DESIGN ISSUES IN HUMAN VISUAL PERCEPTION EXPERIMENTS ON REGION WARPING

Yang-Wai Chow, Ronald Pose and Matthew Regan

School of Computer Science and Software Engineering, Monash University
Clayton, 3800 Victoria, Australia
yang.wai.chow@csse.monash.edu.au
ronald.pose@csse.monash.edu.au

ABSTRACT

Virtual reality systems are primarily concerned with the presentation of realistic 3D graphics to the user. In light of the fact that the human visual system can only perceive a finite amount of detail, it is therefore possible to reach a balance in the tradeoff between human perception of detail and computational load required for rendering. Priority rendering is a technique designed to reduce the overall rendering load for an address recalculation pipeline virtual reality system. This system was developed to reduce user perceived latency during head rotations in head mounted display virtual reality systems. Large object segmentation and region warping were methods introduced to priority rendering in order to further enhance the overall rendering load reductions and to deal with various visual artefacts. This paper discusses the issues and considerations involved in designing human factors experiments concerning human perception to region warping.

KEYWORDS

Distortions, human visual perception, priority rendering, region warping.

1. INTRODUCTION

Immersive Head Mounted Display (HMD) virtual reality systems attempt to place a user in realistically believable synthetic virtual worlds. This is done by presenting the user with computer generated 3D graphics from the user's vantage point in real-time. A major component of presence in a virtual environment, the sense of actually existing and interacting in the alternate reality created by the system, lies in the visual realism of the scene that the system presents to the user. In order to construct an illusion of reality, a variety of considerations are involved in the designing of virtual reality systems, ranging from technical issues to human factors.

One of the design challenges faced by HMD virtual reality system developers is to overcome the end-to-end latency, defined as the time between a user's actions and when those actions are reflected by the display, suffered by the user in such systems. The end-to-end latency has a great impact on a user's performance in the virtual environment and can consequently destroy the illusion of reality that the system attempts to create. The effects and characteristics of a virtual reality system's latency, from a user's perception, have been well documented by a number of researchers [Mania et al. 2004; Meehan et al. 2003].

Visual realism of a virtual environment depends on the amount of detail and complexity portrayed in the scene. The use of techniques such as lighting models and artistic enhancements to the scene, have been found to improve the perceived realism of virtual environments [Greenburg et al. 1997; Longhurst et al. 2003]. However the more detail contained in a scene, the more computationally expensive the processing and rendering requirements of the scene. This can result in lengthy delays to the system's display, further adding to the latency experienced by a user. In view of the fact that the human visual system can only perceive a limited amount of detail, attempts have been made to reduce scene complexity by removing non-perceptible components from the computer generated graphics [Reddy 2001].

In order to achieve adequate performance in designing virtual reality systems from resource limitations in terms of processing power and speed of computations, there needs to be a tradeoff between desired performance and computational load. Due to these limitations, it is desirable to decrease the level of detail found in scenes whilst still maintaining acceptable visual realism. Typically, this will involve a compromise between human factors and computational efficiency [Wann and Mon-Williams 1996].

A graphics display architecture known as the Address Recalculation Pipeline (ARP) [Regan and Pose 1993] has been designed specifically for HMD virtual reality systems in order to reduced head rotational latency suffered by these systems. Priority rendering is a technique developed for use in conjunction with this system, to reduce the overall rendering load and consequently allow for the rendering of more complex and realistic scenes [Regan and Pose 1994]. The segmentation of large virtual world objects for priority rendering has been found to further reduce the overall rendering load. However unwanted scene tearing artefacts emerged as a byproduct of using this approach. Region warping was devised by introducing slight distortions to the scene, in order to hide the scene tearing artefacts [Chow et al. 2005].

In view of the fact that the usability of the system depends on a user's perception of realism, it is important to determine whether these distortions are perceivable to a user, as well as how these distortions might affect the visual realism of a virtual environment. A number of human factors considerations regarding the implementation of priority rendering have been presented in Pose and Regan [1994]. This paper discusses issues that have to be contemplated in the designing of experiments for the purpose of characterizing the human visual perception with respect to region warping.

2. BACKGROUND

This section presents some background to the ARP virtual reality system and priority rendering, as well as the basis for the implementation of region warping.

The ARP is a graphics display controller designed to reduce the latency experienced by a user during user head rotations in HMD virtual reality systems. This hardware architecture is fundamentally different from conventional systems in that viewport orientation mapping is performed post rendering, rather than as the scene is being rendered. In this manner, the user's head orientation is detached from rendering process. This means that the usually lengthy rendering time is removed from the latency experienced by a user during head rotations, as the latency is now bound to the display device's update rate and no longer tightly coupled to the rendering frame rate.

Using this approach, the scene that encapsulates the user's head has to be pre-rendered to display memory. Viewport mapping is then performed on relevant portions of the scene already residing in display memory, based on the latest up-to-date user head orientation information obtained from the head tracking device. This gave rise to a rendering technique known as priority rendering, to be used in conjunction with the ARP virtual reality system in order to reduce the overall rendering load.

The concept behind priority rendering is that when the user's head orientation changes, it might not be necessary to re-renderer everything in the scene that surrounds a user's head. In general, sections of a scene located closer to the user will appear to change more rapidly compared to sections further away from the user. Therefore by using multiple display memories and multiple renderers, different sections of a scene that surrounds the user can be rendered onto the separate display memories at different update rates. These images on the different display memories can then be combined to form an image of the complete scene. Figure 1 shows an illustration of image composition, where visible sections of the images from the separate display memories are overlaid to form the final image of the scene. This significantly reduces the overall rendering load of the system, thereby potentially allowing for the rendering of more complex scenes. For further information regarding the ARP and priority rendering refer to Regan and Pose [1993; 1994].

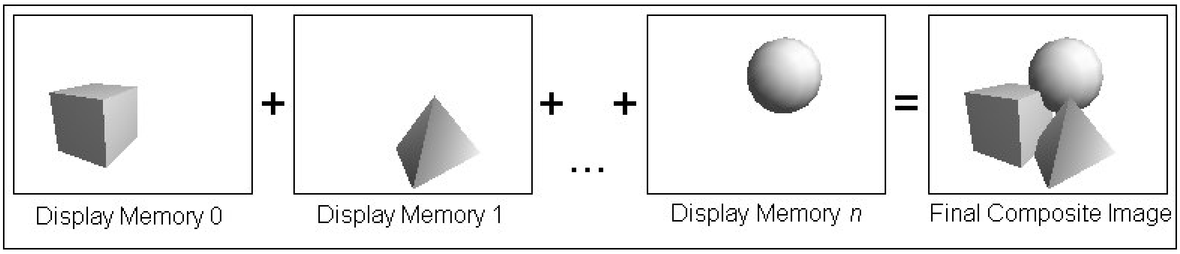

Figure 1: Image Composition.

It has been shown that the overall rendering load can be been further reduced by segmenting large objects in the virtual world into smaller objects for priority rendering. Each of these smaller segments may then be treated as individual objects and can therefore be updated individually onto separate display memories at different update rates. A method of spatially dividing the virtual world into regions has been used to ease the management of virtual world objects for priority rendering, and the region boundaries have been used as the criterion for the segmentation of large objects [Chow et al. 2005].

Virtual world object segmentation however introduces scene tearing artefacts, which appear at certain instances as a user translates through the scene. These unwanted artefacts materialize between the shared vertices of segmented objects, and emerge as a result of updating different segments of the same object at different times. A solution to the problem, known as region warping, was devised in order to mask the unwanted scene tearing artefacts.

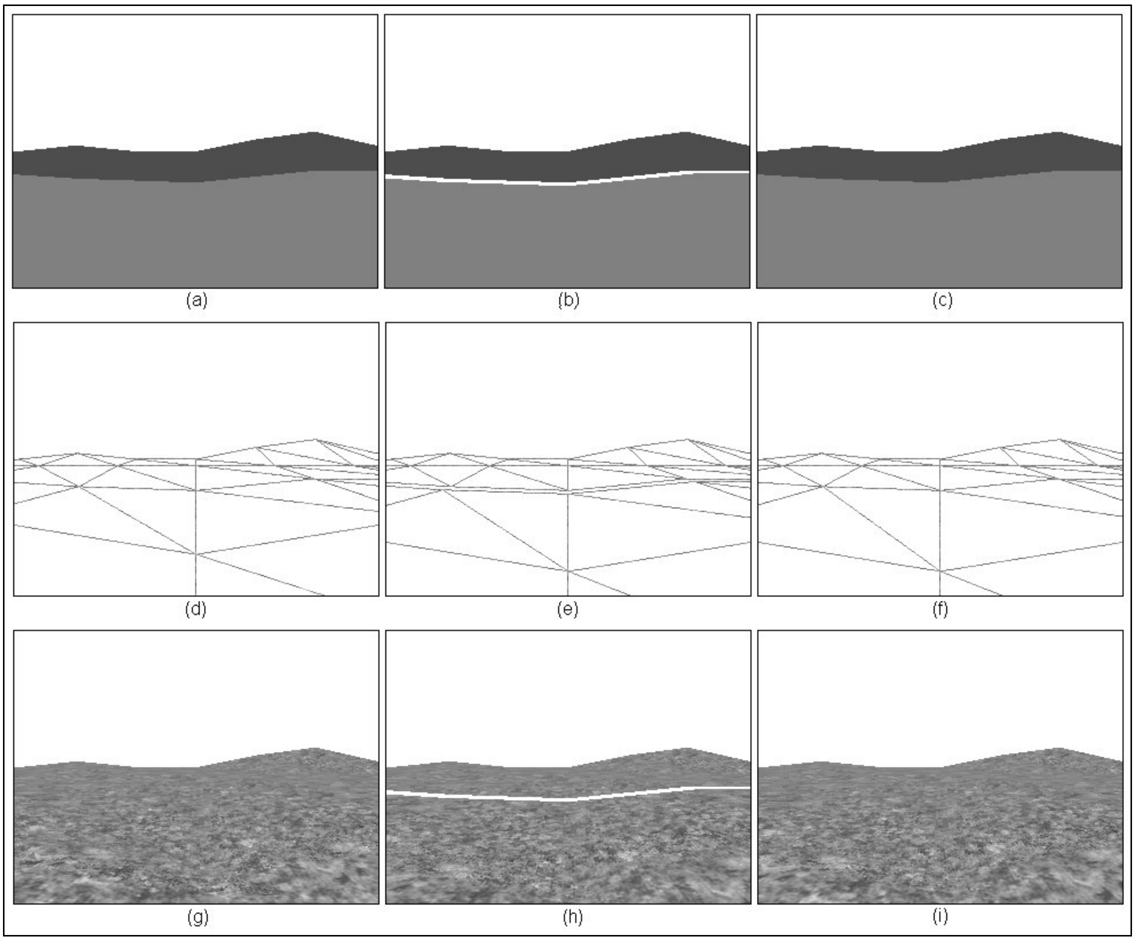

Figure 2: Illustration of tearing artefacts and region warping.

Figure 2 is a depiction of a simple fractal terrain scene used for illustration purposes in explaining scene tearing and region warping. In the figure, (a) – (c) were rendered using 2 distinct colors to differentiate the different regions rendered onto two different display memories; the closer region was rendered onto a display memory with a higher update rate while the further region was rendered at a slower update rate. (d) – (f) were rendered in wireframe so that the tearing and region warping effects can clearly be seen, and finally (g) – (i) show the fully textured scene.

Figure 2 (a), (d) and (g) were rendered at an arbitrary time t. At time $t + 1$, the user has moved to a new position and the closer section of the scene is re-rendered. However the further section will only be updated at a later stage and its display will remain the same as at time t. Therefore compositing the images from the two display memories will give rise to the tearing artefact which can be seen in (b), (e) and (h). Using the region warping technique, the vertices across the closer region are essentially perturbed slightly to force the shared vertices between the segments to align, the end result of region warping is shown in (c), (f) and (i). In this respect, the further region does not have to be re-rendered as often as the closer region, thus reducing the overall rendering load. However, this introduces slight distortions in the scene as a result of implementing region warping.

It has been observed that even in virtual reality systems that have perfect head tracking, a user will perceive virtual world objects to warp and shift when the user moves his/her head [Wartell et al. 1999]. It is therefore conceivable that the user might not be able to perceive any differences due to the slight distortions cause by region warping. This approach attempts to take advantage of a user's inattention to slight distortions in the scene in order to reduce the overall rendering load of the system, and potentially be able to utilize these rendering load reductions to render more detailed and realistic scenes.

3. RELATED WORK

A number of human perception experiments have been conducted on various perceptually based computer graphics rendering techniques. These techniques have been used with the intention of reducing the computational effort and time required in the rendering of a scene or the generating of images, by taking advantage of certain limitations in the psychophysical aspects of the human visual perception. Examples of these human visual perception exploits can also be found in image and video compression techniques such as JPEG and MPEG coding. These lossy compression techniques decrease the amount of required data by reducing less important information such as the chrominance color component, whilst not greatly affecting the quality in the images or videos from a user's perception.

Level of Detail (LOD) management is a perceptually based graphics technique used to optimize a system's performance by removing less or non-perceptible details from the scene, as the human visual system is not able to perceive all the detail in its environment. Watson et al. [1997] have experimented on the effects of degrading peripheral detail in a HMD virtual reality system's display, with regards to user performance. This was done by using high-detailed resolution insets and degrading the HMD's peripheral resolution. In their experiment the user was required to perform a visual search task. Based on their human perception experiments, they found that degrading the peripheral detail had no adverse effects on the visual search task performance.

Other LOD methods have also been used to display geometric objects at different resolutions, for example, by varying the number of polygons constituting an object. The implementation of these different levels of complexity such as using simpler object meshes, can significantly improve a system's rendering performance. The challenge however is to be able to switch between the various LOD without the user being aware of any visual alterations. Numerous selection criteria have been proposed and implemented in object LOD management techniques, these include the use of objects' distance, size, velocity and eccentricity [Reddy 1998].

In a virtual environment, a user will typically focus on conspicuous objects in the scene and will ignore or pay less attention to details in the rest of the environment. This is a fundamental feature of the human visual system known as inattention blindness. The center of the retina known as the fovea has the densest concentration of color sensitive cones in the human eye. The visual angle covered by the fovea however is very small, and everything that lies outside this foveal angle is perceived as blurred or unclear. Visual perception experiments have also been conducted using visual attention models to predict where the user will be looking in the scene and to selectively concentrate computational effort to these foveal regions of the scene [Chalmers et al. 2003].

Visual attention models have also been used in LOD management methods. Brown et al. [2003] have developed a theoretical framework for an attention based LOD management system. This system is design to use the attentive features of a scene such as size, position, motion and luminance, in order to predict a viewer's eye movements. In this way, the system can adaptively select the complexity level for an object to be rendered based on the object's visual importance.

These experiments have shown that the human visual system has a finite perception to detail in a scene, and that it is possible for computer graphics systems to take advantage of these human perception limitations to improve computational efficiency and performance.

4. DESIGN ISSUES

There are a number of issues that have to be considered in the designing of human factors experiments for understanding and characterizing a user's visual perception of region warping. These various design issues are presented below.

4.1 Stereoscopy versus monoscopy

As part of the human visual system, stereo vision is one of the main factors that contribute to a user's perception of realism in a virtual world [Lo and Chalmers 2003]. Stereoscopic display systems provide certain depth cues that are required by the human visual system to obtain depth information from the scene. Experiments have shown that the absence of visual depth cues in monoscopic systems can greatly affect the speed and accuracy of task performances [Rosenberg 1993]. It is possible that distortions caused by region warping might appear more prominently to the user when a virtual environment is viewed in stereo, as compare to when the same environment is viewed on a monoscopic display.

4.2 Distance from the user

A human eye perceives less detail in objects located further away from the user. In addition, distant objects will also be smaller in size from the user's perspective. Human factors experiments have shown that due to various reasons a user cannot accurately perceived distances from the display of a virtual reality system [Plumert et al. 2004; Willemsen et al. 2004]. In addition, various LOD management techniques have been successfully used by others to degrade the complexity of distant objects with little or no apparent adverse effects to the user. It might then be feasible to concentrate region warping distortions further away from the user, at a distance where such distortions will be less perceptible to the user.

4.3 Head tracked displays

When a person moves his/her head, based on feedback from the motoric senses the human brain will expect the image of the surrounding environment captured on the retina to change in certain ways. This same expectation also applies to a user using a virtual reality application. This means that a user will have a pre-defined expectation of how the virtual environment is suppose to alter with respect to the user's head movements. In light of this, a user might perceive region warping differently on a head tracked system as opposed to a non-head tracked system. The speed at which a user's translation through the scene might also have an effect on the level of perceived distortions, as the human eye is less sensitive to details that move rapidly over the retina [Reddy 2001].

4.4 Scene complexity

In a complex scene, the human visual system depends of the retina to cope with the wealth of information contained within the scene [Chalmers et al. 2003]. It is therefore not surprising if a scene with a high level of detail will distract and overwhelm a user's perception to the region warping distortions. A user's focus of visual attention might also be drawn towards certain areas of interest in the virtual environment. In this respect, it is conceivable that the level of perceive distortion might vary from scene to scene depending on the amount of detail encompassed within the scene.

4.5 Virtual environment characteristics

The makeup of the virtual world might also affect a user's perception to distortions. Warping effects might appear differently in natural random scenes compared to scenes containing rigidly structured architecture. For example, the distortions caused by region warping might not be perceivable in an outdoor landscape scene containing natural objects such as rocks and plants, but could potentially be perceived in the walkthrough of a building made up of structured objects like walls and pillars. This is because rigidly structured objects are usually made up of well defined geometry, whereas natural objects have certain random unpredictable characteristics that might cause distortions to be less noticeable.

4.6 Scene familiarity

A part of the human visual system is governed by experiences acquired in the real world, as the human brain interprets images presented to the retina based upon these experiences. An example of this can be observed from optical illusions, where the human brain tends to interpret 2D images from a 3D perspective, or how a person might infer depth information from the use of shadows and lighting in a 2D image. This is illustrated in figure 3 where the human visual system perceives depth information from the picture, due to real world experiences with shadows and lighting. Therefore based on experiences if a person is presented with a familiar scene, the human brain might potentially overlook errors in the scene by unconsciously correcting or 'filling in the details'. Conversely the opposite might apply, in that the due to scene familiarity a person might be particularly prone to noticing discrepancies between the observed virtual scene and the scene recalled from memory.

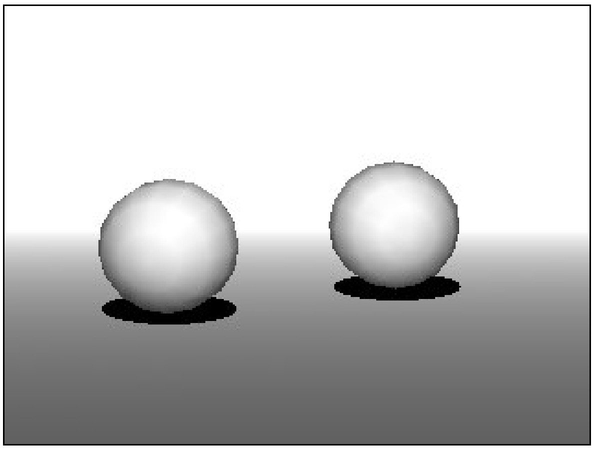

Figure 3: Visual depth cues perceived through the use of shadows, shading and lighting.

4.7 Ability to interact with surrounding environment

It is possible that a user's ability to interact with his/her surrounding environment might make a difference to the user's perception to details. The ability to be able to interact with the environment rather than merely observing the scene from a visual perspective adds to the feeling of presence in a virtual environment. The method and extent of the user's interaction ability is also important, as the user's mindset might be focused on manipulating certain objects in the virtual environment or navigating through the scene. At the same time, the user's focus of attention might also be consumed with how to operate the interaction device in order to achieve the desired outcome.

5. EXPERIMENTAL METHODOLOGY

Human factors experiments for each of the design issues described in the previous section can be performed individually by using a variety of different methodologies. These methods range from the use of various hardware devices to using virtual environments made up of different characteristics, as well as varying certain aspects of the virtual scene. This section discusses how experiments might be designed to test the considerations listed above as well as how we intend to setup our experiments.

In order to test a user's monocular and binocular perception to the region warping distortions, as described in section 4.1, a number of hardware display devices may be used. Examples of such display devices include HMDs with single and/or dual input channels for mono input and/or for individual left and right eye inputs, stereoscopic glasses for desktop virtual reality and CAVE systems, etc. We intend to setup experiment using HMDs with dual input channels to test a user's perception to distortions when the scene is displayed in both mono and stereo. The user will also be required to view the scene on a large screen display, with and without the use of stereoscopic glasses. The purpose of using the different hardware equipment is to test whether a human will perceive distortions differently on HMD systems compared to systems with large screen displays.

A user's perception to distances in the virtual environment will also depend on the type of display device used by the virtual reality system. It has been found that a user is more likely to underestimate distances when using HMDs, possibly due to its limited Field of View (FOV), as compared to large screen immersive displays [Plumert et al. 2004]. Therefore the same experimental setup using the HMD and large screen displays mentioned above, will also be used for testing the issue raised in section 4.2. For the purposes of investigating distortion perception with respect to distance, the user will be presented the same virtual environment with the warping effect varied at different distances from the user's viewpoint.

Different hardware devices are also required when testing human visual perception whilst interacting in the virtual environment. These interaction devices might consist of using the keyboard and mouse, joysticks, data gloves, pointers, magnetic or mechanical head trackers and so on. For our purposes we will be combining sections 4.3 and 4.7 in our experiments, where a head tracked HMD system will be setup for the user to interact with the scene based upon his/her head orientation. The keyboard and mouse will be used for navigating through the scene. In order to compare the cases with and without head tracking, the user will also be presented with a fixed navigation path through the scene, like watching a movie played out on the display device.

All these experiments will be repeated using a number of different virtual environment scenarios, for the purposes mentioned in sections 4.4, 4.5 and 4.6. A variety of diverse virtual worlds will be created, comprising of an assortment of different characteristics, in order to understand whether these various virtual environment characteristics affect the user's perception the warping effects. The various scenarios that we propose to create will be that of an imaginary building, a familiar building modeled from a real world building, an imaginary scene of another planet and a forest scene. These scenes range from highly structured environments with objects made up of rigid geometries, to random nature scenes containing organic objects. These environments also include familiar everyday life scenarios as well as fictional imaginary environments that cannot be experienced in reality.

To test for scene familiarity, the user will be exposed to the real environment in the real world, and then be presented with the virtual environment modeled from the actual scene. The user will also be asked to perform certain tasks in the real and virtual scenarios, in order to enhance his/her familiarity with the environments. The complexity level in these scenes will be varied by manipulating and adjusting the level of detail of objects presented in the scenes, and also by adding and removing virtual world objects to and from the virtual environments. The speed at which the user is permitted to translate through the virtual scene will also be regulated and varied, to ascertain whether translational speed will make any difference to the user's visual perception.

6. CONCLUSION

Region warping attempts to take advantage of limitations in the human visual system, in order to further reduce the overall rendering load in an ARP virtual reality system with priority rendering. Before this technique can effectively be used in conjunction with large object segmentation, it is important to determine and characterize the level of distortions perceivable by the user. While the image processing techniques we have developed appear to work well on computer monitors and display good performance in analytical measures, the real test can only be with the real human perception system.

This paper discusses design issues involved in developing experiments to study the human visual perception to region warping and to possibly ascertain a tolerable distortion threshold. This paper has also outlined our work and the approaches we will be using to verify its validity in the entail of a virtual reality system. We are currently in the process of setting up human factors experiments based on the presented design considerations and methodologies. We encourage others to undertake independent verification studies or to collaborate with us.

REFERENCES

Brown, R., Cooper, L. and Pham, B., 2003. Visual Attention-Based Polygon Level of Detail Management. *Proceedings of the 1st International Conference on Computer Graphics and Interactive Techniques in Australasia and South East Asia*, Melbourne, Australia, pp. 55-62.

Chalmers, A., Cater, K. and Mafioli, D., 2003. Visual Attention Models for Producing High Fidelity Graphics Efficiently. *Proceedings of the 19th Spring Conference on Computer Graphics (SCCG)*, Budmerice Castle, Slovak Republic, pp. 47-54.

Chow, Y.W., Pose, R. and Regan, M., 2005. Large Object Segmentation with Region Priority Rendering. *Accepted to The 28th Australasian Computer Science Conference 2005 (ACSC '05)*, Newcastle, NSW, Australia.

Greenburg, D.P., Torrance, K.E., Shirley, P., Arvo, J., Ferwerda, J., Pattanaik, S.N., Lafortune, E., Walter, B., Foo, S.C. and Trumbore, B., 1997. A Framework for Realistic Image Synthesis. *Proceedings of SIGGRAPH '97, in Computer Graphics, Annual Conference Series*, pp. 477-494.

Lo, C.H., and Chalmers, A., 2003. Stereo Vision for Computer Graphics: The Effect that Stereo Vision has on Human Judgments in Visual Realism. *Proceedings of the 19th Spring Conference on Computer Graphics (SCCG)*, Budmerice Castle, Slovak Republic, pp. 109-117.

Longhurst, P., Ledda, P. and Chalmers, A., 2003. Psychophysically based Artistic Techniques for Increased Perceived Realism of Virtual Environments. *Proceedings of the 2nd International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa (AFRIGRAPH '03)*, Cape Town, South Africa, pp. 123-132.

Mania, K., Adelstein, B.D., Ellis, S.R. and Hill, M.I., 2004. Perceptual Sensitivity to Head Tracking Latency in Virtual Environments with Varying Degrees of Scene Complexity. *Proceedings of the 1st Symposium on Applied Perception in Graphics and Visualization (APGV)*, Los Angeles, California, United States, pp. 39-47.

Meehan, M., Razzaque, S., Whitton, M.C. and Brooks, F.P., 2003. Effect of Latency on Presence in Stressful Virtual Environments. *Proceedings of IEEE Virtual Reality (IEEE VR '03)*, Los Angeles, California, United States, pp. 141-148.

Plumert, J.M., Kearney, J.K. and Cremer, J.F., 2004. Distance Perception in Real and Virtual Environments. *Proceedings of the 1st Symposium on Applied Perception in Graphics and Visualization (APGV)*, Los Angeles, California, United States, pp. 27-34.

Pose, R. and Regan, M., 1994. Techniques for Reducing Latency with Architectural Support and Consideration of Human Factors. *Lecture Notes in Computer Science 1077. Multimedia, Hypermedia, and Virtual Reality - Models, Systems, and Applications. First International Conference, MHVR'94 Selected Papers*, Springer-Verlag, Berlin Germany, pp.117-129.

Reddy, M., 1998. Specification and Evaluation of Level of Detail Selection Criteria. *Virtual Reality: Research, Development and Application*, Vol. 3, No. 2, pp. 132-143.

Reddy, M., 2001. Perceptually Optimized 3D Graphics. *IEEE Computer Graphics and Applications*, Vol. 21, No. 5, pp. 68-75.

Regan, M. and Pose, R., 1993. An Interactive Graphics Display Architecture. *Proceedings of the Virtual Reality Annual International Symposium (IEEE VRAIS '93)*, Seattle, United States, pp. 293-299.

Regan, M. and Pose, R., 1994. Priority Rendering with a Virtual Reality Address Recalculation Pipeline. *Proceedings of SIGGRAPH '94, in Computer Graphics, Annual Conference Series*, pp. 155-162.

Rosenberg, L.B., 1993. The Effects of Interocular Distance upon Operator Performance using Stereoscopic Displays to Perform Virtual Depth Tasks. *Proceedings of IEEE Virtual Reality Annual International Symposium (IEEE VRAIS '93)*, Seattle, United States, pp. 27-32.

Wann, J. and Mon-Williams, M., 1996. Design Issues in Virtual Environment Displays: A Formal Appraisal from the Perspective of Human Perception. *IEE Colloquium on Virtual Reality - User Issues*, pp. 2/1-2/4.

Wartell, Z., Hodges, L.F., and Ribarsky, W., 1999. Balancing Fusion, Image Depth and Distortion in Stereoscopic Head-Tracked Displays. *Proceedings of SIGGRAPH '99, in Computer Graphics, Annual Conference Series*, pp. 351-358.

Watson, B., Walker, N., Hodges, L.F. and Worden, A., 1997. Managing Level of Detail through Peripheral Degradation: Effects on Search Performance with a Head-Mounted Display. *ACM Transactions on Human-Computer Interaction*. Vol. 4, No. 4, pp. 323-346.

Willemsen, P., Colton, M.B., Creem-Regehr, S.H. and Thompson, W.B., 2004. The Effects of Head-Mounted Display Mechanics on Distance Judgments in Virtual Environments. *Proceedings of the 1st Symposium on Applied Perception in Graphics and Visualization (APGV)*, Los Angeles, California, United States, pp. 35-38.