

Perception of Volumetric Characters' Eye-Gaze Direction in Head-Mounted Displays

Andrew MacQuarrie, Anthony Steed*

Department of Computer Science, University College London, United Kingdom



Figure 1: The seven avatar angles (AAs): volumetric reconstructions of a human with eye gaze directed at various angles. AAs are reported with angles to the avatar's left being negative. Above, left to right, these models show the avatar looking 45°, 30°, 15°, 0°, -15°, -30° and -45°.

ABSTRACT

Volumetric capture allows the creation of near-video-quality content that can be explored with six degrees of freedom. Due to limitations in these experiences, such as the content being fixed at the point of filming, an understanding of eye-gaze awareness is critical. A repeated measures experiment was conducted that explored users' ability to evaluate where a volumetrically captured avatar (VCA) was looking. Wearing one of two head-mounted displays (HMDs), 36 participants rotated a VCA to look at a target. The HMD resolution, target position, and VCA's eye-gaze direction were varied. Results did not show a difference in accuracy between HMD resolutions, while the task became significantly harder for target locations further away from the user. In contrast to real-world studies, participants consistently misjudged eye-gaze direction based on target location, but not based on the avatar's head turn direction. Implications are discussed, as results for VCAs viewed in HMDs appear to differ from face-to-face scenarios.

Keywords: User study, Virtual reality, Gaze perception

Index Terms: 1.1.1 [Human-centered computing]—Empirical studies in HCI; 1.1.3.2 [Human-centered computing]: Interaction devices—Displays and imagers; 1.1.4 [Human-centered computing]—HCI theory, concepts and models

1 INTRODUCTION

Volumetric capture is the creation of visual media content that can be explored with six degrees of freedom (6DoF), i.e. content that exhibits motion parallax when the user moves their head. While games and other computer generated imagery also allow 6DoF exploration, volumetric capture allows such content to be created without the need for prior or post-hoc 3D modelling, as the content is filmed from real-world performances in a similar way to video. This allows the creation of very realistic media. There are several technologies that allow this type of content to be created. In this work, a “free-viewpoint video” approach was used, using a reconstruction technique similar to [10]. In free-viewpoint video, a character performance is filmed from multiple perspectives simultaneously. These views are then reconstructed into an animated textured 3D mesh.

*andrew.macquarrie.13@ucl.ac.uk, a.steed@ucl.ac.uk

Cinematic VR experiences, i.e. narrative-led VR, are starting to make use of volumetrically captured avatars (VCAs) (e.g. [27]). These experiences allow users to explore with 6DoF while maintaining near-video-quality realism of character performances. Additionally, the use of classic film-production techniques may make this pipeline more accessible to producers who have extensive film experience.

An issue with volumetric performances is that once they are captured the content is fixed. Specifically, the VCA's eye-gaze is fixed in whatever direction the actor was looking during filming. For VCAs who address the user, this eye-line may be broken as soon as the user changes position. This may damage the illusion of reality, or give the impression the avatar is unintelligent or uninterested. This issue relates to dyadic or mutual gaze – when people assess to what extent someone appears to be looking at them.

It is also important to understand how well users can establish triadic gaze, i.e. eye gaze directed towards an object of shared attention. To create VCAs using free-viewpoint video techniques, characters are frequently filmed separately using green-screen, and then assembled together into a complete scene in software. This technique often relies on carefully planning the scene in advance, and affixing eye-line markers outside of the capture volume to provide actors with targets to look at during filming. As volumetric filming often relies on long takes due to restrictions on cuts in VR, ensuring actors hit these targets while standing in exactly the right position can be difficult. To establish if such production limitations will impact the experience, it is important to know how accurately users can detect triadic eye gaze when watching these performances in a VR display.

In this work, we explore how well participants can assess the eye-gaze direction of a VCA when viewed in a HMD. While vertical eye gaze may also be an important consideration for VCAs, here we explore only horizontal eye-gaze perception. A repeated measures experiment was conducted with 36 participants, in which participants repeatedly rotated a VCA to look at a target. The VCA's gaze direction and the target location were varied. As it was expected that the clarity of the VCA's eyes might also have an impact, participants completed the tasks in two HMDs with different resolution screens.

Contrary to expectations, the HMD resolution was not found to significantly impact accuracy. The task became significantly harder for target locations further away from the user, while the gaze direction of the avatar was also found to impact accuracy. Absolute error angles, used to measure accuracy, increased 2.5 fold from the best to worst performing condition. In contrast to real-world studies, eye-gaze direction was consistently misjudged by approximately

10% based on target location, but not based on the avatar's head turn direction. We discuss the practical implications these unexpected results may have for volumetric VR content creation, and suggest that further investigation may help reveal details on the cognitive processes involved in eye-gaze understanding.

2 RELATED WORK

2.1 Volumetric capture

Captured, near-photorealistic content that supports motion parallax can be achieved in a number of ways. Light-field capture, in which a dense grid of cameras capture content that allows 6DoF viewing within a constrained region, has existed for decades [24]. The technology continues to develop today, with commercial camera companies such as Lytro producing content with impressive fidelity [22].

In this work, we use an alternative form of 3D capture called free-viewpoint video. Specifically, our content was filmed with Dimension Studio [12]. Dimension Studio are a Microsoft partner, and use a volumetric pipeline similar to that described in [10]. In this technique, a central performance is captured by a surrounding array of inward-facing RGB and infrared (IR) cameras which form a cylindrical shape. The space inside this cylinder is referred to as the "capture volume". Chroma key compositing, otherwise known as "green screen", is used in tandem with depth-from-stereo techniques to segment the performance from the background for every camera view. 3D reconstruction algorithms are then used to create a 3D mesh of the performance for every frame. This mesh is then textured using data from the RGB cameras.

To ensure repeatability, in the supplementary materials we provide the 3D reconstructed model for every direction the avatar looked in our experiment.

2.2 Assessment of eye-gaze direction

A number of studies have investigated perception of eye-gaze direction. Almost all of these are based on three seminal papers from the 1960s. While different works use different terminology to refer to the participant assessing eye gaze and the person whose eye gaze is being assessed, we refer to them here as the "receiver" and the "sender" respectively.

In 1963, Gibson and Pick published an experiment that explored mutual gaze, i.e. gaze in which two people are looking at each other [17]. In their work, the receiver had to assess if they felt the sender was looking at them, while the sender looked at seven targets spaced 10cm apart horizontally across the bridge of the receiver's nose. The sender and the receiver were 2m apart. This was repeated with the sender's face facing forward, as well as turned left 30° and right 30°. Their results indicated that the head turn had a significant effect on where the receiver felt the user was looking, with the mean value (what they termed "common error" (CE)) being shifted in the direction of the head turn. They did not find a difference in acuity between head directions, as assessed by comparing standard deviations (SDs). Gibson and Pick assessed that, for dyadic gaze, humans can perceive 1mm differences in the position of the iris at a distance of 2m. This was considered to represent a "precision" (standard deviation of the responses for which the receiver felt they were being looked at) of 2.8° [6].

In 1967, Cline performed experiments exploring human acuity in triadic gaze tasks, i.e. tasks in which the sender is looking at a mutually visible target not located by the receiver's eyes [9]. Triadic gaze is generally considered harder for humans to assess, perhaps because mutual gaze can be determined via symmetry in the eyes. Similarly to Gibson and Pick, Cline found that the direction of the sender's head created a shift in the CE. Unlike Gibson and Pick, however, Cline noted that an increased divergence between the sender's head and eye directions also lowered acuity, as assessed by an increase in SD values. The precision Cline reported has been

assessed to be slightly better than that of Gibson and Pick, at 0.7° for mutual gaze and 3.1° for triadic [6]. In Cline's work, the sender and the receiver were 1.22m apart.

The work of Anstis et al. in 1969 explored eye-gaze perception with no shared object of attention, i.e. the receiver could not see the target the sender was looking at [2]. Similarly to previous work, a CE was found based on the sender's head-turn direction. While there has been some disagreement in the literature about in which direction this CE shift occurs (for a review, see [28]), we follow the conclusion of [28] and assume the shift in CE is in the direction of the sender's head turn. Anstis et al. found that users tended to overestimate the CE for triadic gaze, regardless of head turn.

Multiple other studies have improved our understanding of eye-gaze perception. It has been shown that the same acuity is achieved for static stimuli as for stimuli in which the sender fixating the target was visible to the receiver [25,38]. A study by Ellgring and Cranach indicated that the receiver could improve their accuracy with practice, with this effect being strongest for fixation points away from the receiver's face [14]. Mutual gaze perception was found to be asymmetric by Chen, with participants being less sensitive to eye contact when the sender looked below their eye than above or to the side of the eye [8]. Illumination conditions have been shown to impact the perceived orientation of the face, which may also impact eye-gaze assessment [40,44]. While Anstis et al. argued that eye-gaze direction is determined by assessment of the discrepancy between the sender's iris/sclera configuration, others have cast doubt on this [25,38]. Despite multiple studies investigating eye gaze assessment in humans, to our knowledge there is currently no accepted model that underpins how we interpret where a person is looking.

2.3 Eye gaze in telecommunication

Eye gaze is one of the most powerful non-verbal communication cues that humans use, and a number of social psychology researchers have explored gaze patterns during human interactions [3,21]. These works have established that eye gaze is a rich communication tool, encompassing multiple functions such as controlling turn-taking during conversations, communicating emotion, providing feedback and establishing objects of mutual attention.

Eye contact has also been shown to be an important factor in mediated communications [43]. It has been suggested that eye gaze is important for inferring the hidden mental states of others [4]. Inaccurate eye-gaze portrayals have been shown to impact users ability to detect lies [37] and decrease trust [29,32].

While videoconferencing is increasingly popular, such communication mechanisms have been shown to offer poor spatial portrayals of eye-gaze [1,41]. Significant work has been done on eye-gaze correction in videoconferencing (e.g. [19,42,45]). Despite these works, correction of eye gaze in telecommunications remain uncommon in practice. This may indicate that these techniques do not yet fully facilitate natural eye-gaze communication.

Although this work does not use real-time volumetric capture, recent advances have shown that volumetric techniques are possible in real-time [30]. This opens up the possibility of using this technology for telecommunication, although currently this speed comes at the cost of fidelity. For use in telecommunications, volumetric reconstruction may offer an opportunity to communicate naturally using multiple non-verbal cues over large distances.

2.4 Eye gaze in virtual reality

A number of papers have explored multiple factors around the impact of eye gaze in computer-generated avatars. Work by Cassell et al. explored the importance of eye-gaze direction in computer-generated conversational agents [7]. Realistic saccade behavior in avatars has been modeled using eye-tracking data [23,26]. Studies have also shown the importance of eye-lid behavior, with blinks [39] and realistic lid saccades [35,36] both improving realism.

Garau et al. explored how gaze is used to inform turn-taking during conversations with computer-generated avatars on screens and VR displays [15, 16]. In these works, avatars that exhibited human-like turn-taking behavior outperformed avatars with random eye gaze on metrics such as conversational involvement and avatar realism.

Roberts et al. compared collaborative VR against video conferencing for their ability to communicate eye gaze [33]. Avatars were computer generated in real time, with the avatar’s eye direction being constructed using an eye tracker on the sender. They concluded that, while video conferencing can be configured to support eye gaze in constrained situations, VR allowed movement of the sender and receiver while maintaining gaze-based communication.

Pan et al. explored participant’s ability to assess eye-gaze direction of a filmed human when presented on a situated display [31]. A flat and spherical display were compared. Their results indicated that a spherical display provided the best performance outside of the real-world condition. However, their displays were driven by 11 cameras arranged 15° apart around the sender, and no 3D reconstruction was used. The video feed from the camera closest to the receiver’s position was displayed. As a result, the display did not fully support 6DoF viewing.

While these works explore the impact of eye gaze in VR settings, they all use either computer generated avatars or simple video feeds of the sender. Evaluation of the eye-gaze direction of volumetrically captured avatars has previously been investigated by Roberts et al. in their work “Estimating the Gaze of a Virtuality Human” [34]. In this work, participants rotated a volumetrically captured 3D reconstruction of a human until they felt he was looking at them. In this way, this experiment only explored mutual gaze. The avatar was displayed life-size on a screen in front of the participant. As motion tracking of the receiver’s head was not used, however, this task does not support 6DoF viewing. It was argued that, as the avatar was displayed on a flat screen, the Mona Lisa effect would mean participant’s impression of where the avatar was looking would not be impacted by their movement. The Mona Lisa effect, in which the gaze of a human in a painting or photograph appears to follow you around the room, occurs because the viewer’s perspective of the painting does not update the painting’s viewport. This is somewhat similar to looking at a character on TV, so arguably does not represent how humans will perceive eye-gaze direction in VR contexts, as the viewport of the VR display updates as the user moves.

3 METHOD

As discussed in section 1, VCA performances are currently fixed at the point of filming. Due to the issues associated with using eye-line targets during filming discussed in section 1, it is important to develop an understanding of HMD users’ ability to assess the eye-gaze direction of VCAs. The impact of the display in use is also important, as these experiences are often consumed on HMDs where limited resolutions may impact a user’s ability to judge eye-gaze direction.

To explore the perception of eye gaze in HMDs, a repeated measures experiment was conducted. An avatar was created using volumetric reconstruction, as discussed in the section 3.3. Participants were asked to rotate this avatar using a hand controller until they felt it was looking at a target. The target was represented as a yellow cylinder, floating at eye-height 1m away from the avatar. The experiment explored the relationship between the location of the target, the angle of the VCA’s gaze direction, and the resolution of the HMD.

3.1 Subjects

The study was approved by the UCL Research Ethics Committee (project ID 4547/012). All participants were recruited via a participant pool website. A total of 37 participants took part, although data from one was excluded due to issues with depth perception. Of

the remaining 36 participants, 18 were male and 18 female. The mean age of participants was 29.4 (SD=10.3). All participants self-assessed as having normal or corrected to normal vision.

3.2 Independent variables

3.2.1 Head-mounted display

Previous work has indicated that the perception of eye gaze in humans is highly accurate (e.g. [9, 17]). To achieve this level of accuracy, it has been argued that humans must rely on the perception of tiny differences in the location of the iris. For example, Cline proposed that at a distance of 1.22m, participants were able to discern a 0.18mm displacement of the iris during dyadic gaze perception (an angle of 0.01° from the receiver’s position) [9]. The requirement of accurate representations of the iris indicated that the resolution of the HMD may impact the assessment of eye gaze in these displays. To explore the impact of HMD resolution on gaze perception, the task was performed in two HMDs with different resolutions.

The Vive and the Vive Pro were chosen as these two devices can be considered functionally equivalent in many regards outside of the screen resolution. As the names of these two devices is similar and could lead to confusion, we refer to them throughout as the “Regular Vive” (REG) and the “Vive Pro” (PRO). The REG has a resolution of 2160x1200 (448 PPI) while the PRO has a resolution of 2880x1600 (615 PPI). At a distance of 2m from the participant, each of the avatar’s eyes subtends approximately 0.86° horizontally. We calculate that the 3cm by 1cm area around each of the eyes will be displayed using approximately 55 pixels in the REG and 100 pixels in the PRO. As the PRO is backwards compatible with the tracking system and input controllers of the REG, these aspects of the two devices can be considered similar. The lighthouse tracking system version 1.0 was used for both HMDs. The left and right hand controllers of a pair of REG controllers were each paired to one of the two HMDs; as the controllers are ambidextrous, this did not present any usability issues. The optical systems of both devices are also similar, with both providing a 110° field-of-view using Fresnel lenses. As audio was not required for the task, the built-in headphones of the PRO were removed.

The major remaining difference between the two HMDs is the design of the head strap, with the PRO allowing users to adjust the fit more easily themselves using an adjustment dial on the back of the device. This difference should be considered, as the ease of use of the PRO allows users to ensure a good fit, and therefore perhaps a sharper image. Care was taken during the fitting of both HMDs, however, to ensure participants were comfortable and that the image quality was good. This process is described in section 3.5.

3.2.2 Avatar angle

To explore the impact of the angle of eye gaze with respect to the avatar, several models were used with the avatar looking at equally spaced targets. We call these the *avatar angles* (AAs), as they represent the different angles the avatar was looking. The following AAs were used: -45°, -30°, -15°, 0°, 15°, 30° and 45°. An image showing the seven AAs is shown in Figure 1.

3.2.3 Target angle

Separate from the angle of the eye-gaze direction of the avatar, the location of the target participants were asked to rotate the avatar to look at also varied in position. Targets were placed at -60°, -45°, -30°, -15°, 0°, 15°, 30°, 45° and 60° with respect to the avatar. We refer to these as the *target angles* (TAs). To reduce the impact of differences in the convergence mismatch – discussed further in section 3.3 below – target locations remained at a constant 1m distance from the point between the avatar’s eyes. TAs are reported such that negative angles are to the participant’s left, as shown in Figure 2.

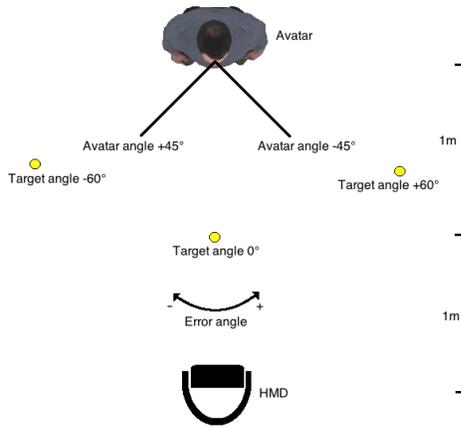


Figure 2: Diagram of angle directions (positive/negative) for target angles, avatar angles and error angles. Also shown is the distance between the virtual targets, avatar and user.

3.3 Stimuli

Volumetric reconstruction was used to create 3D models from a captured performance. The reconstruction methods used were similar to those described in [10]. A professional actor was hired for the capture, who was paid for his time and rights to the resulting media. Our actor self-reported as having balanced eyes (e.g. no strabismus). To ensure repeatability of our experiment, in the supplementary material we provide the 3D model and reconstructed eye-gaze ray for each AA model. Images of the avatar model looking to each of the AAs is shown in Figure 1. Although a volumetric video was reconstructed, we only presented frames of this video to participants (i.e. the stimuli were static).

During filming, gaze targets were placed on the wall of the studio just outside the capture volume. This is a limitation of the capture technology, as placing targets inside the volume may obscure camera views of the actor. The gaze targets were all approximately 3.5m from the actor during filming, while the virtual targets were 1m from the avatar. As the overall gaze direction may be being triangulated from the direction of the left and right eyes, this does introduce the possibility that the results are impacted by this convergence mismatch. However due to the large baselines involved between the eyes and the targets, and the relatively small distance between the eyes, the impact of this issue on the angular metrics is likely to be small. We discuss the implications of this mismatch in section 4.4.

3.4 Hypotheses

We explore two properties related to the error angle. First, we explore the mean error angle. It might be expected that mean error angles would sit around 0° , indicating that participant answers were evenly distributed around the correct value. Previous work, however, has indicated that guesses are often shifted in a consistent direction, leading to a non-zero mean (e.g. [9]). As a result, the mean of error angles is frequently referred to as the “common error” (CE), a convention that we follow here.

Secondly, we examine the absolute error angle. In previous work, the absolute error angle was taken to indicate accuracy, as higher absolute error angles indicate participants were less able to perceive where the avatar was looking [34]. We propose the following hypotheses for these metrics in regards to the independent variables:

3.4.1 HMD condition

It is expected that a higher resolution display will lead to an improvement in task accuracy, as improved detail in the eye is expected to lead to an improved ability to perceive eye-gaze direction. We

calculate that the 3cm by 1cm area around each of the avatar’s eyes will, at a distance of 2m, be displayed using approximately 55 pixels in the REG and 100 pixels in the PRO. The model texture for the same area is approximately 1925 pixels.

H1: Increased HMD resolution will lead to improved accuracy in task performance, measured by absolute error angles.

3.4.2 Target angle

In line with previous findings [9], we expect the CE to be 0° for all TAs. Although Anstis et al. found an overestimation of CE based on TA [2], this was for targets that were not visible to the receiver, and therefore represents a different scenario.

H2: CE will be 0° for all TAs.

Previous work has shown that humans are better at assessing gaze for targets closer to their eyes, e.g. Cline showed that accuracy decreased as TA increased [9]. We expect task difficulty to increase as TAs diverge from 0° .

H3: Gaze perception accuracy will decrease as TAs diverge from 0° , measured by absolute error angle.

3.4.3 Avatar angle

As indicated in previous studies (e.g. [2,9]), it is expected that the CE will be influenced by the head-turn direction of the avatar, which in our case is always in the direction of the avatar’s eye gaze. As a result, we expect to see a CE in error angle that is pulled in the direction of the head turn, i.e. negative for AAs greater than 0° , and positive for AAs below 0° . The direction of all angle types is shown in Figure 2.

H4: CE will show an overshoot effect in the direction of the avatar’s head turn.

Similarly to what Cline described [9], we expect that task difficulty will increase as the AA moves away from 0° , as the gaze and head directions diverge more strongly.

H5: Gaze perception accuracy will decrease as AAs diverge from 0° , measured by absolute error angle.

3.5 Procedure

Participants were asked to read an information sheet and complete a consent form and pre-experiment questionnaire. Details in the information sheet relating to simulator sickness and their right to stop at any time were then reinforced verbally. The task was then explained to participants, and the input controller demonstrated.

Participants were then asked to stand in a marked spot on the floor, facing in a specific direction. SteamVR was calibrated so standing in the given spot meant all participants were in the same position and orientation, centered within the SteamVR welcome menu. Participants were then given the first of the two HMDs, with the order of exposure to the PRO and the REG being counterbalanced through oscillation between subjects. Participants put on the first HMD, and were given the hand controller. Both HMDs were driven by the same desktop PC running Windows 10. The PC had a 3.4GHz Intel i7-6700 CPU, 32GB of RAM, and a NVIDIA GeForce GTX 980 graphics card. The software was implemented using version 2017.4.2f2 of the Unity game engine. This setup achieved a consistent frame rate of 90FPS.

Once participants had put on the first HMD, they were then asked to read the text in front of them in the SteamVR welcome menu. They were asked to comment on the clarity of the text, and if it was blurry were given advice about how to adjust the HMD location on the face to improve focus.

Both HMDs presented the environment stereoscopically. The interpupillary distance (IPD) of the REG was set to 64.5mm, while due to a setup error, the IPD of the PRO was set to 63.8mm. Participants were told that they were allowed to move their head from side to side, or even up and down, in order to help them identify the

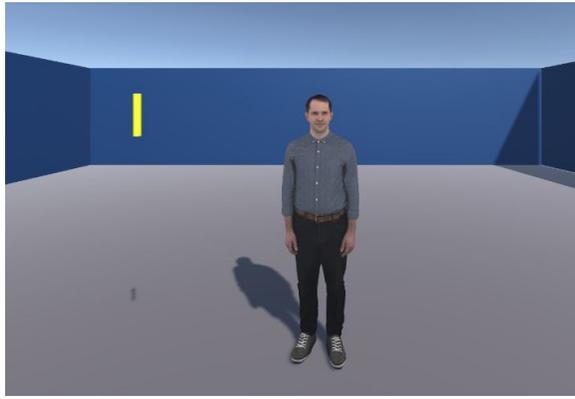


Figure 3: Image of the experimental task. The user was asked to rotate the avatar until they believed he was looking at the target, represented as a yellow cylinder floating at eye height. In the above image, the TA is -60° and the AA is -15° .

position of the target or to improve their understanding of the scene. Participants were told, however, not take any steps in any direction. The position of the HMD was recorded throughout the experiment.

On initialization, participants were asked to stand in a neutral position facing forward. When the software loaded, the HMD in the virtual scene was positioned to be at exactly eye level with the avatar, regardless of the height of the participant. This meant that the floor level of the virtual scene was not always consistent with the physical floor level. The avatar height was 1.73m.

Participants first started in a training scene, to allow them to experience the task and get used to the control mechanism. The training scene was identical to the experimental scene. An image of the training scene is shown in Figure 3. On the input controller, participants pushed down on the trackpad on one side to rotate the avatar one way, and on the other side of the trackpad to rotate the avatar the other way. Participants were only able to rotate the avatar on the horizontal axis. The avatar rotated around the point between its eyes for all AA models.

When the subject believed the avatar was looking at the target they pulled the trigger on the hand controller. This answer was recorded by the system. The target was represented as a yellow cylinder floating at eye height, 1m from the avatar, as shown in Figure 3. Once the system had recorded the participant's answer, the avatar was instantly rotated to a randomly selected rotation from -50° to $+50^\circ$, the target moved to a new TA, and the avatar model swapped for the next AA. Once the user had completed approximately three such tasks, and the experimenter was convinced the participant had a strong idea of how the task worked, the training session was ended. The participant remained in the HMD, and the experimental environment was loaded.

Participants then worked through all of the presented tasks. Each task represents a possible combination of the AAs and TAs. As there were seven AAs and nine TAs, this resulted in a total of 63 combinations. These combinations were presented in a randomized order. In all cases, the yellow cylinder representing the target was 1m from the avatar, and the distance between the avatar and the subject was 2m.

Once a participant had completed all 63 tasks in a given HMD, they were asked to remove the display. The experimenter checked that the participant was feeling well, and that they were experiencing no ill-effects from the display. No participants indicated any ill-effects. They were then asked to put on the second HMD, and the same procedure for adjusting the HMD was followed. Participants completed the tasks again in a randomized order, resulting in a total of 126 answers for both displays. Once all tasks were completed

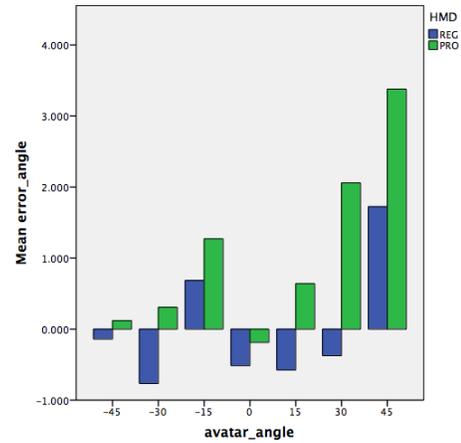


Figure 4: Histogram of common error values (mean of error angles) for all avatar angles. The histogram is split down by HMD. Values are averaged across all target angles using mean values.

the HMD was removed. Participants then answered some interview style questions on the experience. Participants were then debriefed, paid £10 for taking part and dismissed.

3.6 Method of Analysis

In line with previous work, we explore both the common error (CE) and task difficulty. The CE represents the overshoot effect established in the literature, and is indicated by a shift of the mean error angle away from 0° . Task difficulty, as in previous work (e.g. [18, 34]), was measured using the mean of absolute error angles.

4 RESULTS

A mixed models analysis of the data was explored. However, the errors did not have constant variance, as assessed by visual inspection of a plot of residuals versus predicted values. This may be caused by large differences in error angle variation between levels for certain independent variables. Attempts to transform the data did not adjust for this issue.

As a result, the data was analyzed using simpler analysis techniques. To facilitate these analysis techniques, data was aggregated over IVs using mean values. For example, when exploring the impact of the HMD resolution, error angles were aggregated across AAs and TAs.

4.1 Head-mounted display resolution

4.1.1 Common error

It was expected that any impact from the HMD on CE would be symmetric across the 0° AA. This was expected as any overshoot effects were expected to appear equally on both sides, i.e. an overshoot effect for an AA of -30° would appear in the opposite direction for an AA of 30° . However, examining a bar chart of error angles across AAs, this did not appear to be the case. As shown in Figure 4, the CE appears to have been consistently pulled further anti-clockwise in the PRO HMD, while the effect is less strong and frequently in the opposite direction in the REG HMD.

4.1.2 Absolute error

To evaluate how difficult participants found the task, analysis was performed on the absolute error angle. The data contained an outlier, as assessed by visual inspection of the boxplot. As a result, the nonparametric Wilcoxon signed-rank test was used.

Data are medians unless otherwise stated. Of the 36 participants, the PRO HMD elicited an improved (decreased) absolute error angle in 20 participants compared to the REG HMD. The difference

Table 1: Median values and post hoc pairwise comparisons for error angle by target angle. Yellow indicates a significant value. As the table would be diagonally symmetrical, identical values are not shown for simplicity. Likewise, due to symmetry, the column for TA -60° is not required and has been removed.

TA	Median	Pairwise adjusted significance							
		-45°	-30°	-15°	0°	15°	30°	45°	60°
-60°	7.79	1.0	1.0	.001	1.0	1.0	.011	<.0005	<.0005
-45°	6.40		1.0	<.0005	1.0	1.0	.005	<.0005	<.0005
-30°	3.40			.061	1.0	1.0	.354	<.0005	<.0005
-15°	-1.40				1.0	1.0	1.0	1.0	.015
0°	1.84					1.0	.008	<.0005	<.0005
15°	4.27						.008	<.0005	<.0005
30°	-3.09							1.0	.002
45°	-7.98								1.0
60°	-11.89								

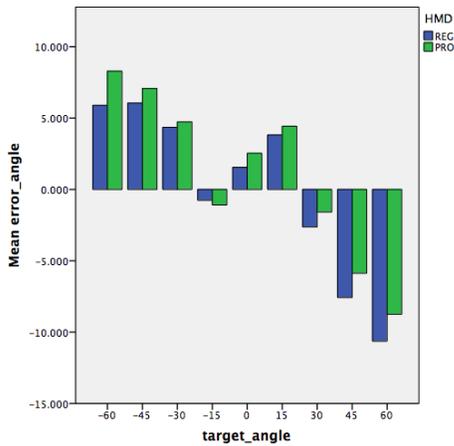


Figure 5: Histogram of common error (mean of error angles) for all target angles. The histogram is split down by HMD. Values are averaged across all avatar angles using means.

scores were symmetrically distributed, as assessed by a histogram. A Wilcoxon signed-rank test determined that there was not a statistically significant median decrease in absolute error angle (-0.439°) when subjects wore the PRO HMD (8.417°) compared to the REG HMD (9.259°), $z = -.935$, $p = .350$.

4.2 Target angle

4.2.1 Common error

There were outliers in the data, as assessed by inspection of a boxplot. As a result, the nonparametric Friedman test was used.

A Friedman test was run to determine if there were differences in CE between TAs. Pairwise comparisons were performed with a Bonferroni correction for multiple comparisons. CE was statistically significantly different at the different TAs, $\chi^2(8) = 127.422$, $p < .0005$. Due to the large number of levels of TAs, the median values and post hoc results are shown in Table 1. The mean values for error angle by TA are shown graphically in Figure 5.

4.2.2 Absolute error

There were outliers in the data, as assessed by inspection of a boxplot. As a result, the nonparametric Friedman test was used.

A Friedman test was run to determine if there were differences in absolute error between TAs. Pairwise comparisons were performed with a Bonferroni correction for multiple comparisons. Absolute error was statistically significantly different at the different TAs, $\chi^2(8) = 146.170$, $p < .0005$. Due to the large number of TAs, the

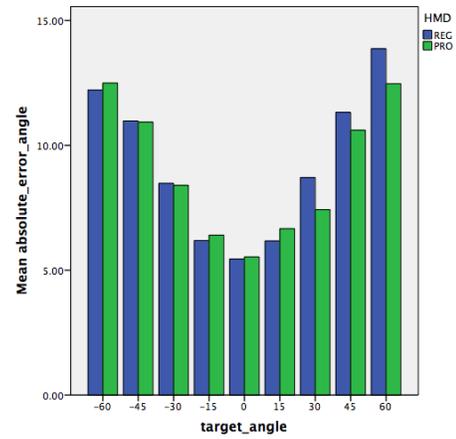


Figure 6: Histogram of mean absolute error angle for all target angles. The histogram is split down by HMD. Values are averaged across all avatar angles using means.

Table 2: Median values and post hoc pairwise comparisons for absolute error angle by target angle. Yellow indicates a significant value. As in Table 1, repeated values and columns are not displayed.

TA	Median	Pairwise adjusted significance							
		-45°	-30°	-15°	0°	15°	30°	45°	60°
-60°	12.86	1.0	.004	<.0005	<.0005	<.0005	.003	1.0	1.0
-45°	10.04		.142	<.0005	<.0005	<.0005	.107	1.0	1.0
-30°	8.10			.093	.005	1.0	1.0	.452	<.0005
-15°	5.77				1.0	1.0	.124	<.0005	<.0005
0°	5.17					1.0	.007	<.0005	<.0005
15°	6.20						1.0	.001	<.0005
30°	7.88							.354	<.0005
45°	10.76								1.0
60°	12.70								

median values and post hoc results are shown in Table 2. The mean of absolute error values by TA is shown graphically in Figure 6.

4.3 Avatar angle

4.3.1 Common error

There were outliers in the data, as assessed by inspection of a boxplot. As a result, the nonparametric Friedman test was used.

A Friedman test was run to determine if there were differences in CE between AAs. CE was not statistically significantly different at the different AAs, $\chi^2(6) = 7.516$, $p = .276$. A histogram of CE across different AAs is shown in Figure 4.

4.3.2 Absolute error

There were outliers in the data, as assessed by inspection of a boxplot. As a result, the nonparametric Friedman test was used.

A Friedman test was run to determine if there were differences in absolute error between AAs. Pairwise comparisons were performed with a Bonferroni correction for multiple comparisons. Absolute error was statistically significantly different at the different AAs, $\chi^2(6) = 93.369$, $p < .0005$. Due to the large number of levels of AAs, the median values and post hoc results are shown in Table 3. Absolute error angles by AA are shown graphically in Figure 7.

4.4 Qualitative feedback

Some participants commented on the realism of the avatar. For most, they enjoyed this realism, although one participant found it unsettling: "He's creepy... it's a little too real" [p. 17]. This may have been due to the static nature of our avatar, with the same

Table 3: Median values and post hoc pairwise comparisons for absolute error angle by avatar angle. Yellow indicates a significant value. As in Table 1, repeated values and columns are not displayed.

AA	Median	Pairwise adjusted significance					
		-30°	-15°	0°	15°	30°	45°
-45°	8.71	1.0	.015	<.0005	.022	.611	.339
-30°	8.81		.611	.010	.801	.015	.008
-15°	7.89			1.0	1.0	<.0005	<.0005
0°	7.43				1.0	<.0005	<.0005
15°	7.45					<.0005	<.0005
30°	10.84						1.0
45°	10.45						

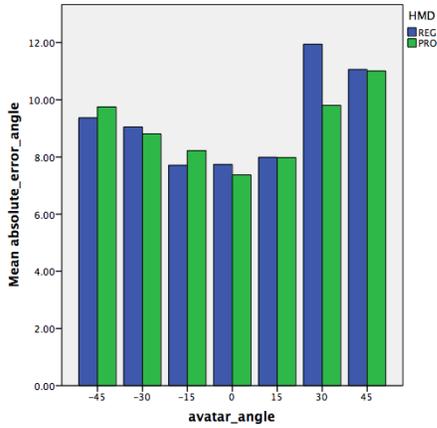


Figure 7: Histogram of mean absolute error angle for all avatar angles. The histogram is split down by HMD. Values are averaged across all target angles using means.

participant commenting, “I keep thinking he’s going to give me a jump scare.”

Eye gaze may play an important role in making these characters appear life-like. One participant noted that the avatar “looked distracted” [p. 27], while another commented, “his gaze was a bit off... a bit spacey, like he was daydreaming” [p. 33]. This may be the result of the convergence mismatch, in which the actor looked at a physical target during filming that was at a different distance from the virtual target in the experiment. Ensuring convincing eye-gaze behavior may be critical for making these performances enjoyable, however. One participant commented that “[the avatar] looked fake when he wasn’t looking at it [the target]. When he was looking at it he looked alive, not dead” [p. 29].

5 DISCUSSION

5.1 Head-mounted display resolution

The results indicate that the HMD did not have effect on how difficult participants found the task, as the absolute error angle between devices did not change significantly. This was in contrast to our expectation, as it was hypothesised (H1) that improved fidelity would improve participants’ ability to perceive the gaze of the avatar.

It is important to note that, in this experiment, we have only checked one distance of the user from the avatar. This was due to indications from a pilot study that changing the distance between the user and the avatar might have led to learning effects. In a pilot study, 10 participants performed a task similar to that described in section 3. In contrast to the final study, the pilot study used only three AAs (-30°, 0°, 30°), while the distance between the participant and the avatar was also varied (1m, 2m, 3m). Identical TAs were used. In the pilot study, there appeared to be a strong learning effect between HMDs. We concluded that this was likely caused by the

low number of AAs used, and the fact that the distance to the avatar varied. By bringing the user closer to the avatar, participants were able to get a better view of the avatar’s eyes, and therefore potentially learn where the eyes of each model were looking. With only three AA models shown, it may have become obvious over time which was being viewed.

To decrease the likelihood of learning effects, we increased the number of AA to seven, and removed the distance between the user and the avatar as an IV. A distance of 2m between the avatar and the user was chosen as it was the mid-point in the distances used in the pilot. The difference in resolution between the two HMDs under test, however, does not appear to have impacted accuracy with participants positioned 2m from the avatar.

The impact of the HMD on the CE is difficult to account for. It could be hypothesized that increased fidelity in the eyes may impact CE as a clearer view of the eyes may have impacted the overshoot effects reported in previous studies on eye gaze. However, any such impact would have been expected to affected the CE symmetrically, e.g. an overshoot to the left should have resulted in an equally large overshoot to the right.

Despite this, the HMD condition appears to have had an effect on CE, with participants consistently providing answers in the PRO that moved the CE to their right in comparison to their answer in the REG HMD. This effect can be seen in Figure 4. While issues around resolution, focus, lens calibration and IPD have all been considered, each of these would likely have resulted in an effect that was equally balanced. While the software executable run for both HMDs was the same, it is possible that underlying differences in the hardware or SDKs may be responsible. It is possible that there are other factors at play that we have not accounted for.

5.2 Target angle

Target angle had a significant effect on both CE and task difficulty. For CE, the effect is shown in Figure 5. Error angles are positive in the anticlockwise direction, as shown in Figure 2. For targets to the participants left (negative TAs), the mean answer was positive. This effect was symmetric across the 0° TA. This means that participants tended to believe the avatar was looking at the target when he was actually looking closer to the participant.

This finding is not consistent with H2, which hypothesized that the CE would be 0° for all TAs. This finding may relate specifically to performing these tasks in VR. Previous work has consistently shown that users tend to underestimate distances in VR (e.g. [11]). A recent study has shown that while underperception of distance may be less of an issue in modern HMDs such as the Vive, the effect is still present [20]. As performing the experimental task correctly also requires the user to evaluate their distance from the target in question, there may be an interplay between assessing the avatar’s eye gaze direction and evaluating the target position. As our results indicate that the user tended to record the eye gaze at a position closer to themselves, this is consistent with VR’s tendency to cause underestimations of distance. This might have important implications for volumetric content, as creators may need to consider how to present eye gaze in a way that is convincing rather than correct. Conversely, eye gaze might be useful to help VR users more accurately resolve object locations.

As shown in Figure 5, the trend of error angles does not hold up for target angles -15°, 0° and 15°. This may be caused by the experimental design. It was decided that the virtual target would be positioned directly in the avatar’s eye line. The user was placed at exactly eye-level with the avatar. This meant that, for a TA of 0°, the target partially obscured the avatar’s face. It was expected that participants would lean slightly to correct for this, but during trials this was not always the case. For TAs -15° and 15°, targets were close enough that the user could lean to create near-dyadic gaze, which can also partially obscure the avatar’s face. Participants did

this often. This may account for results that do not fit the trend for TAs near the user.

The TA had a significant effect on how difficult the task was, with targets further from 0° having larger absolute error values. This finding is consistent with H3. This indicates that content creators need to be more careful about ensuring eye lines are correct when the object of interest is closer to the user, and indeed it may be possible to sacrifice accuracy for distant objects to ensure correctness for nearer ones, for example by subtly rotating the VCA.

5.3 Avatar angle

Avatar angle was not found to have a significant effect on CE, indicating that participants did not exhibit any overshoot effects as suggested in the literature. This finding is not consistent with H4, which hypothesized that the CE would overshoot the target in the direction of the head turn, as shown in Figure 4.

As the model by which humans judge gaze direction is not fully understood, there are multiple factors that may have contributed to this difference from previous real-world studies. Aspects such as the repeated exposure to identical stimuli, the resolution of the HMDs, and depth perception in VR displays are all possible areas for further investigation. Such investigations may even help us understand the underlying model by which humans interpret gaze direction.

The AA was found to significantly impact accuracy levels, as assessed using absolute error angles. This was consistent with H5.

5.4 Comparisons with other works

While Cline found no impact on CE from TA [9], our results indicate CE is shifted closer to the user as TAs diverge from 0°. This is in direct contrast to Anstis et al., who instead found an overestimation effect [2]. Real-world studies have indicated a significant effect from AA on CE. This effect was not present in our data, indicating that HMDs may provide an opportunity to further study human perception of gaze direction. In our study, AA and TA both had a similar effect on accuracy to that found in previous work.

As study designs and reporting differ, drawing precise comparisons to other works is difficult. Here we draw some broad parallels, however, with the caveat that these must be treated with caution. We only list horizontal gaze results. Cline found an accuracy of 0.7° for dyadic and 3.2° for triadic gaze perception, similar to Bock et al. who found a triadic accuracy of 2.3° [6]. Gibson and Pick found a dyadic accuracy of 2.8° [6], while Roberts et al. reported between 3.8° and 5.9° for different AAs [34]. Our accuracy results ranged from 5.17° for a TA of 0°, up to 12.86° for a TA of -60°.

5.5 Limitations

In this study, we have only explored triadic gaze. While in the 0° TA condition the VCA is directed to look at the user's start location, the task requires the participant to direct the VCA's gaze at a target between them, and hence does not correspond to mutual gaze. Such an investigation would be highly valuable, as mutual gaze represents one of the most critical social cues.

As described in section 3.5, the IPD of the HMDs were fixed during the experiment. Previous work has indicated that IPD miscalibration may impact depth perception, although the strength of this effect is debated [5]. As our IPD settings were close to the mean and median human IPD of 63mm [13], we expect any impact on the results to be minimal. In future work, however, we would encourage the precise calibration of IPD settings to ensure accurate depth perception of the target. It is worth noting that some HMDs do not allow the IPD to be set, and our results would reflect performance in these scenarios.

Additionally, while we believe the main difference between the two HMDs used in our experiment was resolution, there are potentially other factors that may have impacted our results. The design of the head strap, the underlying SDKs, lens calibration, etc., could

have had an effect. Future studies may wish to artificially reduce the resolution of a single HMD to ensure consistency, although these results will not reflect the real-world performance of specific HMDs.

There is a tension between allowing users to view the scene with 6DoF, and ensuring consistency between participants. In our experiment, participants were allowed to lean, but told not to take any steps in any direction. Participants took advantage of the 6DoF tracking to varying degrees. We have not yet analysed the HMD tracking data. In future work, we intend to analyse the extent to which participant movement impacts task performance. As we used a repeated measures design, however, we do not believe that differing movement patterns between participants will impact our results. Future studies may wish to control this more strictly, for example by restricting the range of 6DoF movement to a small region, or limiting movement to 3DoF.

As described in section 5.2, there were times when the target obscured one of the VCA's eyes. Participants did not always lean to correct for this. In future studies, it would be better if the target shape or locations were designed to prevent this.

As described in section 3.3, a convergence mismatch exists in our experiment, as the targets using during filming were not the same distance from the actor as the virtual targets were from the VCA. Additionally, volumetric recordings can suffer from artefacts, and it is possible that small artefacts in the eyes could have impacted our results. These are limitations of the current generation of volumetric capture technologies. Further studies would need to be performed, using different volumetric reconstructions and actors, to ensure our findings are generalisable.

6 CONCLUSIONS

In this paper, we explored the perception of eye-gaze direction of volumetrically captured avatars when viewed in a HMD. In a repeated measures experiment, 36 participants repeatedly rotated an avatar until they felt it was looking at a target. The avatar's gaze direction, the location of the target and the HMD were varied.

In contrast to expectation, the resolution of the HMD was not found to affect accuracy levels. The HMD was found to impact the CE, although this effect did not appear to be symmetric. We have not been able to account for this result. Results indicate the task became significantly harder as targets diverged from 0°, i.e. poorer accuracy was found for targets further away from the user. The gaze direction of the avatar was also found to impact accuracy, with larger AAs resulting in lower accuracy.

In contrast to real-world studies, eye-gaze direction was consistently misjudged based on target location, but not based on the avatar's head turn direction. The impact of target location may have been affected by the known phenomenon that users tend to underestimate distances in VR.

Eye gaze perception will likely play an important role in the adoption of volumetric capture as a medium for VR content creation and telecommunications. To this end, an understanding of the perception of VCAs' eye-gaze direction in VR displays is critical. Our results indicate that eye-gaze awareness in VR does not always parallel similar tasks in the real world, and further investigation may help explain the cognitive processes behind eye-gaze understanding.

ACKNOWLEDGMENTS

This work was supported in part by grants EP/N509577/1 and EP/M029263/1 from the UK Engineering and Physical Sciences Research Council (EPSRC). The authors would like to thank Dimension Studio and Digital Catapult for their support with this work.

REFERENCES

- [1] S. R. Acker and S. R. Levitt. Designing videoconference facilities for improved eye contact. *Journal of Broadcasting & Electronic Media*, 31(2):181–191, 1987.

- [2] S. M. Anstis, J. W. Mayhew, and T. Morley. The perception of where a face or television 'portrait' is looking. *The American journal of psychology*, 82(4):474–489, 1969.
- [3] M. Argyle and M. Cook. Gaze and mutual gaze. 1976.
- [4] G. Bailly, S. Raidt, and F. Elisei. Gaze, conversational agents and face-to-face communication. *Speech Communication*, 52(6):598–612, 2010.
- [5] S. Best. Perceptual and oculomotor implications of interpupillary distance settings on a head-mounted virtual display. In *Aerospace and Electronics Conference, 1996. NAECON 1996., Proceedings of the IEEE 1996 National*, vol. 1, pp. 429–434. IEEE, 1996.
- [6] S. W. Bock, P. Dicke, and P. Thier. How precise is gaze following in humans? *Vision research*, 48(7):946–957, 2008.
- [7] J. Cassell, C. Pelachaud, N. Badler, M. Steedman, B. Achorn, T. Becket, B. Douville, S. Prevost, and M. Stone. Animated conversation: rule-based generation of facial expression, gesture & spoken intonation for multiple conversational agents. In *Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, pp. 413–420. ACM, 1994.
- [8] M. Chen. Leveraging the asymmetric sensitivity of eye contact for videoconferencing. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 49–56. ACM, 2002.
- [9] M. G. Cline. The perception of where a person is looking. *The American journal of psychology*, 80(1):41–50, 1967.
- [10] A. Collet, M. Chuang, P. Sweeney, D. Gillett, D. Evseev, D. Calabrese, H. Hoppe, A. Kirk, and S. Sullivan. High-quality streamable free-viewpoint video. *ACM Transactions on Graphics (TOG)*, 34(4):69, 2015.
- [11] 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*, pp. 69–117. Elsevier, 1995.
- [12] Dimension. Dimension studio. <https://www.dimensionstudio.co/>, 2019. Accessed: 8-February-2019.
- [13] N. A. Dodgson. Variation and extrema of human interpupillary distance. In *Stereoscopic Displays and Virtual Reality Systems XI*, vol. 5291, pp. 36–47. International Society for Optics and Photonics, 2004.
- [14] J. H. Ellgring and M. V. Cranach. Processes of learning in the recognition of eye-signals. *European Journal of Social Psychology*, 2(1):33–43, 1972.
- [15] M. Garau, M. Slater, S. Bee, and M. A. Sasse. The impact of eye gaze on communication using humanoid avatars. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 309–316. ACM, 2001.
- [16] M. Garau, M. Slater, V. Vinayagamoorthy, A. Brogni, A. Steed, and M. A. Sasse. The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 529–536. ACM, 2003.
- [17] J. J. Gibson and A. D. Pick. Perception of another person's looking behavior. *The American journal of psychology*, 76(3):386–394, 1963.
- [18] T. Imai, D. Sekiguchi, M. Inami, N. Kawakami, and S. Tachi. Measuring gaze direction perception capability of humans to design human centered communication systems. *Presence: Teleoperators & Virtual Environments*, 15(2):123–138, 2006.
- [19] A. Jones, M. Lang, G. Fyffe, X. Yu, J. Busch, I. McDowall, M. Bolas, and P. Debevec. Achieving eye contact in a one-to-many 3d video teleconferencing system. In *ACM Transactions on Graphics (TOG)*, vol. 28, p. 64. ACM, 2009.
- [20] J. W. Kelly, L. A. Cherep, and Z. D. Siegel. Perceived space in the htc vive. *ACM Transactions on Applied Perception (TAP)*, 15(1):2, 2017.
- [21] A. Kendon. Some functions of gaze-direction in social interaction. *Acta psychologica*, 26:22–63, 1967.
- [22] B. Lang. Lytro's latest VR light-field camera is huge, and hugely improved. <https://www.roadtovr.com/lytro-immerge-latest-light-field-camera-shows-major-gains-in-capture-quality/>. Accessed: 16-September-2018.
- [23] S. P. Lee, J. B. Badler, and N. I. Badler. Eyes alive. In *ACM Transactions on Graphics (TOG)*, vol. 21, pp. 637–644. ACM, 2002.
- [24] M. Levoy and P. Hanrahan. Light field rendering. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pp. 31–42. ACM, 1996.
- [25] C. Lord and M. M. Haith. The perception of eye contact. *Perception & Psychophysics*, 16(3):413–416, 1974.
- [26] X. Ma and Z. Deng. Natural eye motion synthesis by modeling gaze-head coupling. In *Virtual Reality Conference, 2009. VR 2009. IEEE*, pp. 143–150. IEEE, 2009.
- [27] K. Melnick. Blade Runner 2049: Memory Lab brings human actors to cinematic vr. <https://vrscout.com/news/blade-runner-2049-memory-lab-cinematic-vr/>, 2017. Accessed: 15-September-2018.
- [28] P. Moors, K. Verfaillie, T. Daems, I. Pomianowska, and F. Germeys. The effect of head orientation on perceived gaze direction: revisiting gibson and pick (1963) and cline (1967). *Frontiers in psychology*, 7:1191, 2016.
- [29] D. T. Nguyen and J. Canny. Multiview: improving trust in group video conferencing through spatial faithfulness. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 1465–1474. ACM, 2007.
- [30] S. Orts-Escolano, C. Rhemann, S. Fanello, W. Chang, A. Kowdle, Y. Degtyarev, D. Kim, P. L. Davidson, S. Khamis, M. Dou, et al. Holoportation: Virtual 3d teleportation in real-time. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, pp. 741–754. ACM, 2016.
- [31] Y. Pan, O. Oyekoya, and A. Steed. A surround video capture and presentation system for preservation of eye-gaze in teleconferencing applications. *Presence*, 24(1):24–43, 2015.
- [32] Y. Pan, W. Steptoe, and A. Steed. Comparing flat and spherical displays in a trust scenario in avatar-mediated interaction. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pp. 1397–1406. ACM, 2014.
- [33] D. Roberts, R. Wolff, J. Rae, A. Steed, R. Aspin, M. McIntyre, A. Pena, O. Oyekoya, and W. Steptoe. Communicating eye-gaze across a distance: Comparing an eye-gaze enabled immersive collaborative virtual environment, aligned video conferencing, and being together. In *Virtual reality conference, 2009. VR 2009. IEEE*, pp. 135–142. IEEE, 2009.
- [34] D. J. Roberts, J. Rae, T. W. Duckworth, C. M. Moore, and R. Aspin. Estimating the gaze of a virtuality human. *IEEE transactions on visualization and computer graphics*, 19(4):681–690, 2013.
- [35] W. Steptoe, O. Oyekoya, and A. Steed. Eyelid kinematics for virtual characters. *Computer animation and virtual worlds*, 21(3-4):161–171, 2010.
- [36] W. Steptoe and A. Steed. High-fidelity avatar eye-representation. In *Virtual Reality Conference, 2008. VR '08. IEEE*, pp. 111–114. IEEE, 2008.
- [37] W. Steptoe, A. Steed, A. Rovira, and J. Rae. Lie tracking: social presence, truth and deception in avatar-mediated telecommunication. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1039–1048. ACM, 2010.
- [38] L. A. Symons, K. Lee, C. C. Cedrone, and M. Nishimura. What are you looking at? acuity for triadic eye gaze. *The Journal of general psychology*, 131(4):451, 2004.
- [39] K. Takashima, Y. Omori, Y. Yoshimoto, Y. Itoh, Y. Kitamura, and F. Kishino. Effects of avatar's blinking animation on person impressions. In *Proceedings of graphics interface 2008*, pp. 169–176. Canadian Information Processing Society, 2008.
- [40] N. F. Troje and U. Siebeck. Illumination-induced apparent shift in orientation of human heads. *Perception*, 27(6):671–680, 1998.
- [41] R. Vertegaal, G. Van der Veer, and H. Vons. Effects of gaze on multiparty mediated communication. In *Graphics Interface*, pp. 95–102, 2000.
- [42] R. Vertegaal, I. Weevers, C. Sohn, and C. Cheung. Gaze-2: conveying eye contact in group video conferencing using eye-controlled camera direction. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 521–528. ACM, 2003.
- [43] E. Williams. Experimental comparisons of face-to-face and mediated communication: A review. *Psychological Bulletin*, 84(5):963, 1977.
- [44] H. R. Wilson, F. Wilkinson, L.-M. Lin, and M. Castillo. Perception of head orientation. *Vision research*, 40(5):459–472, 2000.
- [45] R. Yang and Z. Zhang. Eye gaze correction with stereovision for video

teleconferencing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(7):956–960, 2004.