶herer T, Schmalstieg D. Efficient and robust radiance transfer for probeless photorealistic augmented reality. In: Proc.IEEE Conf. on Virtual Reality and 3D User Interfaces (IEEE VR 2014); 2014. p. 15–20.

[30] Williams L. Casting curved shadows on curved surfaces. ACM SIGGRAPH Computer Graphics 1978;12(3):270–4.

[31] Adelson E.H.. Checkershadow illusion. 1995. http://persci.mit.edu/gallery/checkershadow.

[32] Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). Biometrika 1965;52(3–4):591–611.

[33] Steel RGD. A rank sum test for comparing all pairs of treatments. Technometrics 1960;2(2):197–207.

[34] Dwass M. Some k-sample rank-order tests. Contributions to Probability and Statistics 1960. https://ci.nii.ac.jp/naid/10018368267/

[35] Mosteller F. Remarks on the method of paired comparisons: III. a test of significance for paired comparisons when equal standard deviations and equal correlations are assumed. Psychometrika 1951;16:207–18.

[36] Kawabe T. Perceptual transparency from cast shadow. Iperception 2019;10(3):14pages.

[37] Fukiage T, Oishi T, Ikeuchi K. Visibility-based blending for real-time applications. In: Proc. IEEE Int. Symp. on Mixed and Augmented Reality (ISMAR); 2014. p. 63–72.

[38] Heinemann EG, Chase S. A quantitative model for simultaneous brightness induction. Vision Res 1995;35(14):2007–20.

[39] Blakeslee B, McCourt ME. A multiscale spatial filtering account of the white effect, simultaneous brightness contrast and grating induction. Vision Res 1999;39(26):4361–77.

[40] Mertens T, Kautz J, Reeth FV. Exposure fusion. In: Proc. 15th Pacific Conference on Computer Graphics and Applications (PG 2007); 2007. p. 382–90.