Virtual human gaze evaluation protocol across photorealism

Joelle Zimmermann
Magic Leap
Los Angeles, CA 90232, USA
jzimmermann@magicleap.com

Thomas Miller
Magic Leap
Los Angeles, CA 90232, USA
tmiller@magicleap.com

John Monos
Magic Leap
Los Angeles, CA 90232, USA
jmonos@magicleap.com

CHI 2020 Extended Abstracts, April 25–30, 2020, Honolulu, HI, USA.

ACM ISBN 978-1-4503-6819-3/20/04.
DOI: https://doi.org/10.1145/3334480.XXXXXXX
*update the above block & DOI per your rightsreview confirmation (provided after acceptance)

Abstract
In the area of social virtual human technology, there is a need for evaluation protocols for different aspects of communication. While the importance of gaze in nonverbal communication is generally accepted, methods for examining gaze realism of virtual humans are lacking. More broadly, the level of photorealism of virtual humans required for effective communication remains an open question. We present a study using a novel paradigm that uses eye gaze in an observer-looker scenario to evaluate gaze across several levels of photorealism in order to guide virtual human development for social AR/VR. Furthermore, we propose measuring gaze perception as a first order behavioral metric for assessing inter-personal communication.

Author Keywords
Virtual human; gaze; perception; mixed reality

Introduction
Gaze plays an important role in social interaction – it reflects emotional responding and attention allocation [20], serves to regulate the flow of conversation, and regulates interpersonal intimacy [2]. The ability to accurately discriminate gaze direction, in particular, mutual gaze, is critical to these social interactions.

Research with human observers and participants shows that people are very accurate at judging another's gaze direction [1, 13]. Accuracy in gaze judgements is greatest during direct gaze, suggesting that humans are especially sensitive to mutual gaze [12]. Seminal studies by Gibson and Pick (1963) and Cline (1967)

Data Analysis

For each subject, we then calculated: 1) the mean offset of perceived look-ats, and 2) SD of offsets of all perceived look-ats (perceived look-ats = look-ats perceived as direct gaze by the user) [15]. The means and SDs represent the center and radius, respectively, of the tolerance range that is considered direct gaze by observers.

Subjective Measures

After the experiment, we asked subjects to complete a questionnaire, rating, for each character, their 1) gaze confidence (i.e. I felt confident when I answered that the character was looking directly in my eyes"), and 2) social presence (i.e. the feeling of being with another). We used the 5-item social presence questionnaire from Bailenson and colleagues [6]; subjects answered each question on a 5 point Likert scale ranging from "Strongly Agree" to "Strongly Disagree". We then averaged the answers across 5 items.

focused on determining the just noticeable deviation of gaze, and the accuracy with which an observer can distinguish between direct and averted gaze. A typical experimental paradigm involves two individuals, an acting "looker", and a participant "observer" - who makes judgements about the looker's gaze direction [14, 16, 15, 12, 1, 11]. The looker changes gaze direction to look at different targets invisible to the observer. The observer is asked where they perceive that the looker is looking. This research shows that people are sensitive to small deviations in another's gaze [9].

While the majority of gaze detection research is performed with a human looker and observer, a virtual human may be able to reproduce the key features of gaze following from naturalistic settings and thus provide insights into relevant features [3]. However, users may be poor at judging gaze direction of virtual humans in VR (i.e. [19], volumetric captured avatars). Virtual human gaze has been shown to impact user's subjective experience and behavior, including: co-presence and liking [5, 18], task performance [5], avoiding collision while walking [17] or interpersonal distance during an interaction [4]. Despite the importance of virtual human gaze, paradigms for evaluating the accuracy of gaze perception in AR/VR are limited.

The contributions of this paper are: 1) the proposal of a novel observer-looker paradigm to evaluate virtual human gaze 2) suggestions for photorealism of social AR/VR characters based on gaze perception. We show a CG character in 3D via a Mixed Reality headset, and assess the accuracy with which an observer can discriminate between direct and averted gaze across varying levels of virtual human photorealism and viewing distance. Research suggests that, with less visual information, observers tend to assume mutual gaze [14]. We therefore hypothesized that observers will have a wider tolerance range with increasing viewing distance and with decreasing character photorealism.

Methods

Twenty-six participants with normal vision wearing a Magic Leap One headset viewed a virtual human in a custom application. Each participant experienced 6 conditions: 3 characters of varying visual fidelity - High fidelity Mica by Magic Leap, medium fidelity Kiteboy by Epic Games, and low fidelity Quinn by Magic Leap (Avatar Chat), and 2 viewing distances: 60cm and 150cm. The character heads were approximately that of an adult human head shown at the participant's height.

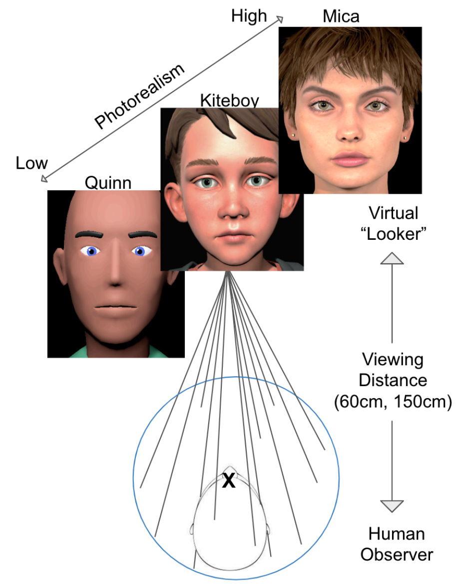

Figure 1: Experiment setup. The virtual looker (high fidelity Mica, medium fidelity Kiteboy, low fidelity Quinn; shown in randomized order) looks at randomly distributed invisible look-at targets around the human observer. The sphere containing the invisible targets is centered around the observer's headpose, demarcated by X. Viewing distance is varied (60cm, 150cm).

Results: Subjective Measures

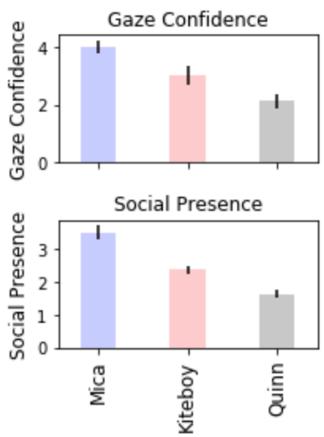

Figure 3: Gaze Confidence (left) and Social Presence (right) scores by character.

A within subjects ANOVA revealed that user's judgements of their Gaze Confidence in the character's gaze increased with increasing photorealism: $F(2,50) = 25.61$, $p < 0.001$, $\eta_p^2 = 0.51$, as did social presence ($F(2,50) = 63.32$, $p < 0.001$, $\eta_p^2 = 0.72$). See Figure 3.

On-device hardware and algorithms were used to predicate the participant's location (denoted by tracking the point between their eyes). The character looked at randomly selected invisible target points on and around the participant, in all directions (X,Y,Z). The random distribution of look-at targets was selected from a distribution with target SD in each direction = 5cm, min/max = -70/70. Virtual characters' gaze was driven by a realistic gaze system with a model of saccadic motion [7]. Characters saccaded between 15 invisible targets, holding each look-at for a duration of 2-3 seconds. Participants held a remote control and were instructed to press the control trigger whenever they felt that the virtual character is looking directly in their eyes. The accuracy of perception can be quantified by the degree with which the "perceived look-ats" (where subjects pressed the trigger), clustered around the observer's headpose (the point between their eyes). Participants also rated their confidence in the other's gaze and social presence (See Sidebar).

Results

We use the term "gaze cone" to describe the range within which observers perceive a looker to be looking directly at them [14].

Gaze Cone Center: We conducted a 2x3 (2 distances X 3 characters) within-subject analysis of variance (ANOVA) on the subject mean cm offsets of perceived look-ats. Partial eta-squared effect sizes (η_p^2) were reported. We found a significant effect of viewing distance $F(1,25) = 28.88$, $p < 0.001$, $\eta_p^2 = 0.54$, character $F(2,50) = 13.34$, $p < 0.001$, $\eta_p^2 = 0.35$, as well as their interaction $F(2,50) = 17.22$, $p < 0.001$, $\eta_p^2 = 0.41$ on the mean offsets of perceived look-ats. A posthoc Tukey test for character revealed significant differences in gaze cone center (MD = mean difference) between Mica and Quinn (MD = 1.38, $p < 0.005$), Kiteboy and Quinn (MD = 1.14, $p < 0.005$), however not between Mica and Kiteboy. Overall, observers made more accurate (closer to 0) judgements when the highly photorealistic character was looking directly at them; the differences between characters was particular apparent at a close viewing distance.

Gaze Cone Width: We conducted a second 2x3 within-subject ANOVA on the subject SD of cm offsets of perceived look-ats to identify differences in gaze cone across conditions. We found a significant effect of viewing distance $F(1,25) = 55.93$, $p < 0.001$, $\eta_p^2 = 0.69$ and of character $F(2,50) = 15.92$, $p < 0.001$, $\eta_p^2 = 0.39$ on the SD of offsets of perceived look-ats. A posthoc Tukey test revealed significant differences between all pairs of characters. Mean differences (MD) are: Mica-Quinn MD = 1.32, $p < 0.005$; Kiteboy-Quinn MD = 0.76, $p < 0.005$; Mica-Kiteboy MD = 0.56, $p < 0.05$. There was also an interaction of distance and character $F(2,50) = 13.36$, $p < 0.01$, $\eta_p^2 = 0.35$. Thus, observers have a smaller tolerance range for what is perceived as direct gaze from a highly photorealistic character, particularly when viewing the character from a near distance.

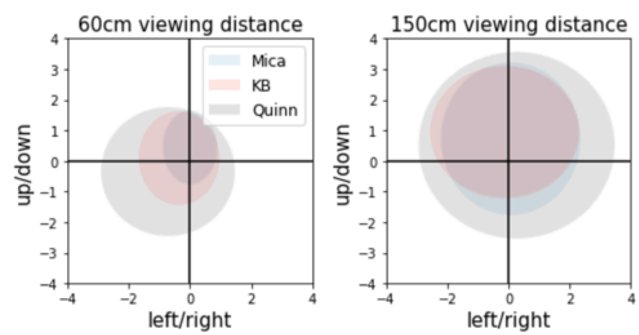

Figure 2: Offsets (in cm) from the observer's headpose (0,0) of all look-ats perceived as direct gaze, from Mica (high fidelity), Kiteboy (medium fidelity), Quinn (low fidelity). The ellipse centers represent the mean offset from the observer's headpose across subjects (i.e. gaze cone center). The width and height of the ellipse represents the mean, across subjects, of all SDs of all look-at offsets perceived as direct gaze (i.e. gaze cone range) in the left/right and up/down direction. The L and R column show results for 60cm and 150 cm viewing distance conditions.

Discussion

In the present study, we describe and test a gaze evaluation paradigm for virtual humans in VR/AR. We show that the observer's tolerance range for discrimination of a virtual human's direct gaze varies across photorealism and viewing distance. We found that: 1) With increasing distance, observers were more likely to assume that the virtual human is looking directly at them, even when they are not (in line with previous work [14, 21]); and 2) With decreasing photorealism, observers were more likely to assume that the virtual character was looking directly at them, even when they were not. This was particularly true when viewing the character from a near distance. At a near distance, the difference between the characters' visual fidelity may be more apparent, whereas at a further distance, visual information is diminished, and observers default to assuming direct gaze when they are unsure. Observers also became more *confident* in their judgements of gaze direction with increasing photorealism. In line with previous work [22], higher social presence was reported.

Together, our findings suggest that photorealistic characters allow for improved gaze perception. Current social VR applications are trending towards more photorealism in computer generated (CG) avatars (at the level of Kiteboy and beyond). The present study suggests that photorealistic characters are indeed a positive choice as a remote communication tool.

Interestingly, some of the present results are in contrast to those of MacQuarrie and Steed (2019), who studied how HMD resolution, distance and head direction of volumetrically captured avatars impacts gaze perception [19]. While distance similarly negatively impacted gaze perception accuracy in the two studies, HMD resolution did not. The difference between studies may be due to the inherent differences in capturing and representing humans (volumetric capture and display vs CG avatar in our study). The fact that participants were poor at judging gaze direction based on eye direction [19] supports this theory.

As eye-tracking technology in AR/VR continues to develop for driving avatar gaze, this protocol can help determine tracking accuracy requirements for varying avatar photorealism. While the present experiment examined modelled gaze, realistic (real-time, eye-tracking) driven gaze can be studied in future work. The tolerance range for judging direct gaze can be taken as the threshold of allowed eye-tracking error for a particular character. Thus, virtual characters can be designed in line with the accuracy of the particular eye-tracking system. Future experiments will additionally examine the impact of head and body direction, as these may have an impact on gaze perception [14, 19], as well as (along with mutual gaze) on emotion understanding and attention allocation [20].

The paradigm can be expanded to examine other aspects of gaze-related communication beyond direct eye contact. For instance, a follow-up test can easily be modified to examine referential gaze perception; the ability to discern visual attention towards a shared virtual object, so that users of social VR are able to communicate about shared content in a natural way.

More broadly, we envision measuring gaze perception as a first order behavioral metric for assessing inter-personal communication between humans and virtual humans, prior to second order metrics that measure outcome. While second order outcome metrics, such as completion time on a joint task [10], provide a high-level assessment of communication success, they do not offer information on the specific aspects of communication in which the system is lacking. The present evaluation paradigm offers a means to assess perception of mutual gaze, referential gaze, or other foundations of nonverbal communication.

References

[1] Stuart Anstis, Peter Mayhew, Peter T Morley. 1969. The perception of where a face or television ‘‘portrait’’ is looking, *American Journal of Psychology* 82: 474–489.

[2] Michael Argyle, Roger Ingham, Florisse Alkema, Margaret McCallin. 2009. The Different Functions of Gaze. *Semiotica*. 7(1): 19-32.

[3] Artin Atabaki, Karolina Marciniak, Peter W. Dicke, Peter Thier. 2015. Assessing the precision of gaze following using a stereoscopic 3D virtual reality setting. *Vision Research*. 112: 68-82.

[4] Jeremy N. Bailenson, Jim Blascovich, Andrew C. Beall, and Jack M Loomis. 2001. Equilibrium revisited: Mutual gaze and personal space in virtual environments. *Presence: Teleoper. Virtual* Environ, 10,6 (December 2001) 583-598.https://doi.org/10.1162/105474601753272844

[5] Jeremy N. Bailenson, Andrew C. Beall, and Jim Blascovich. 2002. Mutual gaze and task performance in shared virtual environments. *Journal of Visualization and Computer Animation* 13, 1-8.

[6] Jeremy N. Bailenson, Jim Blascovich, Beall, Andrew C. Beall and Jack M. Loomis. 2003. Interpersonal distance in immersive virtual environments. *Personality and Social Psychology* Bulletin, 29, 7: 819–833.

[7] James Bancroft, Nafees Bin Zafar, Sean Comer, Takashi Kuribayashi, Jonathan Litt, and Thomas Miller. 2019. Mica: a photoreal character for spatial computing. In *ACM SIGGRAPH 2019 Talks (SIGGRAPH ’19).* 9, 1-2. https://doi.org/10.1145/3306307.3328192

[8] Simon W. Bock, Peter W. Dicke, Peter Their. 2008. How precise is gaze following in humans? *Vision Research*, 48: 946-957

[9] Leanne S. Bohannon, Andrew M. Herbert, Jeff B. Pelz, and Esa M. Rantanen, 2013. Eye contact and video-mediated communication: A review. *Displays*, 34, 2: 177-185.

[10] Elizabeth A. Boyle, Anne H. Anderson, Alison Newlands. 1994. The effects of visibility on dialogue and performance in a cooperative problem solving task, *Language and Speech* 37: 1–20

[11] Milton Chen. 2002. Leveraging the asymmetric sensitivity of eye contact for videoconference. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ’02).* 49–56. https://doi.org/10.1145/503376.503386

[12] Marvin G. Cline. 1967. The perception of where a person is looking, *American Journal of Psychology* 80, 1: 41–51.

[13] Caroline Gale, Andrew F. Monk. 2000. Where am I looking? The accuracy of video-mediated gaze awareness. *Perception & Psychophysics* 62, 586–595

[14] Matthias Gamer, Heiko Hecht. 2007. Are you looking at me? Measuring the cone of gaze. Journal of Experimental Psychology: *Human Perception and Performance*, 33, 3: 705-715.

[15] James J. Gibson, Anne D. Pick. 1963. Perception of another person’s looking behavior, *American Journal of Psychology* 76: 386–394.

[16] Stephen R.H. Langton, Helen Honeyman, Emma Tessler. 2004. The influence of head contour and nose angle on the perception of eye-gaze direction. *Perception and Psychophysics* 66: 752–771.

[17] Sean D. Lynch, Julien Pettré, Julien Bruneau, Richard Kulpa, Armel Crétual, and Anne-Hélène Olivier. 2018. Effect of Virtual Human Gaze Behaviour During an Orthogonal Collision Avoidance Walking Task. *In 2018 IEEE Conference*

on Virtual Reality and 3D User Interfaces (VR), 136-142. doi: 10.1109/VR.2018.8446180

[18] Kim, Kangsoo, Arjun Nagendran, Jeremy N. Bailenson and Greg Welch. 2015. Expectancy Violations Related to a Virtual Human's Joint Gaze Behavior in Real-Virtual Human Interactions. In *28th Annual Conference on Computer Animation and Social Agents*.

[19] Andrew MacQuarrie, Anthony Steed. 2019. Perception of volumetric characters' eye-gaze direction in head-mounted displays. In 2019 *IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*

[20] Linda Marschner, Sebastian Pannasch, Johannes Schulz & Sven-Thomas Graupner. 2015. Social communication with virtual agents: The effects of body and gaze direction on attention and emotional responding in human observers. *International Journal of Psychophysiology, 97*(2), 85–92.

[21] Wade W. Martin, Robert F. Jones. 1982. The accuracy of eye-gaze judgement: A signal detection approach. *British Journal of Social Psychology*, 21(4): 293-299

[22] Katja Zibrek and Rachel McDonnell. 2019. Social presence and place illusion are affected by photorealism in embodied VR. In *Motion, Interaction and Games (MIG '19)*. 13: 1–7. https://doi.org/10.1145/3359566.3360064