Amplifying Realities: Gradual and Seamless Scaling of Visual and Auditory Stimuli in Extended Reality

Zubin Choudhary*
University of Central Florida

Figure 1: Example of an amplified visual stimulus in augmented reality: the "Big Head" technique [4, 5]. From left to right the images show an unaugmented view of two humans at different distances, an AR head that was up-scaled to be able to barely identify the face, the same head up-scaled more to make out the gaze direction (looking right), and even more up-scaled to perceive the facial expression (smiling).

Abstract

Existing literature in the field of extended reality has demonstrated that visual/auditory manipulations can change a person's perception and behavior. For example, a mismatch between the physical self and the virtual self can have psychological and behavioral implications. There are many different approaches that can incur a perceptual change. An under-explored field of research are gradual and subtle manipulations, such as scaling the food one eats or scaling the heads of people one sees. In this position paper, I provide an overview of my prior PhD research focusing on means to gradually and seamlessly scale visual and auditory stimuli in extended reality, and investigations of the corresponding changes in human perception. I discuss future research topics and potential questions to be discussed at the ISMAR 2021 Doctoral Consortium.

Index Terms: Computing methodologies—Computer graphics—Graphics systems and interfaces—Virtual reality; Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Augmented reality; Human-centered computing—Human computer interaction (HCI)—HCI design and evaluation methods—User studies

1 Introduction

A large body of research has focused on identifying and investigating means to change a person's perception and experience of their real or virtual environment using extended reality (XR) technologies. For instance, classical uses of augmented reality (AR) technologies focused on adding/overlaying virtual objects or annotations over a person's view of the real world. Conversely, related work on diminished reality (DR) focused on removing such visual elements from a person's view. While many of these classical applications focused on *discrete* manipulation goals, such as adding an object or removing it, recent work has started to investigate more *gradual* and subtle means for such manipulations.

*e-mail: zubinchoudhary@knights.ucf.edu

For instance, previous work on the *MetaCookie+* in mixed reality (MR) showed that subtle manipulations of the perceived visual size of a real cookie (or other food) while eating could change a person's sense of satiety [11, 12]. In the scope of my PhD research, I introduced a similar technique in XR, called the *Big Head* technique, which is illustrated in Figure 1. I have shown that a human's head can be gradually up-scaled (or down-scaled) in XR to compensate for reduced visual facial cues over long distances, e.g., due to the limited visual acuity of the human eye or caused by the low resolution of XR displays [4, 5], to allow collaborators to identify each other or read their facial expressions or gaze directions, without causing strong Uncanny Valley effects [10]. I further found evidence that audio cues can be gradually scaled in XR to improve a person's ability to hear other people over long distances without strong side-effects on distance perception [4].

The main focus of my PhD research is to explore and investigate novel and creative means to gradually manipulate visual or auditory stimuli in XR as part of a person's perception of their environment, to investigate how subtle changes can affect perceptions within and across modalities, and to advance the theory on XR methodologies and techniques. I started my Computer Science PhD research in 2019, which so far has resulted in two major publications at IEEE VR 2020 [5] and IEEE VR 2021 [4], among others [2, 3, 13].

The remainder of this position paper is structured as follows: Section 2 gives a high-level overview of my research topic. Section 3 shows my previous work in this field. Section 4 discussed my research plan and future work. Section 5 concludes the paper.

2 Amplifying Realities

Extended reality technologies can allow a person to experience realities that may conform to nature or go beyond the limitations of natural space, matter, or biology. Since a person's experience of their environment is created through computer-generated stimuli in virtual reality (VR), and partially through computer-generated (or digitized) stimuli in MR and AR, it is possible to manipulate some of these stimuli using visual/auditory rules and patterns in real time. On a basic theoretical level, it is possible to express many of these possible manipulations in XR through a simple one-dimensional axis (see Figure 2), denoting the normal/natural scale of the visual/auditory stimuli at a scale of one, with a zero scale indicating the point when stimuli vanish, and a gradual interval in between and beyond.

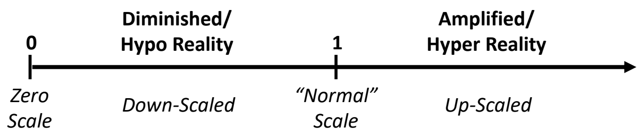

Figure 2: Reality Spectrum

When applied to a person's experience of their environment, in the scope of my PhD research, I have broadly broken down these experiences or realities into three types:

- **Normal Reality:** A person experiences a reality with a one-to-one mapping to nature, e.g., in terms of visual/auditory appearances, physics, and behaviors.
- **Amplified or Hyper Reality:** This denotes an experience characterized by amplified stimuli, i.e., increased or up-scaled. The word *hyper* originates from ancient Greek, meaning “over,” implying excess or exaggeration. These experiences may, for instance, be characterized by magnified visuals, amplified ranges of bodily motion, such as in redirected locomotion research [14], or up-scaled sound volumes.
- **Diminished or Hypo Reality:** This experience is characterized by diminished stimuli. The word *hypo* is an antonym for *hyper*, meaning “under,” implying reduced or shortened. For example, these experiences may involve shrunken visuals or reduced ranges/frequencies of embodied motion among avatars or agents in XR.

As briefly mentioned in the introduction, manipulations leading to hyper or hypo real experiences can be discrete or more subtle and gradual. While a large body of research focused on discrete manipulation goals, my PhD research is exploring the less thoroughly studied field of manipulations aimed at more *gradual* and subtle changes, leveraging parametric models of visual/auditory changes. For instance, a seminal study in this field was performed by Narumi et al. [11], who manipulated the size of a real cookie while eating. They found that a larger visual size made participants eat 10% fewer cookies and a smaller size made them eat 15% more, and they proposed that enlarging food could support individuals in losing weight. Other related work by Kilteni et al. [8] investigated the illusion of an extended virtual arm compared to their real length. They showed that multisensory and sensorimotor information can reconstruct our perception of the body shape, size and symmetry even when this is not consistent with normal body proportions. The above experiments are examples of subtle and gradual manipulations that can affect our perceptions. However, not all gradual manipulations may lead to a perceptual change. Compared to discrete changes in the environment, e.g., adding or removing an object, a gradual change in an object's (or person's) appearance, motion, or pose is not as distinct and noticeable, which suggests the existence of perceptual thresholds for ranges of subtle or overt manipulations. In the following section, I describe my own prior work in this field, which so far focused largely on scaling the head of real humans or virtual avatars/agents, and I discuss some findings.

3 Prior PhD Research

My prior work so far was mainly focused around the so-called *Big Head* technique. This work was inspired by the *Donkey Kong Mode* [6], where heads of game characters were disproportionately up-scaled relative to the rest of the body. This allowed developers to present game characters to players with limited resolution and screen space. I leveraged this general approach to study its uses and perceptual implications in VR and AR.

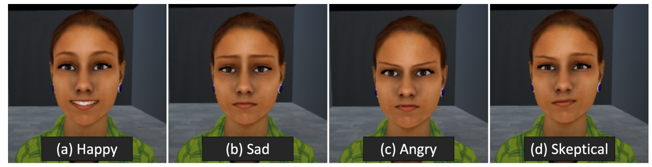

Figure 3: The avatar's four facial expressions used for the experiment: (a) happy, (b) sad, (c) angry, and (d) skeptical.

Figure 4: Scaling methods: Fixed Neck Height and Fixed Eye Height

3.1 Virtual Big Heads

In a paper published at IEEE VR 2020 [5], I investigated the application of this technique to recover lost facial cues over distance in VR. I hypothesised that the *Big Head* technique could be a viable option to recover the lost visual cues in VR, as smaller visual cues, such as facial expressions, become difficult to perceive over longer distances due to the low resolution of current-state VR head-mounted displays (HMDs). The study design comprised of four facial expressions: happy, sad, angry, and skeptical (see Figure 3), and the avatar appeared at five different distances, considering the proxemics in social interactions based on Hall's Proxemics theory [7]. I tested the *Big Head* technique with two scaling methods (see Figure 4): one that maintained a fixed neck height of the avatar and one that maintained a fixed eye height. For the virtual environment, I created a simulated hallway which resembled my lab, and I used a female virtual avatar with an idle animation while the different facial expressions were presented in a continuous loop. The avatar and the surrounding virtual environment were displayed through an HTC Vive Pro HMD, and participants could adjust the scale of the avatar's head using a Vive Pro controller.

Overall, our results showed that participants increased the avatar's head scale as distance increased beyond three meters, at which point they were unable to recognize the facial expressions of a regularly-sized avatar/agent any longer. I introduced two different scaling methods (see Figure 4), of which they preferred the method that maintains eye height instead of neck height because for the convenience of seeing the avatar's face, and importance of maintaining eye contact. When not concerned with recognizing facial expressions, participants felt more comfortable with head sizes that were up-scaled over their natural size at longer distances. Since no strong effects related to the Uncanny Valley were observed and participants even generally preferred slightly up-scaled heads, it proved to be a useful technique. The technique supports subtle and gradual manipulations and can recover facial cues over long distances without having to compromise on their comfort.

3.2 *Big Head* in Distance Perception

My second study in this domain, published at IEEE VR 2021 [4], evaluated the possible effects of the *Big Head* technique on distance perception in VR. Along with the *Big Head* technique, I also tested two different avatar body representations (see Figure 5): full body and head only, and three different audio scales: up-scaled (louder), accurate, and down-scaled (softer). Based on previous work in the field of VR, it is established that longer distances are often underestimated and shorter distances may be overestimated [9, 15]. I hypothesized that avatars with bigger heads and up-scaled audio volumes would be perceived closer, whereas avatars with smaller heads and down-scaled audio would be perceived farther away. For the virtual environment, I reused the virtual hallway from the previous experiment, and I used a male virtual avatar with an idle animation. The avatar and the surrounding virtual environment were displayed through an Oculus Rift S HMD, and participants used an Oculus touch controller for distance judgments via two protocols: verbal estimations and blind triangulated pointing [1]. I tested this design over three distances: 3, 5, and 7 meters. As established in my previous study, a scaled-up head can incur perceptual changes. For this study, I was interested to investigate a similar scaling, but for audio. I used a voice recording in a continuous loop, and attached it to the virtual avatar. To simulate up-scaled and down-scaled audio volumes, I spatially moved the audio source closer, and further away, for the respective effects. Participants experienced all combinations of three head scales, three audio scales and two body representations over three distances, and estimated their perceived distance of the virtual avatar.

The study revealed a significant effect of head scales on distance judgements but, interestingly, only if the scaled heads were presented as floating objects in VR, but not when they were spatially anchored and attached to a human body at true scale. This means that not only visual facial perception but other perceptions like distances can be affected by non-veridical head scales on virtual humans. However, regarding audio scales, I did not find a noticeable effect on distance judgements for any condition, with overall very similar magnitudes of estimated distances for all tested audio scales, suggesting that even large manipulations of audio volumes had negligible effects on distance perception.

Figure 5: Comparison of scaled heads and their effects on distance perception for two social VR body representations with a full body or a head-only avatar

4 Planned Future Work

My future PhD research will involve a wider range of methods aimed at amplifying static or dynamic visual/auditory stimuli. For instance, while my previous work focused on scaling a person's entire head in VR, future work may focus on scaling facial features or *movements*, e.g., by amplifying the corresponding facial muscles/motions to make the movements more pronounced, making them easier to perceive for people with Autism or over long distances. Similar approaches may be used to scale a person's range of motion, non-human entities, or objects in the environment. These effects may be investigated using a wide range of display technologies including VR, AR, MR, and even CAVE-like setups.

For instance, the following research questions may lead to interesting discussions at the ISMAR 2021 Doctoral Consortium that may help me make an educated decision on future topics and directions for my PhD research:

RQ1 Which application domains may benefit from scaled visual or auditory cues? Which scaled cues may elicit strong/noticeable benefits in different social/professional settings?

RQ2 Can gradual manipulations in hypo or hyper reality support transitions for traditionally discrete manipulations? Can gradual changes over time support perceptual/motor adaptation?

RQ3 What are the perceptual implications in VR/AR if it becomes possible to scale any type of visual/auditory element? Can scalings elicit impossible percepts or experiences?

5 Conclusion

In this position paper, I outlined my previous PhD research on gradual and seamless scaling in extended reality. I presented examples of my previous work, which have already revealed a few interesting methods and techniques for scaling of embodied cues in extended reality, with the promise of a wider range of future basic and applied research topics. I am enthusiastically looking forward to discussing my previous and future research with the organizers and mentors at the ISMAR 2021 Doctoral Consortium.

Acknowledgments

This material includes work supported in part by the National Science Foundation under Award Number 1564065 (Dr. Ephraim P. Glinert, IIS) and Collaborative Award Numbers 1800961, 1800947, and 1800922 (Dr. Ephraim P. Glinert, IIS) to the University of Central Florida, University of Florida, and Stanford University respectively; the Office of Naval Research under Award Numbers N00014-17-1-2927 and N00014-21-1-2578 (Dr. Peter Squire, Code 34); and the AdventHealth Endowed Chair in HealthcareSimulation (Prof. Welch).

References

[1] G. Bruder, F. Argelaguet, A.-H. Olivier, and A. Lécuyer. Cave size matters: Effects of screen distance and parallax on distance estimation in large immersive display setups. *Presence*, 25(1):1–16, 2016.

[2] Z. Choudhary, G. Bruder, and G. Welch. Scaled user embodied representations in virtual andaugmented reality. In *Proceedings of the Workshop on User-Embodied Interaction in Virtual Reality (UIVR)*, pp. 1–3, 2021.

[3] Z. Choudhary, A. Erickson, K. Kim, G. Bruder, and G. Welch. AR Big Heads: Analysis of Human Facial Perception and Comfort of Head Scales in Outdoor Augmented Reality. In *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 1–10, 2021, *under review*.

[4] Z. Choudhary, M. Gottsacker, K. Kim, R. Schubert, J. Stefanucci, G. Bruder, and G. F. Welch. Revisiting distance perception with scaled embodied cues in social virtual reality. In *Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 788–797, 2021.

[5] Z. Choudhary, K. Kim, R. Schubert, G. Bruder, and G. F. Welch. Virtual big heads: Analysis of human perception and comfort of head scales in social virtual reality. In *Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 425–433, 2020.

[6] Giant Bomb. Big Head Mode. `https://www.giantbomb.com/big-head-mode/3015-403/`, Nov 6, 2019.

[7] E. T. Hall, R. L. Birdwhistell, B. Bock, P. Bohannan, A. R. Diebold Jr, M. Durbin, M. S. Edmonson, J. Fischer, D. Hymes, S. T. Kimball, et al. Proxemics [and comments and replies]. *Current anthropology*, 9(2/3):83–108, 1968.

[8] K. Kilteni, J.-M. Normand, M. V. Sanchez-Vives, and M. Slater. Extending body space in immersive virtual reality: a very long arm illusion. *PloS one*, 7(7):e40867, 2012.

[9] J. M. Loomis, J. M. Knapp, et al. Visual perception of egocentric distance in real and virtual environments. *Virtual and adaptive environments*, 11:21–46, 2003.

[10] M. Mori, K. F. MacDorman, and N. Kageki. The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine*, 19(2):98–100, 2012.

[11] T. Narumi, Y. Ban, T. Kajinami, T. Tanikawa, and M. Hirose. Augmented perception of satiety: controlling food consumption by changing apparent size of food with augmented reality. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 109–118, 2012.

[12] T. Narumi, S. Nishizaka, T. Kajinami, T. Tanikawa, and M. Hirose. Meta cookie+: An illusion-based gustatory display. In R. Shumaker, ed., *Virtual and Mixed Reality - New Trends*, pp. 260–269, 2011.

[13] N. Norouzi, K. Kim, G. Bruder, A. Erickson, Z. Choudhary, Y. Li, and G. Welch. A systematic literature review of embodied augmented reality agents in head-mounted display environments. In *Proceedings of the International Conference on Artificial Reality and Telexistence & Eurographics Symposium on Virtual Environments (ICAT-EGVE)*, pp. 101–111, 2020.

[14] S. Razzaque. *Redirected walking*. The University of North Carolina at Chapel Hill, 2005.

[15] R. S. Renner, B. M. Velichkovsky, and J. R. Helmert. The perception of egocentric distances in virtual environments-a review. *ACM Computing Surveys (CSUR)*, 46(2):1–40, 2013.