The Impact of Focus and Context Visualization Techniques on Depth Perception in Optical See-Through Head-Mounted Displays

Alejandro Martin-Gomez ⑩, Jakob Weiss ⑩, Andreas Keller, Ulrich Eck ⑩, Daniel Roth ⑩, and Nassir Navab

Abstract—Estimating the depth of virtual content has proven to be a challenging task in Augmented Reality (AR) applications. Existing studies have shown that the visual system makes use of multiple depth cues to infer the distance of objects, occlusion being one of the most important ones. The ability to generate appropriate occlusions becomes particularly important for AR applications that require the visualization of augmented objects placed below a real surface. Examples of these applications are medical scenarios in which the visualization of anatomical information needs to be observed within the patient's body. In this regard, existing works have proposed several focus and context (*F+C*) approaches to aid users in visualizing this content using Video See-Through (VST) Head-Mounted Displays (HMDs). However, the implementation of these approaches in Optical See-Through (OST) HMDs remains an open question due to the additive characteristics of the display technology. In this article, we, for the first time, design and conduct a user study that compares depth estimation between VST and OST HMDs using existing in-situ visualization methods. Our results show that these visualizations cannot be directly transferred to OST displays without increasing error in depth perception tasks. To tackle this gap, we perform a structured decomposition of the visual properties of AR F+C methods to find best-performing combinations. We propose the use of *chromatic shadows* and *hatching* approaches transferred from computer graphics. In a second study, we perform a factorized analysis of these combinations, showing that varying the shading type and using colored shadows can lead to better depth estimation when using OST HMDs.

Index Terms—Augmented reality, perception, depth estimation, visualization techniques, human computer interaction, design and evaluation methods, user studies

✦

1 INTRODUCTION

THE use of focus and context visualization techniques [3], [23] to facilitate the presentation of hidden components located inside of solid opaque objects have been widely investigated in the past. Such techniques may include cut-aways, cross-sections, exploded views or ghosting. In addition to the provision of cues to visualize hidden content, such techniques aim to promote a more intuitive understanding of the object's construction or composition. In the context of AR, F+C visualizations refer to those techniques that enhance the visualization of an augmented object of high interest and use additional cues to present and ground the augmented object within the real scene. This grounding

effect is particularly important when visualizing objects below real surfaces, as required for example in medical scenarios. Understanding the complex geometry of anatomical structures in neurosurgical interventions [45] or visualizing registered ultrasound imagery in-situ through the generation of virtual windows [1] is particularly important in this area. In industrial scenarios, a virtual exploded view can be seamlessly visualized within the context of a real object [30] to facilitate understanding of spatial relations.

Insights into the perceptual system [5], [6], [7], [40] have found that multiple visual cues are used to generate relative ordering and estimate the depth of observed objects. These cues have a different influence when estimating depth depending on the distance at which the observed objects are located. Three types of spaces, defined by the distance between the observer and the observed objects, have been defined in the past [5]: the *personal space* ($\approx 2m$), the *action space* ($\approx 2 - 30m$), and the *vista space* ($>30m$). Similar ranges have been used in AR to define *near-*, *mid-*, and *far-*distances when investigating depth estimation.

In this context, the strongest depth cue in all spaces is *occlusion*. *Binocular disparities* and *motion parallax* provide strong information in the personal space and their effectiveness decreases in function of the distance. In addition, *convergence and accommodation* provide useful information in personal space. According to [5], when conflicting depth cues are observed, occlusion dominates all other depth cues in all spaces, and only binocular disparity can provide similar information in such situations. Moreover, convergence

- *Alejandro Martin-Gomez and Nassir Navab are with the Chair for Computer Aided Medical Procedures and Augmented Reality, Department of Informatics, Technical University of Munich, 80333 Munich, Germany, and also with the Laboratory for Computer Aided Medical Procedures, Whiting School of Engineering, Johns Hopkins University, Baltimore, MD 21218 USA. E-mail: {alejandro.martin, nassir.navab}@tum.de.*
- *Jakob Weiss, Andreas Keller, and Ulrich Eck are with the Chair for Computer Aided Medical Procedures and Augmented Reality, Department of Informatics, Technical University of Munich, 80333 Munich, Germany. E-mail: {jakob.weiss, andi.keller, ulrich.eck}@tum.de.*
- *Daniel Roth is with the FAU Erlangen-Nürnberg, 91054 Erlangen, Germany. E-mail: daniel.roth@tum.de.*
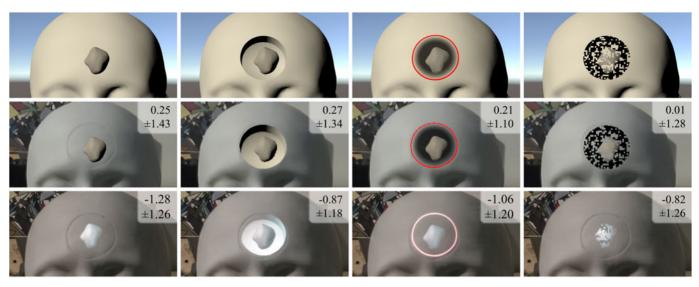
Fig. 1. *Base implementations of focus and context visualization techniques (top row) and their appearance in video- (mid row), and optical- (bottom row) see-through head-mounted displays.* From left to right: *Baseline* overlay without contextual layer, *Virtual Window*, *Contextual Anatomical Mimesis*, and *Virtual Mask*. Mean and standard deviation of corresponding alignment errors of study 1 are presented in centimeters. The OST images are captured using a smartphone camera placed at the eye position. Contrast and brightness have been adjusted for a faithful impression of the overlay as observed by the user.

and accommodation would not dominate other cues as they provide weak information in situations of conflict.

Thus, providing credible *occlusion* information represents a key requirement for AR applications that demand the visualization of augmented content behind the surface of real objects. Failing to provide appropriate additional cues could lead to the augmented content being perceived in front of the real object even when other cues like binocular disparities or motion parallax would suggest otherwise [51].

While F+C visualization techniques can produce the visual effect required to improve the perception of augmented content inside of a real object, for example by generating a virtual hole over its surface [1], they are dependent on the augmentation to fully block the real content to provide consistent occlusion cues. These visualization techniques have been extensively explored for screen-based AR systems such as hand-held devices or VST HMDs as they can effectively render the fully opaque dark colors required for the desired visualization. However, commercially available OST HMDs use additive displays which cannot selectively block or darken the light of the real world, but only project additional light. This leads to augmentations often being perceived slightly semi-transparent, as the light of the real world behind the display still reaches the eye, and the inability to render dark pixels which simply become transparent. This is exemplified in Fig. 1 where the VST and OST conditions use the exact same rendering algorithm yet produce very different outcomes.

In light of these different characteristics we pose the following questions: Can existing methods, which have been developed for VST displays, still be directly applied to OST HMDs? Which characteristics of the visualizations techniques result in problems and which ones provide benefits? Can novel methods be found which improve depth perception specifically on OST HMDs?

1.1 Contributions

To our knowledge, this is the first work that investigates how depth perception in in-situ visualization is affected by the additive displays commonly used in OST HMDs. We propose to decompose existing visualization techniques into different visual properties to provide a useful framework for further exploring the design space of F+C techniques for in-situ visualization in AR. Furthermore, we propose the use of *chromatic shadows* and *hatching* techniques to overcome the limited contrast and occlusion of OST displays. These two concepts were previously discussed in volume visualization [54] and illustrative rendering [46], yet have not been considered for contextual visualization in AR. We use a factorized experimental design that provides a differentiated analysis of the visual properties we have identified, highlighting the positive effects of our novel enhancements.

1.2 Research Strategy and Structure

In this work, we investigate *F+C* visualization techniques in the context of medical applications like many previous works [1], [2], [8], [24], [41]. Based on a literature review (Section 2), we identify three representative techniques for F+C visualization: Virtual Window [1], Contextual Anatomical Mimesis [51] and Virtual Masks [43]. To approach this topic with the scientific rigor necessary for surgical guidance, in Section 3 we evaluate these methods in a direct comparison between OST and VST display modalities. We designed an experimental apparatus that is modeled after a neurosurgical guidance scenario to consider the spatial limitations specific to these cases, and is evaluated regarding its accuracy in a control study. We then compare the visualization techniques and devices in a first user study. This first step of evaluation allows us to understand the differences between the two display modalities and tackle the challenges with respect to different visualizations as well as

internal positioning of the augmented object. On the basis of the resulting insights and analysis from the first study, in Section 4 we propose taking advantage of two visualization modules previously unexplored for augmented reality: *chromatic shadows* and *hatching*. To investigate the impact of each module, we decompose the visualization techniques into basic visual components, considering exterior visualization, interior rendering, and shadow representation and we perform a structured analysis of all combinations. A general discussion result from the insights of both studies, followed by our conclusion, are presented in Sections 5, and 6.

2　BACKGROUND AND RELATED WORK

Visual perception relates to the interpretation of sensory information collected by the human visual system and how this is used to understand the environment. While AR enables users to visualize computer generated content consistently embedded with the real world, several perceptual issues can occur when users observe and interpret an AR environment [7], [32]. These issues can be the result of multiple factors, such as the environment in which the AR system is used, the display technology employed, the design considerations to present the augmentations, and human factors. In this regard, existing works have identified visual acuity, contrast sensitivity, color perception and stereoacuity as four key measures that influence the visual capabilities associated to both, OST- and VST-HMDs [37]. These measures are used to describe how the differences between display technologies can lead to the loss of visual acuity and contrast when using AR HMDs, as well as to the distortion of the colors perceived by the observers.

One of the most challenging tasks in mixed reality relates to the correct estimation of depth [6], [47]. Existing studies have explored the accuracy in depth estimation at reaching distances, ranging 34-50 cm. Results from [52] show that the presence of highly-salient occluding surfaces has an effect on depth judgments, resulting in the underestimation of the virtual object's depth. In addition, results from [56] using *perceptual matching* and *blind reaching* as depth judgement methods for collimated targets in AR show that users can accurately estimate the distance of real objects. However, a systematic overestimation error of 0.50-4.0 centimeters was observed for virtual objects. Further studies explored the effects of focal distance, age, and brightness on estimating the depth of objects within similar reaching distances (33.3 to 50 cm) in AR [53]. Results from this work showed that focal distance and brightness affect depth estimation at reaching distances in AR, while participants were able to accurately estimate the position of the real objects as reported in [56]. More recent studies compared Retinal Projection Displays (RPD) and OST HMDs for depth judgment of real and virtual objects at such distances using a *blind reaching* method [44]. Results from this study show a global underestimation of both real and virtual objects when compared to the real distance regardless the technology used. However, virtual objects show an overestimation in depth when compared to the distance estimated for the real objects using OST. This overestimation is significantly reduced when using RPD. Other studies have investigated depth

perception in AR using not only egocentric but also exocentric methods with OST and VST HMDs [39].

Additional works have been proposed for the visualization of obscured information [14], and for the estimation of depth when real and virtual content is observed at medium-to-far distances [55]. Results from these studies have shown that the perceived depth of the virtual content is consistently underestimated at these distances. Later studies suggested that such effects can be mitigated using directed walking techniques [27], and that these improvements can be result of participants using peripheral optical flow information when using this technique [28].

In addition, multiple works have explored the effects that outdoor environmental conditions such as lighting changes or different backgrounds have over identification tasks, and how they affect user's performance when using OST HMDs. Studies in this field show that user's performance can be significantly affected by background textures, and by the style and color used to present the augmented content to the users [15]. These factors also modify the perception of the colors when the blending of virtual and real content is observed through OST HMDs [16].

In-situ visualization [2], [38], i.e., the augmentation of objects that lie below a real surface, is especially problematic in this regard as the perceived depth is frequently misjudged and biased towards the user [9]. This type of visualization is particularly important in medical AR, where surgeons need to be guided by augmentations within their visual field, indicating anatomical structures within the body that would normally be hidden from direct view. The challenges of depth perception have also been confirmed in studies in this application area, both for augmented microscope guidance [8] as well as for egocentric VST AR systems [51]. These findings highlight the need for perceptual visualizations that are able to convey depth relations between real and virtual objects correctly. An effective way to restore proper depth ordering through occlusion is to introduce a secondary virtual context object that is aligned with the real world surface, such as a virtual cutaway or virtual window [10], [13], [14], [51]. Virtual objects are only visible through the virtual cutaway and are occluded when not viewed at the right angle, thus reinforcing intended depth ordering. As a F+C approach, this provides detailed information on the focus object and spatial context through the virtual cutaway which is aligned with the real surface. Bajura *et al.* [1] have introduced the use of a virtual window in the medical context to show live ultrasound image slices within the patient's body. Later, Bichlmeier *et al.* [2] present a refined method that controls the visibility of the context surface as a function of curvature, angle of incidence and distance from the center of the focus. In the same spirit, opacity can be modulated using only the camera image without explicit knowledge of the surface geometry to create so-called ghosted views. An importance factor is computed for the camera pixels of a VST system using pq-space decomposition [35], gradient and saliency features [59], and then used to control the opacity of the context surface. For endoscopic AR and a known 3D reconstruction of the view, Wang *et al.* [57] extended this to additionally include the depth of the virtual object relative to the surface in the opacity modulation. In industrial settings, the approach of virtual windows

and cutaways has been used by Kalkofen *et al.* [29] to augment a real car with virtual views of the cabin or interior components and Schall *et al.* [50] used cutaway selection boxes to show underground assets such as ducts and cables. For VST AR, *stereoscopic pseudo-transparency* has been proposed [19], [43], a method that uses a random dots mask to generate the impression of seeing the virtual object through many small holes, leading to the real surface being perceived as semi-transparent. In a comparison to virtual window and a semi-transparent variant they find that the virtual mask can improve depth perception.

There is only little work that directly compares F+C visualization techniques that allows for interpretations related to OST HMDs in AR. Previous comparisons between these methods [43], [57] have exclusively used VST HMDs in their experimental setups. Given the drastically different visual appearance of these methods with additive displays discussed in Section 1 (c.f. Fig. 1), it remains unclear whether these effects on depth perception transfer to OST HMDs.

In a recent study, Heinrich *et al.* [24] focus on evaluating depth perception in a projective AR setup. They compare different visualization techniques in a depth sorting task and find a consistent ordering across variables where the best-performing methods are pseudo-chromadepth [49] and supporting lines [33]. The effects of these advanced visualizations were, however, found to be most prominent in a monocular setting. Stereoscopic display strongly improved depth perception across all methods, thereby greatly reducing the performance differences. In their evaluation, they used the full surface of the table as a projection area, augmenting a virtual space into the table with guidance lines to facilitate spacial understanding. In the medical context, there are few scenarios where such a large target area can be used when augmenting guidance onto the patient's body. They also acknowledge that the best-performing methods have additional limitations in the medical context: pseudo-chromadepth occupies the color channel of the augmented object and the supporting lines only work with a fixed reference plane which is not easily available in all scenarios. The work presented here, is focused on investigating methods that only require modification of the space surrounding the augmented object while working in the more constrained space presented by the human body, exemplified by the human head.

3 FOCUS AND CONTEXT TECHNIQUES IN AR

Manifold techniques have been presented for F+C visualization [1], [2], [13], [29], [35], [43], [50], [51], [57]. However, most of these techniques were developed using VST HMDs and rely on creating the illusion of virtual cutaway of the real surface.

Based on our review, we have identified three existing F +C visualizations which have shown promising results on VST and are suitable for adaptation in OST: the *virtual window* [1], *contextual anatomic mimesis* [2] and the *virtual mask* [43] (depicted in Fig. 1). These approaches take advantage of two of the strongest depth cues in personal space, occlusion and motion parallax. To create a sense of proper depth ordering, additional virtual content is used to provide the effect of partially or fully transparent areas over the surface

of a real object. The added virtual area does not conflict with the real surface in terms of depth as it merely sits on top of it. This allows the virtual focus object to be perceived as occluded by the surface of the real object but to be visible to the observer as well. Therefore, the virtual focus object can be displayed with correct depth ordering relative to the virtual transparent area, strengthening the illusion even more when combined with motion parallax.

Virtual Window. The virtual window, first introduced by Bajura, Fuchs and Ohbuchi [1], is the straightforward rendition of a virtual cutaway used to show virtual anatomical content inside of a patient. In a later version by Fuchs *et al.* [13] the rectangular window was adapted to only cut partially into the tissue, instead giving the impression of a hollow body in which the virtual objects can then be rendered freely. In our adaptation, we show the edges of the cutaway to reveal the inside of an otherwise hollow head and use a gray inside surface with diffuse shading.

Contextual Anatomic Mimesis. This visualization technique, introduced by Bichlmeier *et al.* [2], extends the idea of a virtual window into the body, adding subtle visual cues to retain some of the surface features while still allowing a look inside. The opacity of the real surface is controlled by three factors to create a ghosted view: (1) a radial gradient increases opacity towards the edges of the circular window, (2) high curvature areas have higher opacity and (3) modulation by the dot product between view direction and surface normal creates outlines of geometry. These three components, together with an additional explicit circle to delimit the augmentation, create the illusion of a semi-transparent skin surface while still keeping important parts of the original surface as context.

Other ghosting techniques have been proposed for the visualization of hidden content in urban environments [35], or for the visualization of multiple contextual layers [29], and to observe anatomical content in minimally invasive surgery [57], [59]. However their image-based nature does not lend itself to straightforward application for OST.

Virtual Mask. Otsuki *et al.* [43] have proposed to use a less sparse window which retains more contextual information. In their approach, the virtual surface is modulated by a fixed pattern of random dots to control the binary visibility of square patches of the window. As a result of this, parts of the real surface remain fully visible and occlude the virtual object creating strong occlusion cues which reinforce the intended depth ordering when using motion parallax. This effect of looking through many small holes in the surface generates what they call *stereoscopic pseudo-transparency*. In this work, we have implemented this method according to their original concept [42] with dot size of $1/64$ and dot density of 50 percent as evaluated in [43].

The presented visualizations have been primarily designed with VST in mind, so it is not immediately obvious if the same implementation can be used for OST as well. The inability of the additive OST technologies to display dark colors could potentially be compensated by the visual system by relying on other salient depth cues presented by some of the visualizations. We therefore proceed to investigate how the different visualizations are affected when used without modification in OST to show that methods need to be specifically designed for this display modality.
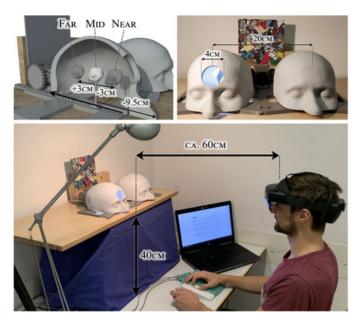
Fig. 2. *The experimental setup for Study 1.* Top: Two 3D printed phantoms of a head were placed in alignment to each other. The left head was used to present the physical target object, the right head was used to present the augmentations. Bottom: A participant executing the study by moving the augmentation on the right head to the perceived position of the physical object in the left head.

3.1 Study 1

Based on the previous discussion of the visualizations and depth cues, we believe that the lack of black color in OST reduces the effectiveness of the occlusion depth cue, therefore overall reducing the effectiveness of the F+C visualizations. Our hypothesis is therefore: (**H1**) *All F+C visualization techniques will result in inferior performance for OST, in comparison to VST.*

3.1.1 Design

The goal of this first study was to investigate the differences between OST and VST, as well as the impact of different visualization techniques and the target position (i.e., depth within the context object) on the perceived depth of a virtual object. Therefore, we conducted a 2(*Display Type*) × 4(*Visualization*) × 3(*Target Position*) within-subjects experiment with a perceptual matching protocol using a method of adjustments [11], [48]. Participants were asked to align a virtual object shown with one of four visualization techniques with a physical reference object placed in three positions within a 3D printed context object. The visualization techniques correspond to those presented in Section 3 and a baseline condition (see Fig. 1).

3.1.2 Apparatus

The physical setup (Fig. 2) is based on a scenario for neurosurgical guidance for tumor resection, where frontal access to the brain is sometimes required. It consisted of two identical 3D printed head replicas based on a human model with hollow interior, together with a 10×10 cm optical marker as well as an HTC Vive Tracker. A servo motor with a 3D printed linear translational stage was used to move a physical tumor within the left head to three defined positions. The positions were 3cm apart with a measured repositioning error of $<1mm$. A circular aperture with a diameter of 4cm allowed a direct line of sight on the physical tumor. The right head was completely closed to provide a low salient physical monochromatic occluder [52]. All components were fixed to a wooden panel to avoid physical movement. The interior of the left head was lit using white LEDs. A desk lamp was used to provide constant exterior lighting conditions ($\approx 175 lx$) throughout the study. Both lights were aligned to match the light direction of the augmentations. The setup included a linear slider to position the virtual objects and a separate button to confirm alignment. Participants were sitting on a chair at a fixed position, approximately at 70 cm from the *Mid* tumor position, and were allowed to move their head from left-to-right, and from front-to-back, but not above the table. This constrained the participant's eye position to a range of approx. 55-75cm from the *Mid* tumor position. The spatial arrangement of the participant's head, printed head models and virtual/physical aperture did not allow participants to see both virtual and real tumors at all the alignment positions. We intentionally designed this as a way to encourage participants to move their heads, thus provoking motion parallax which is an important factor in the effectiveness of the visualizations. A laptop positioned next to the participant was used to collect subjective measures.

To present the augmentations and perform the measurements, we designed an AR application using Unity 3D. An HTC Vive Pro in combination with an attached Stereolabs ZEDStereo camera was used for the VST condition, and a Microsoft HoloLens 2 Headset was used for the OST condition. We used an optical marker with Vuforia Tracking in order to achieve calibration in the OST, whereas in the VST condition, SteamVR 2.0 optical tracking was used.

System Error Assessment. In order to ensure there is no systematic bias and the accuracy of our solution is sufficient, we conducted a control study with $N = 10$ participants (3 female, 7 male, $M_{age} = 26.9$, $SD_{age} = 4.25$) to assess the potential error of the setup. In contrast to the main studies, participants were asked to align the virtual tumor that was shown in the left head, to match the physical target object in the same head in place. The virtual tumor was shown directly on top of the window without F+C techniques. After a training period of 6 trials for each display modality, 6 repetitions at each of the 3 tumor positions were measured from each direction for a total of 30 trials for each HMD. This procedure differed from the main task and was only used to assess the general error of our setup.

The results showed an average positioning error of $M = +0.10$ cm ($SD = 2.61$) for the VST system and $M = +0.14$ cm ($SD = 2.76$) for the OST system, i.e., participants placed the virtual object slightly farther away on average. There was no statistically significant difference between display technologies as determined by two-sample t-tests for alignment error ($t = 0.15$, $p = .879$). We therefore interpreted that there are no indications for a systematic bias due to the tracking or display system.

Just Noticeable Difference. To further ensure that our apparatus was appropriate, we calculated the just noticeable difference (JND) [12], [17], [18] as a means to ensure that the three different positions selected for the placement of the target were distinguishable by the participants during the

alignment task. We used the data from the main study and calculated the JND for each one of the target positions and devices. The highest JND was found at the *Far* position with 1.18cm (OST) and 1.22cm (VST), while the lowest was 0.90cm (OST) at the *Near* position. Thus the 3cm spacing between positions was found to be adequate.

3.1.3   Dependent Measures

Error and Time to Completion. We assessed the (signed) error and absolute error using the position of the virtual tumor relative to the physical reference as logged by the Unity application at the time the users marked a trial as completed. The (signed) error was computed as the difference between the position of the real tumor minus the position of the virtual tumor, using the tracked marker as the reference point. It is important to note here that due to the nature of the experiment (i.e., the participant placing the virtual object instead of indicating its position), this error has the opposite sign to the errors reported in the literature. Nevertheless, the meaning of over- and underestimation of depth will be aligned with past studies: an error = 0 would indicate an accurate judgement; an error < 0 an overestimation on the virtual tumor's depth (the virtual tumor was placed closer to the observer and thus perceived to be further away); and an error > 0 an underestimation in the perceived depth. In addition we measured the time from the presentation of the virtual tumor until trial completion.

Subjective Measures. To derive assumptions of the usability and task load, we used the NASA task load index questionnaire (TLX) [21] using a 21-point rating scale, and we analysed the individual subscales of this questionnaire (Raw-TLX) [22] as a means to avoid the introduction of additional sources of measurement errors [4]. Additionally, we used the single ease question (SEQ) [58] consisting of a 7-point Likert scale ranging from 1-(*Very easy*) to 7-(*Very difficult*) after each condition. To control for any negative effects, we also assessed a virtual reality sickness questionnaire (VRSQ) [31], composed of nine aspects, assessed using a 4-Point rating scale: *None, Slight, Moderate*, and *Severe*.

Control Assessments. To avoid any bias from visual impairments, we assessed a Landolt C-Test (EN ISO 8596) to test for the participants' acuity, an Ishihara Color test for color blindness [26] and a Titmus test for stereo vision. We further noted the interpupillary distance measurement reported by the Hololens 2 for each participant.

3.1.4   Procedure

The study consisted of two main phases: i) an introduction and tutorial phase and ii) the evaluation phase. Fig. 3 depicts the experimental procedure of the study.

On arrival, we welcomed the participants and informed them about the study before presenting them with a consent form and performing the Ishihara test, the Landolt test, and the stereo test to ensure correct or corrected-to-normal vision capabilities. Exclusion criteria included color vision deficiency, impaired stereopsis ($>140'$ angle of stereopsis at 40cm) or a visual acuity below 63 percent (20/32). In the following tutorial, participants could familiarize themselves with the task and input device. The tutorial involved the alignment of two virtual spheres, one with respect to the
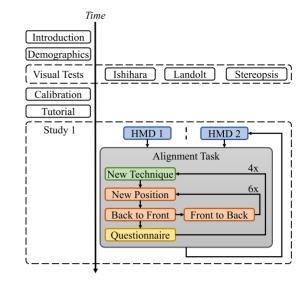

Fig. 3. *The experimental procedure.* Participants were randomly assigned to the order of conditions using a Latin-Square Matrix.

other, using the input device and presented using a desktop screen. Once participants felt comfortable with the input device, the tutorial was finished. The order of appearance of the visualization techniques and target object positions were preudo-randomly assigned using a Latin squared matrix. Half of the participants were asked to start with the OST HMD before switching to the second HMD, while the other half started with the VST HMD. Participants were equipped with the corresponding headset and assisted to wear it comfortably before proceeding to a calibration routine that served to adjust the headset's display according to their interpupillary distance (IPD).

During the evaluation phase, participants completed blocks of trials, one for each *Visualization $\times$ Display Type* combination. In each block, 12 total trials were measured. For each trial, the target object (tumor) was repositioned to one of 3 positions (*near, mid, far*, c.f. Fig. 2). The repositioning of the physical target was rapid, hence participants were not given special instructions during the repositioning phase (<300ms to move from *near* to *far*). For each position a total of four alignment trials were measured. For half of these trials, the virtual tumor was placed behind the target, requiring the participants to move it from back-to-front until they were satisfied. For the other half participants had to move the tumor from front-to-back. The starting position of the virtual tumor was chosen randomly from a range of $-5.5cm \pm 0.5cm$ for front-to-back positioning and $7.5cm \pm 0.5cm$ for back-to-front positioning, relative to the mid position of the physical tumor. This randomization prevents participants from remembering the previous slider positions. Before the participants were able to begin a trial and move the virtual tumor, they had to reset the slider position to the end of its range (front or back depending on the trial mode) as a consistent starting point. The participants were instructed to position the virtual tumor as precisely as possible, favoring precision over time. To be consistent with the *method of adjustments*, participants moved the tumor alternating between two different directions i.e., front-to-back or back-to-front, using the input device as instructed during the tutorial phase. To avoid very prolonged trials, observed

during preliminary experiments, participants were allowed to move the tumor only in one direction per trial.

For each OST block (HoloLens 2), the tracking was realigned with the marker to ensure optimal placement of the augmentation. After completing a block, the participants were asked to take off the headset (HTC Vive) or push the display up (HoloLens 2) to fill out the subjective questionnaires on a laptop. After completing all blocks for one HMD, participants were asked to complete the VRSQ questionnaire before switching to the other HMD.

A total of 96 positioning trials were performed for each participant, resulting from the possible combinations of display technology (2), visualization (4), tumor positions (3), approach directions (2) and approach repetitions (2). The duration of the experiment was between 37 and 89 minutes.

To mitigate the risk of COVID-19 we adopted the following procedure in accordance with local guidelines at the time the study was performed: i) The instructor wore a mask at all times a participant was present. ii) Participants were free to remove the mask while wearing the headset to avoid fogging up the displays. During this time a minimum safety distance was kept. iii) Disposable headset covers were used for the HTC Vive. iv) The disposable covers were replaced after each participant, and the headsets, chair, table and input device were wiped with a disinfectant.

3.1.5 Participants

We recruited 32 participants using mailing lists and campus announcements. Symptoms of or exposure to COVID-19 were a hard exclusion criterion. All 32 recruited participants (12 female, 20 male) had normal or corrected-to normal vision as confirmed by the vision tests. Their age ranged from 20 to 38 years ($M = 26.1$; $SD = 4.6$). On average, participants reported to play video/mobile games for 0.72 hours a day. 25 participants had previous AR experience. Those that had previous experiences reported a daily usage of $M = 0.12h$. Interpupillary distance (IPD) was assessed using the OST HMD. Participants IPD ranged between 63 and $67.5mm$ ($M = 64.74mm$, $SD = 1.48mm$). The study participation was voluntary and could be aborted at any time. German data privacy and anonymity laws were respected. The study data was anonymized and performed in accordance with the declaration of Helsinki and German guidelines.

3.2 Results

Our statistical analysis was performed using the SPSS Statistics software. We used three factorial (*Display Type × Visualization × Position*) repeated measures analysis of variance (ANOVA) as test for the objective assessments, using the mean of the four trials (2x back-to-front approach, 2x front-to-back approach) of each cell in the experiment.

For the subjective measures, two factorial (*Display Type × Visualization*) ANOVAs were calculated. Significance was accepted at $\alpha = 0.05$. We report Greenhouse-Geisser corrected values when the sphericity assumption was violated, assessed by Mauchly tests. We report Bonferroni adjusted post hoc tests for all pairwise comparisons. All the interaction terms were considered in this study, but only the significant ones are reported. We applied outlier correction on the
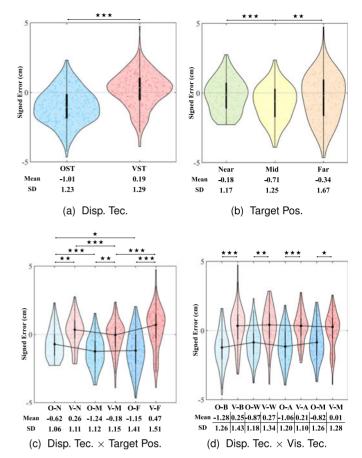

Fig. 4. *Error (Study 1)*. Violin plots depict distribution, median, and CI. Negative values describe a judgement of the virtual tumor closer to the participant in comparison to the actual target position of the real tumor. Results are depicted for Display Technology, (O-*:*OST*, V-*:*VST*), Visualization (*-B:*Baseline*,*-W:*Virtual Window*,*-A:*Contextual Anatomical Mimesis*,*-M:*Virtual Mask*), and Target Position (*-N:*Near*, *-M:*Mid*, *-F:*Far*), as well as their interactions.

objective measures by removing values greater than ± 3 standard deviations from the mean within each experimental condition (i.e., cell, factorial combination). We assume that errors greater than this were not related to a misjudgement of depth but due to intermittent connectivity or tracking failure or participant operating errors. For the alignment error, 21 samples (0.68 percent of the sample points) were removed, while for the time to completion, 64 samples (2.08 percent) were removed. To simplify the reporting, these were replaced with the new mean of the individual condition.

For the figures in the following sections, we use $^\star(p < 0.05)$, $^\star_\star(p < 0.01)$, and $^\star_{\star\star}(p < 0.001)$ to indicate levels of statistical significance. Plots are classic violin plots [25], indicating median with a white middle dot and the 25th and 75th percentiles with black bars. Reported descriptive values denote $mean \pm SD$ unless otherwise stated.

3.2.1 Alignment Error

Error. We investigated the signed error to identify if there were general trends regarding the error direction. We found significant interactions for *Display Type × Visualization*; $F(3, 93) = 3.59$, $p = .017$, $\eta_p^2 = .104$ (see Fig. 4). Post hoc comparisons showed significant differences between the

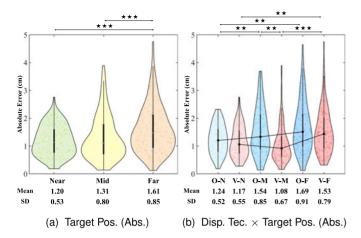(a) Target Pos. (Abs.)   (b) Disp. Tec. × Target Pos. (Abs.)

Fig. 5. *Absolute Error (Study 1).* Results are depicted for Display Technology, (O-*:*OST*, V-*:*VST*), and Target Position (*-N:*Near*, *-M:*Mid*, *-F:*Far*), as well as their interactions. Values in tables depict means and standard deviations in centimeters.

Display Type for all *Visualizations* ($ps \leq 0.023$) whereas the comparisons for the *Visualizations* when using the same *Display Type* were non-significant. However, the *Virtual Mask* showed the lowest alignment errors when compared to the other visualizations using both OST and VST. Interestingly, the alignment errors reported for the *Virtual Window* showed the worst results when using the VST HMD and the second best when using the OST, presenting values almost comparable to those obtained with the *Virtual Mask* (Fig. 4d). Regarding the *Display Type*, all OST conditions revealed that the virtual tumor was placed closer to the observer than the target position ($M = -1.01\ cm$), whereas all VST conditions showed the tumor was placed slightly further than the target position ($M = 0.19\ cm$). This effect was further supported by a main effect for *Display Type*; $F(1, 31) = 16.88$, $p < .001$, $\eta_p^2 = .352$ (Fig. 4a).

A significant interaction for *Display Type×Target Position* was observed; $F(1.4, 41.8) = 20.54$, $p < .001$, $\eta_p^2 = .399$ (Fig. 4c). Pairwise comparisons revealed that the *Mid* distance target position showed the strongest error in alignment (towards the participant) for both display types, corroborated by a main effect for *Target Position*; $F(2, 62) = 8.75$, $p \leq .001$, $\eta_p^2 = .220$ (Fig. 4b). Comparisons showed that the *Mid* target position showed a significant stronger error towards the observer (all $p \leq .004$) than *Near* and *Far*, whereas the *Near* and *Far* did not significantly differ. However, this error was mainly due to the OST, and for the OST, also strongly present for the *Far* position.

Absolute Error. The descriptives and pairwise comparisons for the absolute error are depicted in Fig. 5. There was a statistically significant interaction for *Display Technology×Position*; $F(1.90, 59.01) = 5.85$, $p = .005$, $\eta_p^2 = .159$. Pairwise comparisons revealed a significant difference between both displays for the *Mid* and position ($p = .008$), where the OST showed a higher absolute error compared to VST, but not for *Near* ($p = .523$) or *Far* ($p = .300$). Differences for *Target Position* were observed for the *OST Near* position, showing a higher accuracy than *Mid* and *Far* (all $p \leq .002$), see Fig. 5b. Further, the *VST Target Position* comparisons revealed that the *Far* position showed significantly higher errors than *Near* and *Mid* (all $p \leq .001$).

TABLE 1
Mean Raw-TLX Scores Reported by Participants Grouped by *Display Technology* and *Visualization*

Disp. - Vis.	MD	PD	TD	OP	EF	FL
OST - Baseline	9.78	4.56	4.03	11.84	9.19	6.13
OST - Window	8.56	3.91	3.63	13.28	7.75	5.75
OST - Mimesis	11.19	5.09	4.09	10.38	9.81	6.78
OST - Mask	12.59	5.19	4.53	9.47	10.53	7.28
VST - Baseline	10.81	5.25	3.78	10.94	10.38	6.97
VST - Window	8.78	4.22	3.59	13.50	8.88	6.06
VST - Mimesis	9.81	4.56	3.50	11.94	8.00	5.97
VST - Mask	13.16	5.28	3.78	10.16	10.56	7.88

MD:Mental Demand, *PD*:Physical Demand, *TD*:Temporal Demand, *OP*:Overall Performance, *EF*:Effort, *FL*:Frustration Level.

These results were corroborated by a main effect for *Target Position*; $F(1.41, 43.73) = 26.6$, $p < .001$, $\eta_p^2 = .462$. Consecutive pairwise comparisons showed that the *Far* position evoked the largest absolute alignment error, significantly larger than both other positions ($ps < .001$).

3.2.2 Time to Completion

Analysing the time to completion, we found a strongly significant main effect for *Target Position*; $F(1.37, 42.34) = 79.62$, $p < .001$, $\eta_p^2 = .720$. Pairwise comparisons showed that the *Far* position took the longest ($M = 13.04s$, $SD = 5.34$), followed by *Mid* ($M = 10.17s$, $SD = 3.80$) and *Near* ($M = 9.92s$, $SD = 3.83$; all $p < .001$). No further main or interaction effects were found.

3.2.3 Subjective Measures

NASA TLX. The descriptives of the Raw-TLX scores are reported in Table 1. There was a statistically significant interaction for *Display Technology×Visualization* for Mental Demand; $F(3, 93) = 3.99$, $p = .010$, $\eta_p^2 = .114$. The analysis further revealed a main effect for *Visualization*; $F(3, 93) = 20.20$, $p < .001$, $\eta_p^2 = .395$. Pairwise comparisons showed that using the *Virtual Mask* lead to higher mental demand than the *Baseline*, *Virtual Window*, and *Anatomical Mimesis* (all $p < .001$). And further, when using the *Anatomical Mimesis* ($p = .010$) compared to the *Virtual Window*. A statistically significant interaction for *Display Technology×Visualization* for Effort was also found; $F(2.44, 75.5) = 4.79$, $p = .007$, $\eta_p^2 = .134$. A main effect for *Visualization*; $F(3, 93) = 9.24$, $p < .001$, $\eta_p^2 = .230$ and pairwise comparisons showed that participants required more effort with the *Virtual Mask* than with the *Anatomical Mimesis* ($p = .005$) and *Virtual Window* ($p < .001$). Moreover, we found significant main effect for *Visualization* when analyzing the Physical Demand; $F(2.18, 67.42) = 3.52$, $p = .032$, $\eta_p^2 = .102$, the Overall Performance; $F(3, 93) = 14.00$, $p < .001$, $\eta_p^2 = .311$, and the Frustration Level; $F(3, 93) = 3.94$, $p = .011$, $\eta_p^2 = .113$. Pairwise comparisons showed that the *Virtual Mask* required higher physical demand ($p = .013$), and lead to lower overall perceived performance ($p < .001$), and higher frustration level ($p = .015$) than the *Virtual Window*, see Table 1.

Single Ease Question. Analysing the SEQ, we found significant interactions for *Display Type×Visualization*; $F(3, 93) = 4.95$, $p = .003$, $\eta_p^2 = .138$. These results were corroborated by

main effects for *Display Type*; $F(1,31) = 4.95$, $p = .034$, $\eta_p^2 = .138$, and *Visualization*; $F(3,93) = 24.67$, $p < .001$, $\eta_p^2 = .443$. Pairwise comparisons showed that the task was easier to preform when using the VST HMD ($p = .034$). Regarding the visualizations, the task was perceived to be significantly easier to perform when using the *Virtual Window* ($M = 3.27$, $SD = 1.20$), followed by the *Anatomical Mimesis* ($M = 3.94$, $SD = 1.26$; $p = .002$), *Baseline* ($M = 4.11$, $SD = 1.39$; $p = .002$) and *Virtual Mask* ($M = 4.77$, $SD = 1.37$; $p < .001$).

Virtual Reality Sickness Questionnaire. The scores collected after using each display were analyzed. As expected, participants reported significantly lower scores when using the OST; $F(1,31) = 30.89$, $p < .001$, $\eta_p^2 = .499$, $M = 13.15$, $SD = 9.26$, than the VST ($M = 21.90$, $SD = 12.78$).

3.3 Summary and Discussion

We found that all visualizations perform worse on OST than on VST regarding the error, supporting **H1**: *All F+C Visualization techniques will result in inferior accuracy performance for OST, in comparison to VST*.

There is a significant effect between the display modalities regarding the error. On average, participants perceived the virtual tumor to be further away. Therefore, they placed the virtual object closer to themselves than the real target when using the OST display. These results are consistent with the perceptual matching results reported in [56] for virtual content, but differ from the underestimation observed when a physical obstacle was used to occlude the virtual objects in [52]. However, this difference can be a result of the highly salient obstacle used in [52]. Conversely, participants slightly positioned the virtual tumor farther than the real tumor in the VST condition. The slight offset observed for the VST (M= 0.19cm) is however of a similar order to the results derived from the system error assessment for this condition (M= 0.14cm). Thus, the errors reported may not necessarily be indicative of misjudgments in depth caused by the in-situ visualization. Interestingly, these subtle differences, close to two millimeters, are similar to the results of the perceptual matching experiments reported by [56] for real content.

Specifically for the *Mid* position of the reference, users perceived the virtual tumor significantly closer for both display types than in the *Near* and *Far* conditions, while the absolute positioning error was highest for the *Far* position for both display types. While it is intuitive that the absolute positioning error increases with the distance from the user, the bias in the *Mid* position has no obvious explanation. One interpretation is that the contextual structures that could serve as depth cues are further away.

In the original presentation of the virtual mask [43], the authors found that the use of this visualization, as well as cut-outs, improve depth judgements when compared to a *no masking* condition for the observation of a virtual object placed 0.1 to 2cm behind a surface. These conditions most closely correspond to our *Near* position, comparing the *Virtual Mask* and *Window* to our *Baseline*. While our findings in this condition show similar results for the *OST*, the *VST* condition only presents a similar trend for the *Virtual Mask* and not for the *Virtual Window* (Figs. 1 and 4d).

In terms of subjective measures, *Virtual Window* consistently scored best while *Virtual Mask* scored worst. Results

from the NASA TLX questionnaires show that participants perceived the alignment task to be more mentally and physically demanding, reporting higher effort and frustration levels as well as lower overall performance scores when comparing these two visualizations. Interestingly, based on our observations and despite the task being the same, participants may have perceived the task to be more physically demanding when using the *Virtual Mask* due to it requiring the use of motion parallax to derive additional information. This is in line with comments from participants who did not like the visual appearance of the virtual mask, however the accuracy measures showed that *Virtual Mask* is comparable to *Virtual Window*, improving on the other two. Our interpretation is that in the case of the *Virtual Mask*, the holes add more depth occlusion and parallax depth cues, however the overall increased visual complexity and the lack of background in the OST display made this visualization harder to understand intuitively and thus reduced subjective scores.

4 IMPROVING F+C FOR OST HMDs

The results of our first study have made it evident that there is a need for more specialized visualizations that adapt specifically to OST HMDs. Based on the experience and user feedback from our first study, we have identified two important shortcomings in OST that reduce the effectiveness of the existing techniques: First, the inability to render opaque black in an additive display affects the visual outcome when compared to VST, most strongly changing how *Contextual Anatomic Mimesis* and *Virtual Mask* manifest. Second, the contrast of the display reduces visibility of shading details, for example within the interior surface of *Virtual Window*.

To tackle these problems and motivated by our previous findings, we propose the use of two techniques which have not previously been explored for AR: First, the disappearing dark regions can be mitigated by compressing the intensity range, mapping black to gray. These gray areas provide a better border between non-augmented regions and areas which are augmented but black. To mitigate the lost contrast in luminance that this approach implies, we propose using *chromatic shadows* to introduce additional contrast through chrominance differences. *Chromatic shadows* have previously been used to enhance the contrast of shaded areas in scientific volume visualization [54]. Second, we observe that regions with low luminance still appear translucent because the real world is visible behind the augmentation, whereas very bright overlay regions appear solid. To exploit this effect, we propose the use of *hatching* techniques commonly employed in illustrative rendering [34]. The bright strokes are highly salient, strongly covering the real-world background and thus appearing highly solid. They also have the additional benefit of adding texture to an otherwise featureless surface, conveying the surface shape better than shading details on a low contrast OST. Moreover, hatching techniques provide additional information with motion perspective, which represents the third most significant depth cue in personal space after occlusion and binocular disparity [5]. These two techniques could potentially be combined, i.e., the space between hatches could be filled
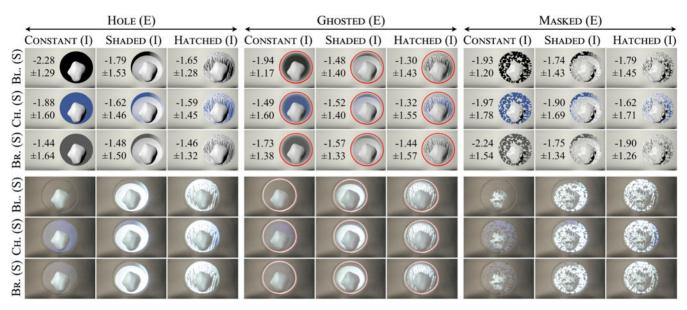
Fig. 6. *All 27 visualization permutations tested in Study 2.* Top: Base implementation in virtual environment, with error $mean \pm SD$ in cm. Bottom: Corresponding appearance of the visualizations in the OST HMD. The images are captured with a smartphone camera placed at the eye position. Contrast and brightness have been adjusted to provide a faithful representation of the real view. (E) Exterior, (I) Interior, (S) Shadow.

with a shadow color to further increase the coverage of the overlay. On further consideration of the methods evaluated in Study 1, we note that these methods are partially orthogonal and could therefore also be mixed together.

In order to perform a systematic evaluation of the characteristics of the existing methods and of combinations with our novel ideas, we have decomposed the techniques into the three orthogonal properties Exterior *Visualization*, Interior *Rendering* and Shadow *Representation*, which characterize the previous methods and allow for natural integration of our new ideas. Using these characteristics allows a factorized analysis of the components, assessing their impact individually. Each of the three visualizations studied in the first part of this work can be described in terms of these three characteristics, as shown in Fig. 6 and explained in Table 2.

Exterior *Visualization.* The way the surface of the real object is visualized is the strongest characteristic difference between the methods evaluated in Study 1. This characteristic essentially describes how the opacity of the real surface is virtually modulated.

Interior *Rendering.* With all cutout techniques, the appearance of the virtual interior can be rendered in different ways. For example, the *Virtual Window* only employed a Phong shading on the interior surface of the window. We propose the use of *hatching* [46], a general surface shading method used to create images that appear similar to pencil drawings. The hatches, which we inverted so they give bright streaks, create a stronger intensity contrast and add additional visual detail in the background.

Shadow *Representation.* A minimum luminance even in fully darkened areas can be ensured by using a method for producing illustrative *chromatic shadows* [54]. This method uses a *shadowiness* parameter S to compute a shadow color parameter relative to the surface color. The shadow color is adapted by partially shifting the natural luminance contrast from a gray-level shading to a chrominance contrast with a shadow color tone that has the same perceptual distance

(measured in CIELAB color space) than the original black shadow color.

$$C_{RGB} = (1 - S)C^O + SC_S. \qquad (1)$$

In our implementation, we use the clamped Lambert term $S = \max(n \cdot L, 0)$ as the *shadowiness* factor. Given that the interior surface color is pure white, the shadow color C_S does not depend on the surface position and we can use a constant shadow color. We evaluate a blue-tinted shadow which performs well according to [54] in their application, and a grayscale shadow to compare whether a tinted shadow has a favorable effect.

4.1 Study 2

Based on our findings from Study 1, and considering the perceptual advantages and visual cues provided by the

TABLE 2
Possible Values for the Three Visualization Properties

Exterior *Visualization*	
hole	A circular cutout hole with a hard edge.
ghosted	Advanced opacity modulation based on curvature, normal, and Gaussian falloff as used in *Contextual Anatomical Mimesis*.
mask	Circular cutout modulated by a binary random stencil texture as in *Virtual Mask*.
Interior *Rendering*	
constant	Constant background color.
shaded	Diffuse shading applied to the interior.
hatched	Illustrative hatching.
Shadow *Representation*	
black	A shadow color of $C_S^1 = (0, 0, 0)$, reduces 1 to standard diffuse shading
chromatic	A blue shadow of color $C_S^2 = (63, 89, 150)$
bright	A gray value $C_S^3 = (89, 89, 89)$, chosen to have the same luminance as C_S^2

techniques described in Section 4, we formulate two hypotheses: *(H2) Using visualization techniques that enhance the contrast between a virtual object and the background observed contribute to better estimate the depth of the object and help to reduce the alignment error when using OST displays,* and **(H3)** *Using visualization techniques that enhance the contrast of shaded areas by ensuring a minimum luminance value contribute to better estimate the depth of the virtual objects observed and help to reduce the alignment error when using OST displays.*

4.1.1 Design

We conducted a 3(EXTERIOR *Visualization*)×3(INTERIOR *Rendering*)×3(SHADOW*Representation*)×3(Target Position) within-subjects follow-up study using the apparatus as in Study 1, limited to the investigation with an OST HMD. The factorial design aimed to decompose the aspects of the visualization techniques investigated in Study 1, and resulted in the 27 different visualization variants (i.e., Study cells) depicted in Fig. 6.

4.1.2 Measures

We used measures analogously to Study 1 regarding the objective assessments: (signed) error, absolute error, and time to completion; as well as analogue vision tests. Since we focused on the analysis of the components of the visualization techniques more closely, we did not assess the NASA TLX questionnaire but introduced a rating for visual attractiveness: *"Overall, I liked the appearance of the visual information provided by this technique"*, answered using a 7-point Likert scale ranging from *1-Strongly agree* to *7-Strongly disagree.*

4.1.3 Procedure

The overall procedure of this study followed the structure of Study 1, including the guidelines for COVID-19 mitigation. To compensate for the increased number of conditions, we reduced the number of repetitions to 2 (1 back-to-front, 1 front-to-back), resulting in a total of 162 trials (one block for each visualization combination (27) consisting of 3 target position and 2 repetitions). The positioning and control of the virtual tumor was adapted to avoid any influence of the linear slider: The position and randomization range at which the tumor can appear was increased to $-9 \pm 1cm$ (front-to-back) and $9 \pm 1cm$ (back-to-front). The movement scaling of the linear slider was also decreased by 45 percent to allow for more precise control. Furthermore, participants were asked to fill the two subjective questions on the laptop while looking through the headset instead of folding up the HMD. This streamlined the experiment and avoided the need for recalibration using the tracked marker, which was now only performed when the user indicated misalignment. The experiment took about 75 minutes, on average.

4.1.4 Participants

Participants were recruited analogue to Study 1, but had no prior exposure to it. Participants that failed the vision tests were excluded. The final sample consisted of 27 participants (9 female, 18 male; age 22 to 34 years; $M = 26.15$; $SD = 3.4$). None reported to have any motor impairments. On average, participants reported to play video/mobile games for 0.84

hours a day and 20 participants had previous experience with AR. Those that had previous experiences reported an average daily usage of $M = 0.14h$. The participants' IPD ranged between 59.3 and 67.7mm ($M = 63.69mm$, $SD = 2.75mm$).
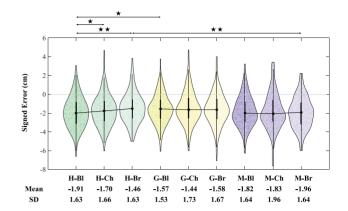

Fig. 7. *Error for* EXTERIOR × SHADOW *.* Lines connect median values. Results are depicted for EXTERIOR, (H-*:*hole*, G-*:*ghosted*, M-*:*mask*), and SHADOW Representation (*-Bl:*black*, *-Ch:*chromatic*, *-Br:*bright*), as well as their interactions. Values in tables depict means and standard deviations in centimeters.

	H-Bl	H-Ch	H-Br	G-Bl	G-Ch	G-Br	M-Bl	M-Ch	M-Br
Mean	-1.91	-1.70	-1.46	-1.57	-1.44	-1.58	-1.82	-1.83	-1.96
SD	1.63	1.66	1.63	1.53	1.73	1.67	1.64	1.96	1.64

4.2 Results

We calculated four-way (EXTERIOR*Visualization*× INTERIOR*Rendering*×SHADOW*Representation*×*Target Position*) repeated measures analyses of variance (ANOVAs) for the objective assessments, aggregating the repetitive trials for each cell. For the subjective measures, three-way (EXTERIOR*Visualization*×INTERIOR*Rendering*×SHADOW*Representation*) *ANOVAs were calculated. Significance was accepted at* $\alpha = 0.05$. *We report Greenhouse-Geisser corrected values when the sphericity assumption was violated, assessed by Mauchly tests. We report Bonferroni adjusted post-hoc tests for the pairwise comparisons. The outlier correction was performed analogous to Study 1. This procedure removed 44 (1.01 percent) of the measures for the alignment accuracy measure and 85 (1.94 percent) measures for time to completion. To simplify the analysis and reporting, these values were replaced with the remaining mean of the specific cell.*

4.2.1 Alignment Error

Error. We found significant interaction for EXTERIOR*Visualization*×SHADOW*Representation*; $F(2.71, 70.51) = 4.23$, $p = .010$, $\eta_p^2 = .140$. With the *Mask* exterior, errors were relatively similar for each shadow representation. The *chromatic* shadow led to the smallest error with the *mask* exterior, whereas the *bright* shadow lead to the smallest error in the synthetic *hole* conditions, see Fig. 7. Overall, the latter led to the lowest error in these combinations.

Additionally, a significant interaction for EXTERIOR*Visualization*×*Target Position* was found; $F(4, 104) = 3.38$, $p = .012$, $\eta_p^2 = .115$, and pairwise comparisons showed that participants performed better with *Ghosted* than *Mask* at *Near* ($p = 0.013$) and *Mid* ($p = 0.025$) positions and better with *Hole* than *Mask* at the *Mid* position ($p = 0.027$). These results can be seen in Fig. 8.

The relatively strong impact of the EXTERIOR*Visualization* was corroborated by a main effect; $F(1.26, 32.84) = 4.71$, $p = .030$, $\eta_p^2 = .153$. Pairwise comparisons showed that, overall,
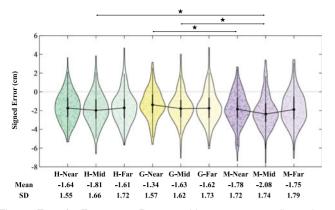
Fig. 8. *Error for* EXTERIOR × POSITION . Lines connect median values. Results are depicted for EXTERIOR, (H-*:*hole*, G-*:*ghosted*, M-*:*mask*), and POSITION (*-Near, *-Mid, *-Far), as well as their interactions. Values in tables depict means and standard deviations in centimeters.

the *mask* evoked the largest error, significantly larger than the *hole* visualization ($p = .007$). Moreover, a main effect for INTERIOR*Rendering* was found; $F(2, 52) = 6.46$, $p = .003$, $\eta_p^2 = .199$. Comparisons revealed that participants showed the largest negative errors with the *constant* rendering, significantly larger than with *hatched* ($p = .015$). These results are summarized in Figs. 9a and 9b. Overall, the *ghosted black* visualization in combination with the *hatching* technique revealed the smallest error ($M = -1.30$, $SD = 1.43$) when compared to all the other combinations (see Fig. 6).

Absolute Error. Our statistical analysis for absolute error revealed significant interaction for EXTERIOR*Visualization*×*Target Position*; $F(4, 104) = 2.93$, $p = .024$, $\eta_p^2 = .101$. A main effect for EXTERIOR*Visualization*; $F(1.44, 37.55) = 7.79$, $p = .004$, $\eta_p^2 = .230$ was found. Pairwise comparisons showed that the *mask* evoked the largest absolute error, significantly larger than the *ghosted* ($p = .012$) and the *hole* ($p = .003$), see Fig. 10a.

In addition, our statistical analysis revealed a main effect for INTERIOR*Rendering*; $F(1.50, 39.09) = 7.25$, $p = .004$, $\eta_p^2 = .218$. Absolute errors for *hatched* were significantly lower compared to *constant* ($p = .014$). See Fig. 10b.

4.2.2 Time to Completion

During the analysis of the time to completion, a strong significant main effect for *Target Position* was found; $F(2, 52) =$
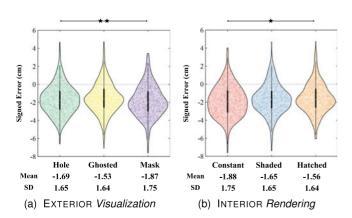

Fig. 9. *Error (Study 2).* Results are depicted for (a) EXTERIOR*Visualization*, and (b) INTERIOR*Rendering*. Values in tables depict means and standard deviations in centimeters.
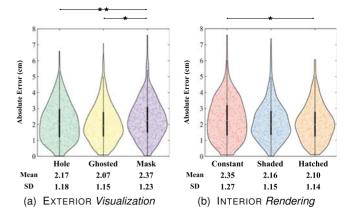

Fig. 10. *Absolute Error (Study 2).* Results are depicted for (a) EXTERIOR*Visualization*, and (b) INTERIOR*Rendering*. Values in tables depict means and standard deviations in centimeters.

25.35, $p < .001$, $\eta_p^2 = .494$. Pairwise comparisons showed that the *Far* position took the longest ($M = 11.89$, $SD = 7.54$), followed by *Mid* ($M = 10.77$, $SD = 6.62$; $p < .001$) and *Near* ($M = 10.61$, $SD = 6.49$; $p < .001$). No further main or interaction effects were found.

4.2.3 Subjective Measures

Single Ease Question. Results for the SEQ (Table 3) revealed significant interactions between EXTERIOR*Visualization*×SHADOW*Representation*; $F(4, 104) = 4.04$, $p = .004$, $\eta_p^2 = .135$. THESE RESULTS WERE CORROBORATED BY MAIN EFFECTS FOR EXTERIOR-VISUALIZATION; $F(1.26, 32.83) = 24.72$, $p < .001$, $\eta_p^2 = .487$. PAIRWISE COMPARISONS SHOWED THAT PARTICIPANTS PERCEIVED THE TASK EASIER TO COMPLETE WHEN USING THE *HOLE* ($M = 2.83$, $SD = 1.30$; $p < .001$) THAN THE *GHOSTED* ($M = 3.63$, $SD = 1.44$) AND THE *MASK* ($M = 3.86$, $SD = 1.56$). IN ADDITION, OUR ANALYSIS REVEALED A SIGNIFICANT MAIN EFFECT FOR SHADOW*Representation*; $F(2, 52) = 14.051$, $p < .001$, $\eta_p^2 = .351$. POSTERIOR PAIRWISE COMPARISONS REVEALED THAT USERS FIND THE ALIGNMENT TASK EASIER TO ACHIEVE WITH *CHROMATIC* SHADOWS ($M = 3.20$, $SD = 1.40$) THAN WITH *BLACK* ($M = 3.56$, $SD = 1.56$; $p < .001$) AND *BRIGHT* ($M = 3.55$, $SD = 1.52$; $p = .001$) REPRESENTATIONS.

Moreover, significant interactions for INTERIOR*Rendering*×SHADOW*Representation*; $F(4, 104) = 3.87$, $p = .006$, $\eta_p^2 = .130$ were found. Main effects for INTERIOR*Rendering*; $F(2, 52) = 19.33$, $p < .001$, $\eta_p^2 = .351$, revealed that users

TABLE 3
SEQ Results for Study 2, Reported as $mean \pm SD$ in Terms of EXTERIOR Visualization, INTERIOR Rendering (Co-*:*constant*, Sh-*:*shaded*, Ha-*:*hatched*) and SHADOW Representation (*-Bl:*black*, *-Ch:*chromatic*, *-Br:*bright*)

	Hole	Ghosted	Mask
Co-Bl	3.67 ±1.44	4.44 ±1.62	4.85 ±1.58
Co-Ch	2.74 ±1.14	4.04 ±1.60	3.89 ±1.75
Co-Br	3.63 ±1.47	4.41 ±1.28	4.74 ±1.69
Sh-Bl	2.59 ±1.28	3.11 ±1.26	3.81 ±1.33
Sh-Ch	2.67 ±1.12	3.26 ±1.17	3.07 ±1.02
Sh-Br	2.74 ±1.26	3.67 ±1.31	3.33 ±1.33
Ha-Bl	2.52 ±1.00	3.07 ±1.15	4.00 ±1.49
Ha-Ch	2.37 ±0.99	3.30 ±1.24	3.44 ±1.40
Ha-Br	2.52 ±1.17	3.33 ±1.39	3.59 ±1.28

found it easier to complete alignment when using the *hatched* ($M = 3.13$, $SD = 1.35$) and *shaded* ($M = 3.14$, $SD = 1.30$) rendering compared with the *constant* ($M = 4.05$, $SD = 1.64$; $p < .001$).

Visual Attractiveness. Results obtained after analysing the participants' opinion regarding visual attractiveness showed significant interaction for INTERIOR*Rendering*×SHA-SHADOW*Representation*; $F(4, 104) = 4.46$, $p = .002$, $\eta_p^2 = .146$. A MAIN EFFECT FOR INTERIOR*Rendering* REVEALED THAT PARTICIPANTS FOUND THE *SHADED* ($M = 3.35$, $SD = 1.54$; $p < .001$) AND THE *HATCHED* ($M = 3.45$, $SD = 1.50$; $p = .004$) RENDERINGS MORE VISUALLY APPEALING THAN THE *CONSTANT* ($M = 4.17$, $SD = 1.63$).

Moreover, our test revealed a strong significant main effect for EXTERIOR*Visualization*; $F(2, 52) = 27.06$, $p < .001$, $\eta_p^2 = .510$. Posterior pairwise comparisons showed that participants found the *mask* ($M = 4.30$, $SD = 1.59$) less visually appealing than the *ghosted* ($M = 3.73$, $SD = 1.49$; $p = .005$) and the *hole* ($M = 2.95$, $SD = 1.41$; $p < .001$), as well as the *ghosted* than the *hole* ($p < .001$).

4.3 Summary and Discussion

In terms of EXTERIOR*Visualization*, we have found that *mask* performs badly for error measures as well as both subjective (usability related) metrics. Even the combinations which add a brighter background cannot improve on the error metrics, and the subjective metrics show the same behavior. The *hole* and *ghosted* perform similarly for error metrics, however participants seem to prefer *hole* visualizations over using *ghosted* techniques. This is an interesting finding given that *hole* is part of the earliest works for in-situ AR, whereas *ghosted* and *mask* are far more recent methods based on careful arguments and considerations on visual perception.

The analysis of the INTERIOR*Rendering* parameter revealed that the addition of interior geometry (*shaded* or *hatched*) is preferred by the participants and *hatched* modifications outperform *constant* shading in terms of error, confirming **H2** as the combinations using *hatching* reduced the alignment error.

SHADOW representation interacts with the EXTERIOR method, however the descriptives in Fig. 6 suggest for *hole*, *bright* works best, *ghosted* works well with a *bright* shadow and *masked* cannot be improved by *chromatic* shadows. This only partially supports **H3** as the visualizations with *chromatic* shadows helped to improve the alignment error only in some cases. *Chromatic* shadows showed a positive effect for SEQ, therefore, it could be beneficial to apply it in circumstances where usability is of strong importance.

From the descriptives presented in Fig. 6, we derive two general recommendations in the average error, yet not easily shown through statistical analysis: the HOLE exterior seems to benefit most from brightening of the dark colors, and in that case INTERIOR rendering plays only a minor role. When considering the use of GHOSTED exterior, one should consider combining it with a HATCHED interior, in which case the shadow color does not have a strong influence.

5 GENERAL DISCUSSION

We have provided a direct comparison of VST and OST displays in terms of depth perception with F+C techniques, where objects are shown below a real object's surface. Our

comparison between HMDs in Study 1 shows that depth judgments using F+C techniques for in-situ visualization are more accurate with VST HMDs.

Further, we found that the proposed techniques have different impact on the estimation of depth when using the same display technology. A decomposed analysis of these visualization techniques in our second study showed that *interior hatching* can provide useful cues to improve the estimation of depth with OST HMDs, as confirmed by both signed and absolute errors. Furthermore, the use of *chromatic shadows* showed significant improvements for subjective scores.

Results from our Study 2 show that the two novel visual components proposed in this work reduce the perceived complexity of the task and increase the visual attractiveness of the augmentations on AR displays without increasing and in many cases decreasing error metrics in the estimation of depth during perceptual matching tasks. This suggests that *interior hatching* and *chromatic shadows* can be effectively used to improve in-situ visualization with OST HMDs. Our studies also provide evidence that the masking method proposed by Otsuki *et al.* [42] presents adequate depth cues and produce similar alignment errors as the other methods. However, the subjective scores from both studies indicate that users do not find the masking visually appealing and they seem to perceive the perceptual matching task to be harder when using this technique. Thus, based on our results and observations, EXTERIOR rendering approaches such as the *window* or *ghosting* should be preferred for the presentation of in-situ content at near-field distances in AR.

Overall, results from our study showed that the virtual objects tend to be consistently perceived further away from the observer when using OST HMDs. This is derived from the fact that participants placed the virtual object closer to themselves than the real target. These results differ from the findings reported in studies that explored how physical occluders affect the perceived depth of virtual content at near-field distances in AR [52]. A significant difference between these studies is that their experimental setup used a textured occluder with high saliency while our monochromatic 3D printed head only had few salient features. This suggests that the texture of the occluder influences the user's perception when estimating the depth of virtual content placed behind a real object. However, our study protocol was not designed to explore the effects of the occluder texture over the estimation of depth, but rather to provide a comparative view on the visualization techniques proposed in this work. Therefore, further investigations need to be conducted in this regard.

5.1 Limitations

We intentionally limited our studies by not considering visualization techniques that modify the way the augmented object is rendered, which could represent another factor to be considered in a factorial design. Techniques like depth-encoding outlines [20] or pseudo chromadepth [49] have been shown to effectively aid perception [24] and interesting combinations with our hatching techniques might provide additional benefits. Further interesting aspects include different cutaway geometries [36], other illustrative surface shading techniques [34] or even animated surface visualizations. In this regard, we constrained

our second study to a select number of parameters to avoid participant fatigue in a prolonged experiment. It is also important to note that in the presented studies, our implementation considered the adjustment of the brightness as a means to ensure a minimum luminance and enhance the visibility of the virtual content when using the OST displays. Recent studies have shown that the brightness of the virtual content influences the accuracy achieved in near-field depth matching tasks [53]. However, these studies involved the judgement of virtual content that was not presented inside real objects. Therefore, further studies need to be conducted in this regard.

It is important to emphasize that the interpretation of our results is limited to the specific characteristics of the headsets used. Different headsets cover a wide range of intrinsic parameters that can influence depth perception and therefore results might vary for other devices with different FoV, focal planes, display resolutions, brightness and other characteristics. In the context of our studies, the comparably low angular resolution of the VST device used might have negatively impacted how the *Virtual Mask* performs, as the mask cutouts caused some aliasing at the pixel boundaries and we suspect that at least the subjective metrics could improve with a higher resolution. Furthermore, we expect that additive displays with different overall brightness will likely benefit from adapting the luminance of the shadow color C_s of the SHADOW representations accordingly. Our visualizations do not strongly rely on color accuracy of the displays, however it will be an interesting point for future research to investigate whether spatial perception is affected by displays with limited color uniformity like the display employed by the Hololens 2. In addition, VST and OST differ fundamentally with respect to accommodation, making a direct comparison between the two relatively complex. While in VST the virtual and real content observed is consistent in terms of accommodation, focusing on virtual and real objects in OST can potentially require different accommodation, leading to another inconsistent depth cue.

Moreover, the size of the cohorts recruited for our two studies was driven by considering the balanced randomization of the experiments through a Latin-Square Matrix in a repeated measure fashion. Considering the complexity of the models and the multiple levels investigated in our studies, future studies should extend the present work with larger samples, focusing on combinations and factor interactions, to further substantiate our results.

Lastly, our scenario is limited to the near-field, that represents a prototypical distance for medical scenarios. As a result of this, we did not investigate depth estimation at mid-to-far distances, larger separation between the target positions, nor larger object geometries.

6 CONCLUSION

Developing techniques for in-situ visualization is often guided by a specific application case with unique constraints. The conceptual decomposition of *F+C* visualization techniques presented in this work can be used to explore the design space of such AR visualizations for both OST and VST technologies, potentially adding new variants or novel dimensions to the proposed scheme.

In this regard, we believe that users benefit from the adaptation of visualization techniques designed to provide optimal visual cues for specific tasks not only in the context of medical applications, but also for various near-field AR approaches that require the in-situ visualization of structures within real objects.

Therefore, we hope that making explicit the visual dimensions involved in the design of the visualization techniques will assist researchers in the development of novel techniques suited to specific needs.

Moreover, we hope that the structured analysis used to develop and evaluate the new visualizations techniques presented in this work, can also serve as a model for future extensions and contribute to form the basis for future investigations of in-situ visualization techniques.

REFERENCES

[1] B. Michael, F. Henry, and O. Ryutarou, "Merging virtual objects with the real world: Seeing ultrasound imagery within the patient," *ACM SIGGRAPH Comput. Graph.*, vol. 26, no. 2, pp. 203–210, 1992.

[2] C. Bichlmeier, F. Wimmer, S. M. Heining, and N. Navab, "Contextual anatomic mimesis hybrid in-situ visualization method for improving multi-sensory depth perception in medical augmented reality," *in Proc. 6th IEEE/ACM Int. Symp. Mixed Augmented Reality*, 2007, pp. 129–138.

[3] S. Bjork and J. Redstrom, "Redefining the focus and context of focus+context visualization," in *Proc. IEEE Symp. Inf. Vis.*, 2000, pp. 85–89.

[4] E. A. Bustamante and R. D. Spain, "Measurement invariance of the Nasa TLX," *Proc. Hum. Factors Ergonom. Soc. Annu. Meeting*, vol. 52, no. 19, pp. 1522–1526, 2008.

[5] J. E. Cutting and P. M. Vishton, "Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth," in *Perception of Space and Motion*, Cambridge, MA, USA: Academic Press, 1995, pp. 69–117.

[6] Cutting, J. E., "Reconceiving perceptual space," in *Looking Into Pictures: An Interdisciplinary Approach to Pictorial Space*, Cambridge, MA, USA: MIT Press, 2003, pp. 215–238.

[7] D. Drascic and P. Milgram, "Perceptual issues in augmented reality," in *Proc. Stereoscopic Displays Virtual Reality Syst. III*, 1996, pp. 123–134.

[8] P. J. Edwards, L. G. Johnson, D. J. Hawkes, M. R. Fenlon, A. J. Strong, and M. J. Gleeson, "Clinical experience and perception in stereo augmented reality surgical navigation," *in Proc. Int. Workshop Med. Imag. Virtual Reality*, 2004, pp. 369–376.

[9] S. R. Ellis and B. M. Menges,. "Localization of virtual objects in the near visual field," *Hum. Factors: J. Hum. Factors Ergonom. Soc.*, vol. 40, no. 3, pp. 415–431, 1998.

[10] N. Elmqvist, U. Assarsson, and P. Tsigas, "Employing dynamic transparency for 3D occlusion management: Design issues and evaluation," in *Proc. IFIP Conf. Hum.-Comput. Interact.*, 2007, pp. 532–545.

[11] T. Engen *et al.*, "I psychophysics: Discrimination and detection. II. psychophysics: Scaling methods," in *Woodworth & Schlosberg's Experimental Psychology*. 3rd ed. J. W. Kling & Lorrin A. Riggs, and 17th contributors. New York, USA: Holt, Rinehart, and Winston, 1971, pp. 11–86.

[12] J. A. Ferwerda, "Psychophysics 101: How to run perception experiments in computer graphics," in *Proc. ACM SIGGRAPH Classes*, 2008, pp. 1–60.

[13] H. Fuchs *et al.*, "Augmented reality visualization for laparoscopic surgery," in *Proc. Int. Conf. Med. Imag. Comput. Comput.-Assis. Intervention*, 1998, pp. 934–943.

[14] C. Furmanski, R. Azuma, and M. Daily, "Augmented-reality visualizations guided by cognition: Perceptual heuristics for combining visible and obscured information," in *Proc. Int. Symp. Mixed Augmented Reality*, 2002, pp. 215–320.

[15] J. L. Gabbard, J. E. Swan, D. Hix, S. Kim, and G. Fitch, "Active text drawing styles for outdoor augmented reality: A user-based study and design implications,' in *Proc. IEEE Virtual Reality Conf.*, 2007, pp. 35–42.

[16] J. L. Gabbard, J. E. Swan, J. Zedlitz, and W. W. Winchester, "More than meets the eye: An engineering study to empirically examine the blending of real and virtual color spaces,' in *Proc. IEEE Virtual Reality Conf.*, 2010, pp. 79–86.

[17] G. A. Gescheider, *Psychophysics: Method, Theory, and Application*. Hillsdale, NJ, USA: Lawrence Erlbaum Associates, 1985.

[18] G. A. Gescheider, *Psychophysics: The Fundamentals*. London, U.K.: Taylor and Francis, 2013.

[19] S. Ghasemi, M. Otsuki, P. Milgram, and R. Chellali, "Use of random dot patterns in achieving X-ray vision for near-field applications of stereoscopic video-based augmented reality displays," *Presence*, vol. 26, no. 1, pp. 42–65, Feb. 2017.

[20] C. Hansen, J. Wieferich, F. Ritter, C. Rieder, and H.-O. Peitgen, "Illustrative visualization of 3D planning models for augmented reality in liver surgery," *Int. J. Comput. Assis. Radiol. Surg.*, vol. 5, no. 2, pp. 133–141, 2010.

[21] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (task load index): Results of empirical and theoretical research," *Adv. Psychol.*, vol. 52, pp. 139–183, 1988.

[22] S. G. Hart, "NASA-task load index (NASA-TLX); 20 years later," in *Proc. Hum. Factors Ergonom. Soc. Annu. Meeting*, vol. 50, pp. 904–908, 2006.

[23] H. Hauser, "Generalizing focus+context visualization," *Sci. Vis.: Vis. Extraction Knowl. Data*, pp. 305–327, 2006.

[24] H. Florian, K. Bornemann, K. Lawonn, and C. Hansen, "Depth perception in projective augmented reality: An evaluation of advanced visualization techniques," in *Proc. 25th ACM Symp. Virtual Reality Softw. Technol.*, 2019, pp. 1–11.

[25] J. L. Hintze and R. D. Nelson, "Violin plots: A box plot-density trace synergism," *Amer. Stat.*, vol. 52, no. 2, pp. 181–184, 1998.

[26] S. Ishihara, *Test for Colour-Blindness*. Tokyo, Japan: Kanehara Shuppan Co., 1985.

[27] J. A. Jones, J. E. Swan, G. Singh, E. Kolstad, and S. R. Ellis, "The effects of virtual reality, augmented reality, and motion parallax on egocentric depth perception," in *Proc. 5th Symp. Appl. Perception Graph. Vis.*, 2008, pp. 9–14.

[28] J. A. Jones, J. E. Swan, G. Singh, and S. R. Ellis, "Peripheral visual information and its effect on distance judgments in virtual and augmented environments," in *Proc. ACM SIGGRAPH Symp. Appl. Perception Graph. Vis.*, 2011, pp. 29–36.

[29] D. Kalkofen, E. Mendez, and D. Schmalstieg, "Interactive focus and context visualization for augmented reality," in *Proc. 6th IEEE ACM Int. Symp. Mixed Augmented Reality*, 2007, pp. 191–201.

[30] D. Kalkofen, M. Tatzgern, and D. Schmalstieg, "Explosion diagrams in augmented reality," in *Proc. IEEE Virtual Reality Conf.*, 2009, pp. 71–78.

[31] H. K. Kim, J. Park, Y. Choi, and M. Choe, "Virtual reality sickness questionnaire (VRSQ): Motion sickness measurement index in a virtual reality environment," *Appl. Ergonom.*, vol. 69, pp. 66–73, 2018.

[32] E. Kruijff, J. E. Swan, and S. Feiner, "Perceptual issues in augmented reality revisited," in *Proc. IEEE Int. Symp. Mixed and Augmented Reality*, 2010, pp. 3–12.

[33] K. Lawon, M. Luz, B. Preim, and C. Hansen, "Illustrative visualization of vascular models for static 2D representations," in *Proc. Int. Conf. Med. Imag. Comput. Comput.-Assis. Intervention*, 2015, pp. 399–406.

[34] K. Lawonn, I. Viola, B. Preim, and T. Isenberg, "A survey of surface-based illustrative rendering for visualization," *Comput. Graph. Forum*, vol. 37, no. 6, pp. 205–234, 2018.

[35] M. Lerotic, A. J. Chung, G. Mylonas, and G.-Z. Yang, "PQ-space based non-photorealistic rendering for augmented reality," in *Proc. Int. Conf. Med. Imag. Comput. Comput.-Assis. Intervention*, 2007, pp. 102–109.

[36] W. Li, L. Ritter, M. Agrawala, B. Curless, and D. Salesin, "Interactive cutaway illustrations of complex 3D models," *ACM Trans. Graph.*, vol. 26, no. 3, pp. 31–es, 2007.

[37] M. A. Livingston, J. L. Gabbard, J. E. SwanII, C. M. Sibley, and J. H. Barrow, "Basic perception in head-worn augmented reality displays," *Human Factors in Augmented Reality Environ.*, New York, NY, USA: Springer, 2013, pp. 35–65.

[38] M. A. Livingston, A. Dey, C. Sandor, and B. H. Thomas, "Pursuit of X-ray vision for augmented reality," *Human Factors in Augmented Reality Environments*, New York, NY, USA: Springer, 2013, pp. 67–107.

[39] D. Medeiros, M. Sousa, D. Mendes, A. Raposo, and J. Jorge, "Perceiving depth: Optical versus video see-through," in *Proc. 22nd ACM Conf. Virtual Reality Softw. Technol.*, 2016, pp. 237–240.

[40] S. Nagata, "How to reinforce perception of depth in single two-dimensional pictures," in *Pictorial Commun. in Real Virtual Environ.*, 1991, pp. 527–545.

[41] St. Nicolau, L. Soler, D. Mutter, and J. Marescaux, "Augmented reality in laparoscopic surgical oncology," *Surg. Oncol.*, vol. 20, no. 3, pp. 189–201, 2011.

[42] M. Otsuki and P. Milgram, "Psychophysical exploration of stereoscopic pseudo-transparency," in *Proc. IEEE Int. Symp. Mixed Augmented Reality*, 2013, pp. 283–284.

[43] M. Otsuki, H. Kuzuoka, and P. Milgram, "Analysis of depth perception with virtual mask in stereoscopic AR," in *Proc. 25th Int. Conf. Artif. Reality Telexistence 20th Eurograph. Symp. Virtual Environments*, 2015, pp. 45–52.

[44] E. Peillard, Y. Itoh, G. Moreau, J. -M. Normand, A. Lyer, and F. Argelaguet, "Can retinal projection displays improve spatial perception in augmented reality?," in *Proc. 19th IEEE Int. Symp. Mixed Augmented Reality*, 2020, pp. 80–89.

[45] P. E. Pelargos *et al.*, "Utilizing virtual and augmented reality for educational and clinical enhancements in neurosurgery," *J. Clin. Neurosci.*, vol. 35, pp. 1–4, 2017.

[46] E. Praun, H. Hoppe, M. Webb, and A. Finkelstein, "Real-time hatching," in *Proc. 28th Annu. Conf. Comput. Graph. Interactive Techn.*, 2001, pp. 581–586.

[47] R. S. Renner, B. M. Velichkovsky, and J. R. Helmert, "The perception of egocentric distances in virtual environments - A review," *ACM Comput. Surveys*, vol. 46, no. 2, pp. 1–40, 2013.

[48] J. P. Rolland, C. Meyer, K. Arthur, and E. Rinalducci, "Method of adjustments versus method of constant stimuli in the quantification of accuracy and precision of rendered depth in head-mounted displays," *Presence: Teleoperators Virtual Environ.*, vol. 11, no. 6, pp. 610–625, 2002.

[49] T. Ropinski, F. Steinicke, and K. Hinrichs, "Visually supporting depth perception in angiography imaging," in *Proc. Int. Symp. Smart Graph.*, 2006, pp. 90–104.

[50] G. Schall *et al.*, "Handheld augmented reality for underground infrastructure visualization," *Pers. Ubiquitous Comput.*, vol. 13, no. 4, pp. 281–291, 2009.

[51] T. Sielhorst, C. Bichlmeier, S. M. Heining, and N. Navab, "Depth perception—A major issue in medical AR: Evaluation study by twenty surgeons," in *Proc. Int. Conf. Med. Imag. Comput. Comput.-Assis. Intervention*, 2006, pp. 364–372.

[52] G. Singh, J. E. Swan, J. A. Jones, and S. R. Ellis, "Depth judgment measures and occluding surfaces in near-field augmented reality," in *Proc. 7th Symp. Appl. Perception Graph. Vis.*, 2010, pp. 149–156.

[53] G. Singh, S. R. Ellis and J. E. Swan, "The effect of focal distance, age, and brightness on near-field augmented reality depth matching," *IEEE Trans. Vis. Comput. Graph.*, vol. 26, no. 2, pp. 1385–1398, Feb. 2020.

[54] V. Šoltészová, D. Patel, and I. Viola, "Chromatic shadows for improved perception," in *Proc. ACM SIGGRAPH/Eurograph. Symp. Non-Photorealistic Animation Rendering*, 2011, pp. 105–116.

[55] J. E. Swan, A. Jones, E. Kolstad, M. A. Livingston, and H. S. Smallman, "Egocentric depth judgments in optical, see-through augmented reality," *IEEE Trans. Vis. Comput. Graph.*, vol. 13, no. 3, pp. 429–442, May/Jun. 2007.

[56] J. E. Swan, G. Singh, and S. R. Ellis, "Matching and reaching depth judgments with real and augmented reality targets," *IEEE Trans. Vis. Comput. Graph.*, vol. 21, no. 11, pp. 1289–1298, Nov. 2015.

[57] R. Wang, Z. Geng, Z. Zhang, and R. Pei, "Visualization techniques for augmented reality in endoscopic surgery," in *Proc. Int. Conf. Med. Imag. Augmented Reality*, 2016, pp. 129–138.

[58] W. Wetzlinger, A. Auinger, and M. Drflinger, "Comparing effectiveness, efficiency, ease of use, usability and user experience when using tablets and laptops," in *Proc. Int. Conf. Des., User Experience Usability*, 2014, pp. 402–412.

[59] S. Zollmann, D. Kalkofen, E. Mendez, and G. Reitmayr, "Image-based ghostings for single layer occlusions in augmented reality," in *Proc. IEEE Int. Symp. Mixed Augmented Reality*, 2010, pp. 19–26.

Alejandro Martin-Gomez received the BSc degree in electronic engineering from Instituto Technologico de Aguascalientes, Mexico, the MSc degree in electronic engineering from Universidad Autonoma de San Luis Potosi, Mexico, and is currently working toward the PhD degree with the Department of Informatics of the Technical University of Munich. He is currently a research assistant with Johns Hopkins University. His main research interests include the improvement of visual perception for augmented reality and its applications in interventional medicine. He was the Chair for Computer Aided Medical Procedures and Augmented Reality. He was the recipient of the Academic Excellence Award by the National Association of Schools and Faculties of Engineering, Mexico.

Jakob Weiss received the BSc and MSc degrees in informatics from the Technical University of Munich in 2013 and 2016, respectively. Between 2016 and 2020, he was a research scientist and a doctoral candidate with the TUM Chair for Computer Aided Medical Procedures. He is currently a visualization specialist with OneProjects Design and Innovation GmbH, Munich, Germany. His research interests include improving guidance systems for medical applications through visual computing, including topics like medical volume processing and rendering, computer graphics, and augmented and virtual reality.

Andreas Keller received the BSc degree in informatics from the Technical University of Munich in 2020 and is currently working toward the MSc degree in informatics. He is currently a research assistant with the Technical University of Munich, Germany. His main research interests include computer graphics and virtual and augmented reality technologies.

Ulrich Eck was born in Munich, Germany. He received the PhD degree in computer and information science from the Magic Vision Lab, University of South Australia. He is currently leading the research activities with the NARVIS laboratory of the Chair for Computer Aided Medical Procedures and Augmented Reality at the Technical University of Munich. His research interests include multimodal augmented reality systems for medical procedures, training, and education.

Daniel Roth received the master's degree in media and imaging technology from the University of Applied Sciences, Cologne, TH Köln and the doctoral degree in computer science in 2019 from the University of Würzburg, Germany, where the main subjects included intrapersonal, interpersonal, and hybrid aspects of virtual reality. He is currently an assistant professor of human-centered computing and extended reality with the Department of Artificial Intelligence in Biomedical Engineering, Friedrich-Alexander Universität Erlangen-Nürnberg. He was a senior research scientist with the Technical University of Munich, Germany. His research interests include human-computer interaction, virtual- and augmented reality, and AI in the context of physical and mental health as well as computer aided interventions.

Nassir Navab received the PhD degree from INRIA and University of Paris XI, France. He is currently a full professor and the director of the Laboratory for Computer-Aided Medical Procedures, Johns Hopkins University and the Technical University of Munich. He also has secondary faculty appointments with both affiliated Medical Schools. He was a postdoctoral fellow with MIT Media Laboratory before joining Siemens Corporate Research (SCR) in 1994. He is the author of hundreds of peer-reviewed scientific papers, with more than 36,000 citations and an h-index of 88 as of November 2020, the author of more than 30 awarded papers, including 11 at MICCAI, five at IPCAI, and three at the IEEE ISMAR, and the inventor of 50 granted US Patents and more than 50 International ones. His current research interests include medical augmented reality, computer-aided surgery, medical robotics, and machine learning. He was a distinguished member with SCR. He was the recipient of the Siemens Inventor of the Year Award in 2001, the SMIT Society Technology Award in 2010 for the introduction of Camera Augmented Mobile C-arm and Freehand SPECT technologies, and the 'ten years lasting impact Award' of the IEEE ISMAR in 2015. In 2012, he was elected as a Fellow of the MICCAI Society. From 2007 to 2012 and 2014 to 2017, he was a member of the board of directors of the MICCAI Society and is currently on the steering committee of the IEEE Symposium on Mixed and Augmented Reality and Information Processing in Computer-Assisted Interventions.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/csdl.