Perceptual issues in designing augmented reality

Aditya Acharya
University of Birmingham
axa384@bham.ac.uk

ABSTRACT

Augmented reality is based on combining the content of the real world with the content of the virtual world. Although these systems enhance the virtual experience by adding elements of the real world, they are also prone to the limitation in the capability of accurately displaying relevant depth cues. As a result, we see perceptual issues interfere in performing task at hand using these systems. In this paper, we discuss some of the known depth issues in augmented reality. We then divert our attention to some specific issues related to environment with induced motion. Also, we look at shadows cue as a relevant solution for establishing perceptually correct spatial relationship between the environment and the virtual object.

Keywords

Human perception, augmented reality, depth perception

1. INTRODUCTION

Over the years, Augmented Reality (AR) and Virtual Reality (VR) technology have become very popular research areas, with achievements in the area of computer technology, graphics, sensor technology, ergonomics and man-machine interaction theory. However, researchers and practitioners are still trying to solve many fundamental problems related to designing an efficient AR system [1]. These problems range from issues related to technological aspect, for example, registration problem due to object tracking limitations; to perceptually correct display, for example, issues related to depth perception, illumination, field of view, etc. In either case, presence of such issues leads to poor task performance for the users of the AR systems [16].

Drascic, D. and P. Milgram in 1996 identified and classified perceptual issues in AR systems [6]. Their research revolved around issues pertaining to stereoscopic Head Mounted Displays. However, over the years we have seen different other platforms emerge for AR systems, e.g., mobile and tablet

Figure 1: City Lens augmented reality application by nokia

devices, video projector and Holograms. Even though Head mounted displays till date remain the popular and dominant platforms for AR systems, the emergence of these new platforms have widened the problem space for AR systems [16].

In this report we will try to look at perceptual problems encountered in these three platforms:

- Head mounted displays
- Mobile and Tablet devices
- Video Projectors

The scope of this report is limited to identifying some of the perceptual problems related to depth. In conclusion we try and discuss our understanding to this problem and possible areas of interest.

2. PERCEPTION

Perception can simply be defined as interpretation of sensory stimuli by the human brain. Gregory in 1997 gave an interesting approach to understanding human perception. He explains perception as a hypothesis which involves making inferences of the environment we see by utilizing the past knowledge and experience a person has of that environment [8]. Drascic, D. and P. Milgram(1996) described a proposition that follows from this idea, i.e., "the importance of consistency among the various cues which is used to infer information about our visual world"[6]. In addition, incorrect perception of the environment is caused due to inconsistency or conflict among those cues. In the next section we look at
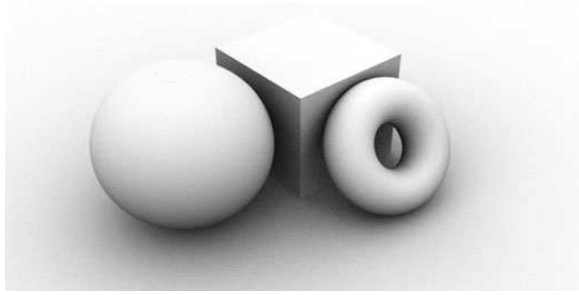
Figure 2: Example of Objects occluding

these different cues that are available in the real world that helps us correctly perceive the environment around us.

3.   OVERVIEW OF DEPTH CUES

Humans perceive depth in sensed images using, what is called as depth cues. These cues can be classified as extra retinal cues (information gained from other than retina), monocular cues (information gained from one eye) and binocular cues (information gained using both eyes). In real world, human visual system automatically uses all the available cues to accurately interpret depth information in a sensed image [27].

Monocular cues. These are features that give an impression of objects being farther or nearer to user on a flat surface. These cues are often used by artists in their painting to depict depth. Cues under this classification consist of:

- **Occlusion:** Objects that appear to block the view of other objects are perceived to human eyes as closer.

- **Linear Perspective:** Parallel lines moving farther away from the viewer appears to converge or come close together.

- **Relative Size:** In an image object of larger size are perceived closer as compared to object of smaller size.

- **Texture Gradient:** Object farther away from the viewer is perceived with lesser details or lower texture as compared to objects closer to the viewer.

- **Atmospheric Perspective:** Object farther away from the viewer seems to be blurred and with higher concentration of colour blue due to atmospheric distortion of light particles.

- **Relative Motion Parallax:** Object farther away from the viewer seems to follow the viewer and objects closer move out of perspective.

- **Shadows and Shades:** Object brighter appears to be closer and compared to the darker objects. Also, Object whose shadow falls on the other object appears to be closer.

Figure 3: Example of linear perspective
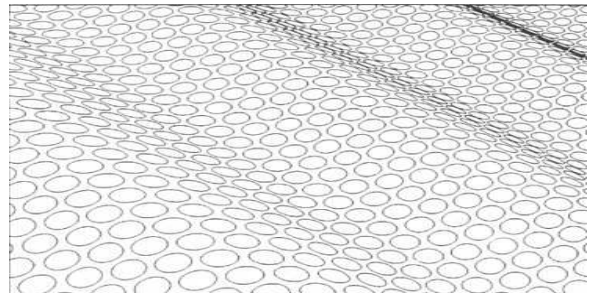
Figure 4: Texture grading showing depth.
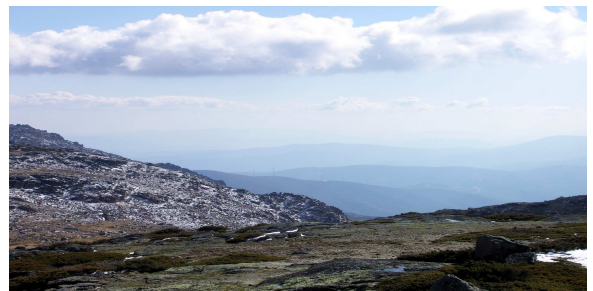
Figure 5: Atmospheric blur showing depth.

Extra-Retinal Cues. These are the cues that rely on the physical changes of the human eye to accommodate depth information. These cues are very weak and work at short distances. Cues under this classification consist of:

- **Accommodation:** it is the tension of the muscle that changes the focal length of the lens of eye to focus objects at different distances.

- **Convergence:** Objects closer to the viewer, the human eye tends to come close to each or points inwards. This change in direction of human eye is defined as convergence.

Binocular Cues. These cues rely on images captured by both the eyes. The left and right eyes capture two different perspective of environment around them. By using the difference in the two images, i.e., disparity in the images, a strong sense of depth is obtained.

4. PERCEPTUAL ISSUES IN AUGMENTED REALITY

Researchers have identified several issues that affect visual perception, which in turn affects the cognition or understanding the augmented content [16]. The goal here is to understand different perceptual issues that are encountered when a digital content is augmented in a real world. These problems are very common in augmented reality application where developers tend to misjudge the spatial relationship between the environment and the digital content. Thus, leads to various perceptually incorrect augmentations. In this section we look at some of the perceptual issues that cause depth distortion.

4.0.1 Distortion Due to Environment Structure

One of the critical areas where perceptual problems frequently originate is the environment it uses to augment. Lappin, Shelton and Rieser (2006) demonstrates that both accuracy and the precision of the visually perceived distance are conditioned on the structure and context of the environment being perceived [17]. One of the major problems faced in a structured environment is 'Clutter', which is defined as a state with excess items, or their organisation in an environment leads to degradation of visual search performance [24]. Also, a cluttered environment is difficult to augment as it limits the surface perception, i.e., foreground-background perception and scene understanding (in reference to occlusion relationship between objects in an environment perceived by the user). False perception of depth is seen when objects that have to be rendered behind an object appears in front, which causes false depth ordering in scene and gives an impression that object doesn't belong in the environment. Video Projectors are mostly influenced by the structure of environment, where characteristics like the angle of surface that is being used to project or texture of the surface may lead to depth distortion [16].

4.0.2 Distortion due to Field of View
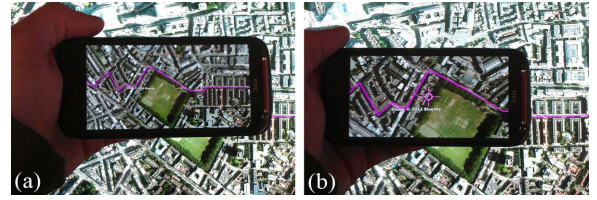
Figure 6: Perceptual issue due to FOV caused by (a)device perspective and (b)user-perspective

Field of view (FOV) is referred to the viewing area of a human eye. The inception of head mounted display and hand held devices, brought restriction to the human field of view, which in turn makes visual task difficult [29]. However, limited field of view may not cause problem in perceiving depth in head mounted display [15]. But, this issue becomes complex when it comes to hand held display, since the visual space is now divided into the hand held display and the human eye. Humans have a horizontal FOV of 104° to 94° for each eye (over 180° approx) [15], while hand-held devices have a relatively small FOV. This limited view of the world by the hand held devices is then used for augmenting the digital information. The consequence of this is two variations of a dual-view situation in turn cause disparity between the smaller and incorrectly scaled augmented view as compared to surrounding (180° vision) perceived by the viewer [16][4].

4.0.3 Distortion due to Resolution, Display size and Contrast

Contrast is defined as the difference in colour and intensity of light captured on an object to other objects. Ichihara, Kitagawa and Akutsu (2007) showed that region in an image with low contrast reduce the depth perception of a viewer [13]. Contrast in augmented reality plays an important role for depth perception. The input information in devices for augmenting the environment is taken from the camera. Here the limitation is caused by the "micro-contrast of the lens, which is the level of differentiation between smaller details that have an increasingly similar tonal value that the lens can capture" [16]. Resolution and display size of devices, on the other hand, have similar effects on depth perception. Users viewing augmented environment on devices with smaller display size and lower resolution had difficulty in perceiving depth information [5]. Contrast and resolution is a major problem in video projectors due to the dependency on nature of the surface [16].

4.0.4 Discussion

Kruijff and others (2010) in their review proposed solutions for the above mentioned problems. Solutions ranging from using x-ray in cluttered environment to clipping of augmented objects to avoid interference in occlusion [16]. Although, the proposed solutions solve some of the problems, there are still issues yet to be addressed. For example, adapting colour schemes dynamically for different environmental conditions to correctly perceive depth is still an open area for research.

Among all the issues, the role of motion (tracking) in augmented reality and the perceptual issues it bring along is still

one of the major unsolved problems in this area [28]. Lack of understating of motion in the environment induces problem of registration (refers to adding of content on the surface of the environment). Also, the fact that the real world environment is not in control of the developer, it causes miscalculation of various aspects of the environment, e.g. , light intensity, orientation of user, etc. From a perceptual point of view motion plays an important role in understanding depth. For example, with change in position of the viewer the cues that were previously dominating the environment, and, which were used to render the virtual object would no longer be the dominating cue which makes it difficult to accurately interpret distance or depth [23].

In our study we found two areas of interest, i.e. , depth issues related to change in position of user and perceiving location of the augmented object in the spatial environment when the objects in the environment are moving. We term this kind of environment as dynamic environment (classifying in relation to motion in environment). Augmenting objects in dynamic environment is very difficult, especially when it comes to perception. Issues that we discussed above become a very common problem in dynamic environment. A lot of work has been done from the point of view of registering objects in such environments [16]. We look at these problems from depth perspective and the cues that can help resolve to some extent these problems. In the next section we talk about how Shadow cue can be an influential contributor to understanding and perceiving depth and distance in dynamic environment.

5. SHADOWS IN DYNAMIC ENVIRONMENT

Shadows are defined as the area with absence of direct light due to obstruction from another object. In dynamic environment where motion is implicit understanding spatial arrangement of objects becomes difficult. Shadows in such environment become a useful tool in understanding spatial relationship between the object and the environment. Also, shadows play an important role of in portraying virtual objects as realistic as possible [26]. In order to understand the effectiveness of shadows, we first need to understand the type of shadows perceived by humans in natural environment. Shadows can be classified into two groups, i.e., cast shadows and attached shadows [18]. Attached shadows are those which are formed when an object obstructs the light falling on itself. In contrast, cast shadows are those which are formed when one object occludes another object from the light source, which results in an image being formed on the occluded object [18].

Shadows are effective means of perceiving distance in an environment and also perceiving motion and change in direction of motion of an object. Helen Hu and others(2002)in their experiments showed the effectiveness of shadows in estimating distance in an immersive environment. Where, shadows in combination with stereoscopic vision showed a significant effect in estimating distance and shadows alone were also effective without stereo [11]. On the other hand, Kersten and others (1997) in their experiments using 3D graphics showed that shadows clearly provide very salient cues for the relative spatial position of objects in an environment, particularly when an object and its shadow are moving [14]. We saw many empirical evidence of how shad-
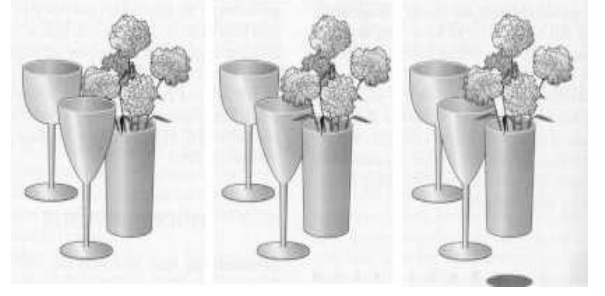
Figure 7: Middle: Contradicting Occlusion and Relative Height Cues, Right: Shadows resolving the contradiction.

ows are effective in understanding distance and spatial arrangement of objects in motion.

Looking from a neural psychological perspective, Mamassian and others (1998) study showed that the perceptual constraints plays an important role in understanding the relative motion in a scene. The constraints mentioned involve positions of the viewpoint, the light source, the casting object and the background surface [18]. They also showed that cast shadows have a strong impact on the perception of spatial layout, the reason being that the straightforward assumptions allow the visual system to make reliable inferences about the relative position of objects [18]. Among the results shown by Mamassian (1988), there are still few unanswered questions like, What are the mechanisms which makes us select a particular constraints that is used by our visual system? Also, does the attached shadows are processed in the occipito-temporal stream and cast shadows in the occipito-parietal stream [18]? In regards to the view proposed by Goodale, Melvyn and Milner, i.e., there exist two streams of visual processing in which occipito-temporal stream deals with the shape of the objects and the occipito-parietal stream deals with the location in space [7].

5.0.5 *Limitations of shadow rendering in Augmented Reality*

In the previous section we saw the usefulness of shadows in perceiving depth information. But, among all its advantages we still face the problem of rendering shadows in real world environment. The problem depends on two factors for shadow rendering, i.e., the lighting condition of the environment [10] and the intent of real object (e.g., humans in environment)[12]. These two factors make it extremely difficult to achieve accurate rendering of shadows in real world environment. Although there have been some success in achieving real-time rendering of virtual objects in real world environment using shadow volumes [10]. This technique was derived from how shadows are rendered in virtual reality using shadow volume technique. Although, shadow rendering in virtual reality is quite successful, we still need to consider the fact that in virtual reality illumination, source of light and the content is in control of the graphic designer and some modifications are required to implement these algorithms in augmented reality.

5.0.6 *Inspiration from Art*

Art and artists for many years have been major contributors in understanding perception. Looking back into annuls of history, Leonardo da Vinci in 1490 described how artists could use light and shade to show perceptions of three-dimensional information in a two-dimensional paintings [18]. While there are constraints in utilizing virtual reality techniques in rendering shadows, we look at how artists resolve such issues. Specifically, we look at methodology behind retouching paintings which is used to recreate paintings. It is a method that is hugely successful and is practiced in galleries to preserve paintings. For example, Raphael a Italian High Renaissance painter successful restored the Madonna of Foligno [19].

Augmented reality is built up on the philosophy that information in real world is obscure and hard to perceive to accomplish a task. Virtual objects are augmented in the environment to fill in/recreate or to highlight particular information. On the similar principal, retouching of painting deals with recreating of an image once again, that has under gone wear-and-tear or loss of paint over time. The similarity in principle of both can help us in directing towards a possible solution. On our analysis of a case study by Bailao (2010) on methodology behind retouching paintings, we found that a great emphasis is given on understanding the painting that is to be retouched. Psychological analysis like gestalts law of contrast is used to understand the colour, contrast, illumination perception. Only after the analysis, a technique of retouching is proposed for example, mimetic technique and also the tools to be used [3]. The importance here in retouching is given to the amount of time spent in understanding the meaning behind the painting and understanding the psychological cues present in the painting.

A logical extension to this understanding can be applied to object rendering in dynamic environment. Here, a proposed principle that could be used to render shadows, is to first analyse the environment in question, i.e., to identify the location at which the object has to be rendered. Here, Gestalts law can be useful in understanding the relationship between existing objects before augmenting the virtual object. The next step could be to identify as similar object in a relationship to perform analysis, on the object to identify the contrast, colour and angle at which light falls on object, etc. After the completion of the analysis, the next step would be to apply real time techniques to render the virtual object and its shadow. While existing techniques like shadow volume focuses more on the object that is to be rendered rather than the environment it would render in [10].

6. DISCUSSION AND CONCLUSION

In this report, we looked at some of the main depth related perceptual issues in augmented reality on range of different platforms. While, the problems ranged from conflict of cues due to incorrect rendering to ergonomics, each platform is tailored to some specific problems. This draws our attention to the fact that solutions to the problem have to be specific to the platform than a general one.

We focused our attention more on perceptual impact of motion in an environment where change in position can influence the dominating cue. Many empirical evidences have been shown with regards to shadows being an effective cue to correctly perceive spatial relationship of an object with its environment. But, looking from psychological perspective we found more unanswered questions than solutions. There still remains a huge gap in understanding this phenomenon from psychological perspective. We argue that this gap in understanding perceptual issues from a psychological perspective is one of areas that require more study and focus for augmented reality to improve. Azuma(1997) in this paper drew attention to the fact that psychological studies can help improve augmented reality by helping us understand questions like how much registration error is detectable and tolerable by the viewer that will impact the performance of the systemv[2]. Filling up this gap will help us model systems that are cogent with human neural system. Where, impact of psychological studies in improvement of task performance has previously been seen in menu designing [22]. For example, miller's law of maximum number of items a human can hold in his memory [21].

Shadows have also been a major contributor in infusing realism in the augmented virtual object [25]. Realism is an important aspect in making the environment immersive, which is an essential aspect from a usability point of view in augmented reality. With the motive of infusing realism, developer often neglect the context of the environment in which object is being augmented. For example, realism is a crucial factor when augmenting objects while the user is driving. In this scenario, the objective is not to divert attention of the user from task at hand to the virtual objects. Here, the object should represent itself as a part of the environment using realistic rendering, e.g., using shadows. On the other hand, in museums augmented objects has to have a clear distinction between objects that are part of history and virtual objects that aid in perceiving them. Here, the objective is to draw attention to the overlaying information which can be done using toon shading [9]. Understanding the context in which objects are rendered is another area of focus to avoid perceptually incorrect augmentation [20].

To conclude, perceptually correct augmented reality still remains a difficult problem which was shown throughout the paper. We identified in our understanding two neglected areas in augmented reality and also suggested a principle to help in rendering of shadows.

7. REFERENCES

[1] *Virtual Reality-Human Computer Interaction*. InTech, 2012.
[2] R. T. Azuma et al. A survey of augmented reality. *Presence*, 6(4):355–385, 1997.
[3] A. Bailão. Application of a methodology for retouching. a case study of a contemporary painting. In *CeROArt. Conservation, exposition, Restauration d'Objets d'Art*. CeROArt asbl, 2010.
[4] K. Čopič Pucihar, P. Coulton, and J. Alexander. Evaluating dual-view perceptual issues in handheld augmented reality: device vs. user perspective rendering. In *Proceedings of the 15th ACM on International conference on multimodal interaction*, pages 381–388. ACM, 2013.
[5] A. Dey, G. Jarvis, C. Sandor, and G. Reitmayr. Tablet versus phone: Depth perception in handheld

augmented reality. In *Mixed and Augmented Reality (ISMAR), 2012 IEEE International Symposium on*, pages 187–196. IEEE, 2012.

[6] D. Drascic and P. Milgram. Perceptual issues in augmented reality. In *Electronic Imaging: Science & Technology*, pages 123–134. International Society for Optics and Photonics, 1996.

[7] M. A. Goodale and A. D. Milner. Separate visual pathways for perception and action. *Trends in neurosciences*, 15(1):20–25, 1992.

[8] R. L. Gregory. *Eye and brain: The psychology of seeing* . Princeton university press, 1997.

[9] M. Haller. Photorealism or/and non-photorealism in augmented reality. In *Proceedings of the 2004 ACM SIGGRAPH international conference on Virtual Reality continuum and its applications in industry*, pages 189–196. ACM, 2004.

[10] M. Haller, S. Drab, and W. Hartmann. A real-time shadow approach for an augmented reality application using shadow volumes. In *Proceedings of the ACM symposium on Virtual reality software and technology*, pages 56–65. ACM, 2003.

[11] H. H. Hu, A. A. Gooch, S. H. Creem-Regehr, and W. B. Thompson. Visual cues for perceiving distances from objects to surfaces. *Presence: Teleoperators and Virtual Environments*, 11(6):652–664, 2002.

[12] C. E. Hughes, J. Konttinen, and S. N. Pattanaik. The future of mixed reality: Issues in illumination and shadows. In *The Interservice/Industry Training, Simulation & Education Conference (I/ITSEC)*, volume 2004. NTSA, 2004.

[13] S. Ichihara, N. Kitagawa, and H. Akutsu. Contrast and depth perception: Effects of texture contrast and area contrast. *PERCEPTION-LONDON-*, 36(5):686, 2007.

[14] D. Kersten, P. Mamassian, D. C. Knill, et al. Moving cast shadows induce apparent motion in depth. *PERCEPTION-LONDON-*, 26:171–192, 1997.

[15] J. M. Knapp and J. M. Loomis. Limited field of view of head-mounted displays is not the cause of distance underestimation in virtual environments. *Presence: Teleoperators and Virtual Environments*, 13(5):572–577, 2004.

[16] E. Kruijff, J. E. Swan II, and S. Feiner. Perceptual issues in augmented reality revisited. In *ISMAR*, volume 9, pages 3–12, 2010.

[17] J. S. Lappin, A. L. Shelton, and J. J. Rieser. Environmental context influences visually perceived distance. *Perception & psychophysics*, 68(4):571–581, 2006.

[18] P. Mamassian, D. C. Knill, and D. Kersten. The perception of cast shadows. *Trends in cognitive sciences*, 2(8):288–295, 1998.

[19] A. McClellan. Raphael's foligno madonna at the louvre in 1800: Restoration and reaction at the dawn of the museum age. *Art Journal*, pages 80–85, 1995.

[20] E. Mendez. On the usage of context for augmented reality visualization. *Phd, Graz University of Technology, Graz, Austria*, 2011.

[21] G. A. Miller. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, 63(2):81, 1956.

[22] K. L. Norman. *The psychology of menu selection: Designing cognitive control at the human/computer interface.* Intellect Books, 1991.

[23] R. A. Rensink, J. K. O'Regan, and J. J. Clark. To see or not to see: The need for attention to perceive changes in scenes. *Psychological science*, 8(5):368–373, 1997.

[24] R. Rosenholtz, Y. Li, and L. Nakano. Measuring visual clutter. *Journal of Vision*, 7(2):17, 2007.

[25] M. Slater, P. Khanna, J. Mortensen, and I. Yu. Visual realism enhances realistic response in an immersive virtual environment. *Computer Graphics and Applications, IEEE*, 29(3):76–84, 2009.

[26] N. Sugano, H. Kato, and K. Tachibana. The effects of shadow representation of virtual objects in augmented reality. In *Mixed and Augmented Reality, 2003. Proceedings. The Second IEEE and ACM International Symposium on*, pages 76–83. IEEE, 2003.

[27] M. Teittinen. Depth cues in the human visual system. *see http://www. hitl. washington. edu/scivw/EVE/III. A*, 1, 1993.

[28] D. Van Krevelen and R. Poelman. A survey of augmented reality technologies, applications and limitations. *International Journal of Virtual Reality*, 9(2):1, 2010.

[29] C. Ware. *Information visualization: perception for design.* Elsevier, 2012.