Ambika Miglani

Spring 2019

Perception of Egocentric Distances in Virtual Spaces

A Study of "Presence" and "Immersion" in Virtual Reality

Veridical perception is becoming increasingly necessary in applications of virtual reality technology. Virtual reality is currently being used interchangeably to describe either a virtual environment on a flat display screen, a room-based system such as a CAVE, or a head mounted display (HMD), all in which there is an element of stereoscopic depth gives the illusion of 3D space. This illusion does not necessarily give rise to images that are accurate or representative of real space, which is unfavourable to the various fields which have use for it. In the field of psychological research, having control over otherwise uncontrollable variables in a 3D space is incredibly useful (such as in the case of phobic stimulus, studying anxiety and avoidance behaviour etc) (Wilson). VR is also used biology for visualising cell structures, in the military for training, and for architectural modelling (Renner). In some of these cases, non-veridical perception of spatial layout and egocentric distance deems the technology less useful that it could be or even removes the utility from it altogether-- which is the reason for the extensive research done to understand the reasoning behind the underestimation of egocentric distance in virtual spaces. The factors that studies believe are affecting distance estimates are measurement methods, technical, compositional, and human factors. Concurrently, the ideas of "presence" and "immersion" are used often to evaluate the effectiveness of virtual reality in allowing the viewer feel as if they are a part of, and interacting with, the environment they are placed with, alongside the continuous investigation into the quality of graphics, and whether or not the sophistication of the visuals presented to the viewer has an impact on the way they perceive spatial layout and distance. *This paper seeks to evaluate the experiments hat study the repeated underestimation and compression of egocentric distance in virtual spaces, compare the potential causes of this phenomenon that have been studied, and evaluate whether these studies support or reject Gibson's ecological approach to perception.*

The argument that the relatively poor quality of computer graphics is the reason for the misperception of distance could be understood through the lens of Gibson's ecological approach to perspective, and also validates it. The roots of this approach are based in principles from Gibson's ecological principles. Amongst other depth cues, the ecological approach often references the importance of ground textures in spatial awareness. Without high quality computer graphics and image rendering, it is not possible to accurately portray textures and shadows as they appear in physical reality. According the Gibson, colors are only defined in the optic array from how they differ in illumination from the adjacent colors, with textures being a conglomerate of various colours in the optic array(Gibson, p.83). For example, if a rugged texture had shadows that were of a lighter tone than they were intended to be, the surface could be perceived as smoother than it would be otherwise, and because of this possibly even longer (i.e. the wrong information would be fed to the viewer) which would have a direct effect on the perception of distance. Lumination leads to the emergence of colors and textures. In this case, the lack of use of appropriate lumination tools in computer graphics is a possible cause of the inaccurate rendering of surfaces and spacing. This is discussed in an architecture study by Beata Stahre, in which the issue of the interaction between light and objects in a scene is analysed. She suggests that the difficulty to solve the radiance problem (as represented by the equation, Radiance= emitted radiance + total reflected radiance), is partially responsible for the difficulty to render accurate 3D spaces, even when people choose the global solution to the radiance problem. Lighting is a key element in the discussion of the ecological approach, as the interaction between light and objects give rise to the ambient optic array, which is used for perception, which is proved in studies by Thomson Et Al. (Renner Et. Al) by the improved accuracy of verbal estimates after increasing the detail of textures in virtual spaces. This explanation works well in sparse, underdeveloped graphics, but does not sufficiently explain the case for environments with dense, developed illumination rendering In both cases, without this interaction, veridical perception is not possible.

The computational and constructivist approaches to perception, in terms of prior knowledge compensating for missing information, invalids the argument that the quality of graphics is a significant reason for underestimating egocentric distance in virtual spaces. However, in experiments studied by Renner et al, the addition of knowledge of the constitution of the object being placed in a virtual space did not improve accuracy of estimates. If distance estimates are regarded as “noticeable differences” (Renner et al), but the comparing the position of a known object in space did not lead to a successful noticing of difference, then it can be argued that the constructivist lens of approaching this issue of visual cues is inapplicable. Alternatively, other studies that have to do with audio and haptic feedback provided positive results, and increased the accuracy of the estimates. While the constructivist approach is not necessarily supported in this case, one can argue that neither is the ecological perspective. By giving observers the exact size of an object, they are inherently provided with the absolute distance information needed. This concept aligns with the ecological concept of affordances, and that each object has an intrinsic set of qualities which provide information to the viewer on how they can interact with said object. According to Gibson, these qualities contain enough visual information for veridical perception, particularly because the retinal projections in both the physical and virtual spaces are similar (Jones Et al). Therefore, the argument that low quality of graphics rendering does not sufficiently explain the cause of misperception of distance, nor does it present enough evidence that supports the ecological, computational, or constructivist approaches. Therefore, the argument that sparse environments composed of graphics low detail graphics leading to consistent underestimation of distance is negligible, and can be studied as supplementary to arguments that deal with the physical feelings of immersion and presence in virtual reality.

Effectiveness of virtual reality is often measured through “immersion” and “presence, and while virtual spaces with sophisticated graphics rendering may have a strong illusion of immersion, it is this lack of presence which is often used to explain underestimation of egocentric distance. Immersion is the measurement of how “ecologically valid” stimuli in VR are, or how the level of “sensory fidelity” of the environment, whereas presence refers to the extent to which the viewer feels as if they are a part of, or interacting with, the environment. With a decreased sense of presence, there is a dissociation between viewer and environment, and the less likely the viewer will perform judgements and actions in the virtual world as they would the in the real world. Presence is essential in constructing an accurate egocentric spatial experience, and the parietal lobe of the brain (Kober et al) integrates the information from different senses to generate an egocentric view of the world that gives a viewer information about their own height within a space. Gibson's ecological approach reflects this idea that an observer calibrates body and environment through the perceived position of their own body, in that the field of view of the animal gives rise to proprioceptive information.

The occluding edges of the nose, cheekbones and eyebrows are closer to the viewer's point of observation than the occluding edges of the arms, legs, hands, and feet, hence providing details on the perception of distance and depth. While using a headset to view environments in virtual reality, the viewer's field of view does not include the occluding edges that are normally used to calibrate and calculate concepts of depth. Cues about height do not necessarily need to be sourced from the view of one's own feet and arms, as Gibson often suggests, as even when the view of feet is occluded in real world experiments, distance estimates are not warped (Renner et al). Methods to increase the sense of presence include adding avatars to the scene to make the viewer (Schubert et al) as a method of increasing interactivity between the viewer and space. Some VR systems already seek to do this. For example, in Marshmallow Laser Feast's "We live in an ocean of air" VR installation, users are equipped with infrared cameras that detect their hands and breath and present them in the virtual environment using miniscule spheres that represent carbon dioxide and oxygen molecules, and were able to see the breath molecules of other participants in the room. As a result, very few collisions between people occur and users are able to understand the boundaries of the space. In this case, although the user still could not see their exact body, placing the hands in the frame gave the viewer information about their own height. While this was not done under experimental conditions, formal studies have been done that show how adding avatars to a scene significantly improves distance estimates (Renner et al), regardless of whether they were static or tracked. It was shown that when the viewer had greater autonomy over their own movements or that they had ownership over the avatar as the avatar was considered in relation to their own body, hence increasing the feeling of presence and providing a more accurate portrayal of egocentric distance. Awareness of one's own body in a space leads to calculations in the geometry of space, and can be considered through a computational lens to be the result of very precise measurements of depth disparity and angles of inclination or declination.

Presence is also affected by the depth disparity between the viewer's two eyes, a topic which strongly resembles the computational approach, although it is also discussed in GIbson's ecological approach. To find depth disparity and hence have a parameter for what a certain depth should look like, subjects must have a view of the occluding edges of their eye sockets. The disparity in images perceived in the two separate eyes form a combined view of the scene that, when compared, can find the mismatch between images and be used to locate oneself in an environment and acquire depth information (Poyade et al) In a virtual environment presented in an HMD, while Gibson explains this in terms of the movement of optical rays, the systematic and linear process of image detection, stitching, and disparity detection are implicative of computational systems. The computational approach and ecological approach are complementary in the case of understanding how a viewer sees themselves in the environment they are placed in, and are supported by the perceptual matching, action-based, and verbal tasks that are carried out in studies that seek to understand presence in VR.

However, some may argue that regardless of the degree of "presence" required to invite the viewer to interact normally with the environment, the immersiveness of the environment defined by the visual, auditory, haptic, and olfactory fidelity (Slater et al) is the most important aspect in terms of veridical perception, as "presence" can just as easily be an emotional response to a space, and be subjective to the observer. Presence, and attempts to increase the feeling of presence, can therefore be seen as independent from the visual calibration process (Interrante et al), and for this reason may be discounted from the attempt to understand compressed depth perception. The arguments that assume the reliance of "presence" on one's internal state contrast with Gibson's studies on the geometry of space. Slater describes presence as "orthogonal to emotional content", and that two are interrelated, so the level of presence felt is dependent on the observers' propensity for immersion themselves depending on their psychological state and their desire to give their attention to the object. It is difficult to distinguish them from one another, suggesting that 'presence' to be an unreliable factor in determining the effectiveness of VR, and hence not necessarily the most accurate method in understanding the misestimation of distance. Even if the arguments of presence and visual calibration are separated, as the study by Interrante et al discusses, the geometry of space dependent on the viewer's angle of inclination and declination have tangible effects on the way they perceive spatial layout, height cues, and depth cues. Contrary to the arguments claiming that presence is nothing but emotional receptiveness to an environment and is not a valid measurement of the effectiveness of VR, the geometry of space is another determinant of depth perception It can affect and be affected by the lack of awareness in of the observer's own body, and has shown to impact they way they act in their environment.

The height of the point of observation is used to yield information about the horizon-distance angle, is then calculate distance to the object. Without the information that places the viewers point of observation at a specific height, it is not possible for this sequence of interpretations to occur. This is evaluated in the paper by Renner et al, and through Messing and Durgin's 2005 experiments shows that if the camera height was distorted to begin with (in their case, lowering the horizon by 1.5 degrees in a virtual space), the angle of declination was warped, and distance estimates were affected. Although there is contradicting evidence in the case of the angle of declination and its effect on distance judgements, studies such as those done by Woods et al in 1983 (Renner et al) show that pincushion distortion, the concept that the image stretches as it moves away from the center, does not have evidence suggesting its effect on distance estimates. In this case, the warped angle of declination serves as a stronger explanation for depth compression and cannot be replaced with the other forms of distortion as reasoning.

It can be argued that the disconnect between accommodation and convergence caused by motion parallax is a major reason for the compression of depth in VR. The concepts of accommodation and convergence are not aligned with Gibson's ecological view of perception. In virtual reality, the object is created through a series of images. While the observer is fixated on this object, the object itself is moving (Bingham et al in Wilson et al), which results in this decoupling. In other words, although the viewer can focus at close range and has power over where they focus their attention to, there is no parallax produced by the eye movement, which restricts how the viewer perceives distance (Messing et al). The accommodative signal is constant and determined by the optics of the HMD, but accommodation on its own is not considered to be a reliable indication of depth, and is heavily dependent on the convergence system, as they form a "body centred match" (Slater et al). The dissociation between accommodation and convergence in addition to the subtle lags that comes from computing the motion of objects in virtual spaces are all thought to have effects on depth perception. These technical systems can be applied to play a role in the feeling of presence, which is also affected by the physical weight of the HMD. HMDs have physical effects, as studied by Willemsen et al, and have proven to distort perception of distance. Gibson's theory of affordances, however, does not support the conclusion that the use of the HMDs cause observers to underestimate distance measurements. While Gibson might argue that the increased weight of the HMD on an observer's head may cause them to overestimate distance, the opposite occurs. However, according to the ecological approach, affordances of objects and spaces are what cause animals to make judgements and create actions. Hence, the affordances of an environment could also be a major factor in depth compression.

The studies evaluated come to a similar conclusion of ambiguity. Nonetheless, attempting to view these studies and theories through the lens of various approaches to perception. Through evaluating the technical and human reasons for the underestimation of distance in virtual spaces, mainly head mounted displays, the idea of "presence" as emerged to be the most prominent factor in judging the effectiveness of VR spaces, and hence explaining the perception of egocentric distance. Presence itself it a conglomerate of the quality of graphics presented in a headset, egocentric awareness, motion parallax, depth disparity, angles of inclination/declination etc., with priority given to factors such as providing the viewer with an indication of their eye height through adding avatars to the space, specifically avatars that are a part of, or attached to, their own bodies. This autonomy and sense of control will then increase the sense of presence. In terms of comparing te ecological and computational approaches, as the experiments done pertaining to depth compression in VR are exploratory and no definite solution has been found, many possibilities remain to explain this. As a result, there is a distribution of theories that support the various approaches to perception, with none of them being dominant in the discussion. The idea of presence is both ecological and computational in nature, with it relying on factors such as proprioception as well as the computation of depth disparity and assuming that the brain computes depth as a relative difference between objects and spaces the viewer is presented with. Some seem to believe that presence is highly emotional, and that in order to be achieved the virtual space must offer the viewer a reason to be present, regardless of their propensity to connect. If this emotional connection is truly so important in this process, which ultimately is needed in larger fields such as psychology research and architectural modelling, the environment must *afford* interaction and connection, and it would be interesting to further investigate the effect of different types emotional stimuli with objects and spaces that afford different levels of interaction on such a task as perception of egocentric distances in virtual reality.

Bibliography

1. Slater, M., Lotto, B. and Arnold, M. (2009). *How we experience immersive virtual environments: the concept of presence and its measurement*. [online] Diposit.ub.edu. Available at: http://diposit.ub.edu/dspace/bitstream/2445/49643/1/631349.pdf
2. .Wilson, J., C., & Alessandro. (2015, August 03). The Use of Virtual Reality in Psychology: A Case Study in Visual Perception. Retrieved from https://www.hindawi.com/journals/cmmm/2015/151702/
3. Interrante, V., Ries, B., Lindquist, J. and Anderson, L. (n.d.). *Elucidating Factors that can Faciliate Veridical Spatial Perception in Immersive Virtual Environments*. [online] Available at: https://www.researchgate.net/publication/236605608_Defining_Reality_in_Virtual_Reality_Exploring_Visual_Appearance_and_Spatial_Experience_Focusing_on_Colour
4. Wilson, J. C, Alessandro. The Use of Virtual Reality in Psychology: A Case Study in Visual Perception. https://www.hindawi.com/journals/cmmm/2015/151702/. Published August 3, 2015.
5. Sutcliffe, Alistair. "Presence, Memory and Interaction in Virtual Environments." *International Journal of Human-Computer Studies*, Academic Press, 17 Jan. 2005, www.sciencedirect.com/science/article/pii/S1071581904001430.
6. Messing, Ross, and Frank H Durgin. *Distance Perception and the Visual Horizon in Head-Mounted Displays*. ACM, 2005, www.cs.rochester.edu/u/rmessing/academic/undergrad/p1-messing.pdf.

7. Rousset, Thomas, and Jean Louis Vercher. “Misperception of Egocentric Distances in Virtual Environments: More a Question of Training than a Technological Issue?” *Displays*, Elsevier, 27 Feb. 2018, www.sciencedirect.com/science/article/pii/S0141938217300732.

8. Messing, Ross. “Distance Perception and Cues to Distance in Virtual Reality .” Vol. 10, no. 7, 2018, pp. 241–264., doi:http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.126.4772&rep=rep1&type=pdf.

9. Renner, R. S., Velichkovsky, B. M., and Helmert, J. R. 2013. The perception of egocentric distances in virtual environments - A review. ACM Comput. Surv. 46, 2, Article 23 (November 2013), 40 pages.DOI: http://dx.doi.org/10.1145/2543581.2543590

10. Willemsen, Peter.,Thompson, William, and Gooch, Amy. (2005). *Effects of Stereo Viewing Conditions on Distance Perception,* School of Computing University of Utah