Faculty of Informatics - Papers (Archive)

Faculty of Engineering and Information Sciences

1-1-2006

Human visual perception of region warping distortions

Yang-Wai Chow
Monash University, caseyc@uow.edu.au

Ronald Pose
Monash University

Matthew Regan
Monash University

James Phillips
Monash University

Recommended Citation

Human visual perception of region warping distortions

Abstract

Interactive virtual reality requires at least 60 frames per second in order to ensure smooth motion. For a good immersive experience, it is also necessary to have low end-to-end latency so that user interaction does not suffer from perceptible delays in images presented to the eyes. The Address Recalculation Pipeline (ARP) architecture reduces end-to-end latency in immersive Head Mounted Display (HMD) virtual reality systems. By using the ARP in conjunction with priority rendering, different sections of the scene are updated at different rates. This reduces the overall rendering load and allows for more complex and realistic scenes. Large object segmentation in conjunction with priority rendering further reduces the overall rendering load. However, scene tearing artefacts potentially emerge and region warping was devised to alleviate this. In compensating for the tearing, region warping introduces slight distortions to the scene.Immersive virtual reality systems have humans as integral parts of the system. While researchers do thorough measurements and evaluation of hardware and software performance, the human experience and perception of the system is often neglected. This paper addresses this important issue. We describe our human visual perceptual experimental methodology in detail and present some initial results. Initial experiments in human visual perception of region warping distortions show interesting characteristics which lead us to propose further experimental investigations to clarify their significance.

Keywords

Disciplines

Physical Sciences and Mathematics

Publication Details

Human Visual Perception of Region Warping Distortions

Yang-Wai Chow[†], Ronald Pose[†], Matthew Regan[†], James Phillips[††]

School of Computer Science and Software Engineering[†] / Department of Psychology[††]
Monash University
Clayton, Victoria 3800, Australia

{yang.wai.chow,rdp,regan}@csse.monash.edu.au, jim.phillips@med.monash.edu.au

Abstract

This paper presents human visual perception experiment results for a computer graphics rendering technique introduced in ACSC'05. In order to achieve a good immersive virtual reality experience, it is necessary to have at least 60 frames per second to ensure smooth motion. It is also necessary to have low end-to-end latency so that user interaction does not suffer from perceptible delays in images presented to the eyes. The Address Recalculation Pipeline (ARP) architecture reduces end-to-end latency in immersive Head Mounted Display (HMD) virtual reality systems. By using the ARP in conjunction with priority rendering, different sections of the scene are updated at different rates. This reduces the overall rendering load and allows for more complex and realistic scenes. Large object segmentation in conjunction with priority rendering further reduces the overall rendering load. However, scene tearing artefacts can emerge when different object segments are updated at different times. Region warping was devised to hide such tearing artefacts. In compensating for the tearing, region warping introduces slight distortions to the scene.

Immersive virtual reality systems have humans as integral parts of the system. While researchers do thorough measurements and evaluation of hardware and software performance, the human experience and perception of the system is often neglected. This paper addresses this important issue. We describe our human visual perceptual experimental methodology in detail and present some initial results. Initial experiments in human visual perception of region warping distortions show interesting characteristics which lead us to propose further experimental investigations to clarify their significance.

Keywords: Address Recalculation Pipeline, object segmentation, priority rendering, region warping, tearing artefacts, visual perception.

1 Introduction

The ultimate goal of virtual reality is to present the user with an illusion of reality within a virtual environment. One of the main aspects involved in the portrayal of a believable virtual environment is the presentation of computer generated 3D imagery to the user. In order to maintain the illusion of reality, virtual reality systems must continually display images from the user's vantage point in real time. Interactive virtual reality requires at least 60 updates per second to achieve good immersion.

Latency is a well recognized problem in virtual reality and teleoperation technology (Ellis et al. 2004). Latency is the lag or time delay between when a user performs an action and when the system responds to that action or when that action is represented by the system. A particular issue in virtual reality is that of the latency between a user moving his/her head, thus changing the user's viewpoint, and that change being reflected in updated images before the user's eyes. Excessive system latency or delays in virtual reality makes the system hard to use and in severe cases this can lead to adverse side effects such as user disorientation, motion sickness, and etc. Even with the fast graphics accelerators available today that can render over 100 frames per second (fps), latency still remains a factor that has to be addressed (Meehan et al. 2003). This is because on conventional systems the update cycle is bound by the need to obtain user head orientation information before rendering can commence.

A hardware architecture known as the Address Recalculation Pipeline (ARP) was designed to reduce user head rotational latency in immersive Head Mounted Display (HMD) virtual reality by detaching user head orientation from the rendering process (Regan and Pose 1993). Priority rendering was developed for use in conjunction with the ARP system in order to reduce the overall rendering load by concentrating rendering power on sections of the scene that appear to change the most. Using priority rendering, different sections of the scene can be updated at different rates. This allows for the rendering of more complex and realistic scenes (Regan and Pose 1994).

Large object segmentation and region priority rendering were introduced to manage objects in the virtual environment as well as to further reduce the overall rendering load (Chow et al. 2005a). However, tearing problems can emerge as a result of updating different segments of the same object at different rates. This problem was overcome by using region warping. Region warping involves the perturbation of object vertices in model space in order to force these vertices to align, thereby hiding scene tearing artefacts. This however introduces slight distortions in the computer generated graphics.

The quality of frames produced by implementing the region warping technique has previously been investigated using mathematical analysis (Chow et al. 2005b). However while mathematical analysis gave an indication of the level of distortion caused by region warping, it did not incorporate the human visual perception of the distortions. It is therefore essential to

conduct perception experiments in order to fully understand the characteristics of region warping distortions with respect to human visual perception. Virtual reality and interactive systems involve humans as integral parts of the systems. All too often researchers thoroughly test and measure the performance of the hardware and software while ignoring the fundamental issue of how humans might perceive and react to the system. This paper addresses this issue by presenting the experimental methodology and results of an experiment investigating the human perception of region warping distortions.

The results and insights gained in the research presented in this paper are also relevant in other areas and applications of computer graphics. There are a number of other perceptually based computer graphics techniques that attempt to optimize a system's performance taking advantage of the limitations in the human visual system. Some of these applications are provided in the related work section of this paper.

2   Previous Work

In order to describe the basis for region warping, this section first provides some background of priority rendering and the Address Recalculation Pipeline system. For more information please refer to Pose and Regan (1994), Regan and Pose (1993, 1994), Chow et al. (2005a).

2.1   The Address Recalculation Pipeline and Priority Rendering

The Address Recalculation Pipeline (ARP) graphics display architecture reduces the end-to-end latency perceived by the user during user head rotations, by implementing a concept known as delayed viewport mapping. In delayed viewport mapping, the scene that encapsulates the user's head is pre-rendered into display memory. Viewport orientation mapping is then performed only when required by mapping relevant sections of the scene already residing in display memory, and is therefore fairly independent of scene complexity and is based on the most up-to-date user head orientation information. Unlike conventional systems where user head orientation has to be obtained prior to rendering, the ARP system effectively detaches user head orientation from the rendering process. Thus in the ARP system, latency is no longer bound by the usually lengthy rendering process.

A rendering technique known as priority rendering was designed to be used in conjunction with the ARP system. Priority rendering takes advantage of the fact that in the ARP system, the scene surrounding the user's head has to be rendered into display memory. In a scene that surrounds the user's head, only dynamically animated objects might constantly be changing, when the user rotates his/her head. Most other objects in the scene will remain the same. When a user translates through a scene, sections of the scene that are closer to the user will appear to change faster than sections that are further away from the user. Therefore by using multiple display memories

and multiple renderers, different sections of the scene can be rendered onto separate display memories that can be updated individually at different rates. The images on the different display memories can then be composited to form an image of the whole scene.

Priority rendering is demand driven rendering, in that an object is not updated until its image in display memory has changed above a tolerable amount. This concentrates rendering power on sections of the scene that are changing the most. In this manner, priority rendering reduces the overall rendering load, thus potentially allowing for the rendering of more complex and realistic scenes. A threshold which defines the minimum feature size of the virtual environment has to be pre-determined. This threshold takes the form of an angle, θ_t. Priority rendering attempts to keep the image in display memory accurate to within this threshold. Figure 1 shows a priority rendering translational validity period estimation. The translational validity period is an estimate of how long the image of an object in display memory will remain valid before requiring an update, based on the user's relative translational speed. The translational validity period estimation was used in the experiment by varying the value of θ_t for the different scenes.
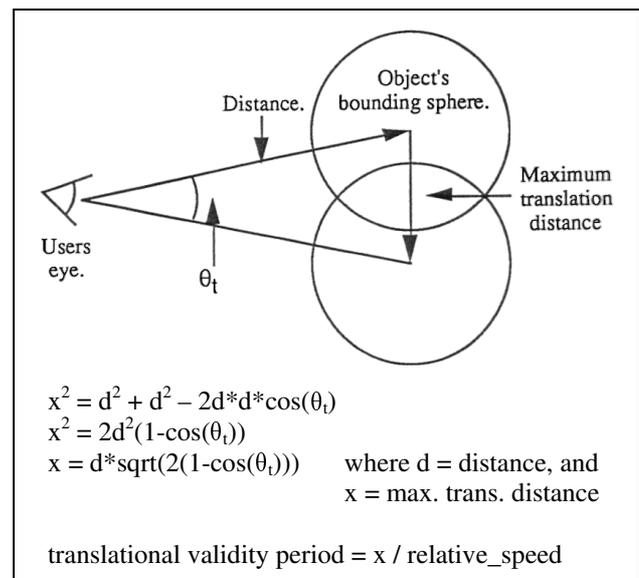

$$x^2 = d^2 + d^2 - 2d*d*\cos(\theta_t)$$
$$x^2 = 2d^2(1-\cos(\theta_t))$$
$$x = d*\sqrt{2(1-\cos(\theta_t))} \qquad \text{where } d = \text{distance, and}$$
$$x = \text{max. trans. distance}$$

translational validity period = x / relative_speed

Figure 1: Translational validity period estimation.

2.2   Large Object Segmentation and Region Priority Rendering

Region priority rendering was introduced to assist object management in the virtual environment for priority rendering. In region priority rendering, objects in the virtual world are spatially divided and grouped into different regions. Using this technique, entire regions of objects can then be assigned to the separate display memories based on the region that they are located in. This avoids having to calculate individual object validity periods and sort objects based on these validity periods before assigning individual objects to the separate display memories. Figure 2 shows a top-down example of region to display memory allocations for region priority

rendering. This allocation strategy was employed for the virtual environment used in the experiment.

Large object segmentation was devised to further reduce the overall rendering load in priority rendering. Large virtual world objects were segmented so that different segments of these objects could be updated on separate display memories at different update rates. In this manner, if the image of a section of a large object became invalid in display memory and required an update, only this section of the object would have to be updated instead of having to re-render the whole object.

2.3 Tearing and Region Warping

The implementation of large object segmentation with region priority rendering however causes an adverse visual artefact, a form of scene tearing. Scene tearing artefacts occur along the shared vertices of objects' segments. User head rotations will not cause the scene surrounding the user's head to appear to change much. The problem occurs when the user translates through the scene. Tearing artefacts emerge as a result of discrepancies between the shared vertices of an object's segments that are updated on separate display memories at different update rates, whilst the user is translating through the scene. This tearing problem can destroy the illusion of reality that the virtual reality system attempts to present to the user.
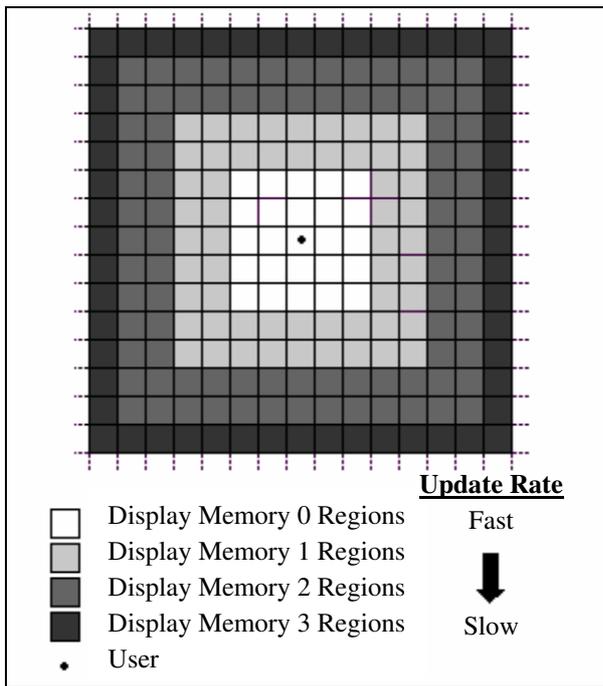

Figure 2: 2D top-down view of example region to display memory allocations.

Region warping was devised in order to alleviate the tearing problem. Region warping involves the perturbation of shared object vertices in order to force these vertices to align, thereby eliminating any discrepancy between the vertices during user translations. Two methods of interpolation for region warping were introduced in Chow et al. (2005b), linear interpolation and quadratic interpolation. Before warping could be

performed, the exact amount of perturbation had to be known. Vertices in the regions had to be normalized based on their distance from the user's region. This is illustrated in figure 3.

Normalization was performed using what can be seen as concentric squares, centered on the circumference of the square based region the user was located in. Interpolation of vertex perturbations could then be conducted with these normalized values. In linearly interpolated region warping, these normalized values were used without any alteration, whereas in quadratic interpolation the square of the normalized values was used. Both linear and quadratic region warping methods were tested in this experiment. In Chow et al. (2005b) it was concluded that the distortions in the frames generated using quadratic region warping were concentrated further away from the user and were more similar to normally rendered frames, compared to frames rendered using linear region warping. This conclusion however was based on mathematical comparisons and did not reflect human visual perception of the distortions.
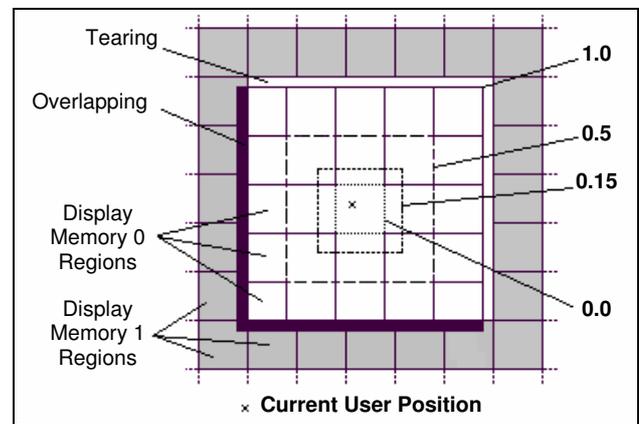

Figure 3: 2D top-down view illustrating the normalization of region vertices.

3 Related Work

Similar concepts of using multiple renderers and/or multiple display memories have also been designed and developed by other researchers.

NVIDIA Corporation's Scalable Link Interface (SLI) technology combines the rendering power of two Graphics Processing Units (GPUs) in a single system (NVIDIA 2005). The SLI system has a rendering mode called Split Frame Rendering (SFR) which allows a frame to be divided into two portions (top and bottom) and rendered separately on each GPU. Software drivers are used to dynamically share and balance the load between the two GPUs. Each GPU then renders one of the two sections, before the sections are digitally composited to form the whole frame. By clipping the scene into 2 portions, the system attempts to avoid the processing of all the vertices in a frame on both GPUs.

A 3D graphics and multimedia hardware architecture codenamed Talisman was designed by researchers at Microsoft Corporation (Toburg and Kajiya 1996). One of the main uses of the Talisman architecture was in multimedia applications such as interactive animation. In

smooth animated sequences, most of the display remains the same from frame-to-frame and it would be wasteful to have to re-render the entire frame. The Talisman architecture takes advantage of temporal and spatial similarities between sequential frames, by allowing individually animated objects to be rendered onto independent image layers before being composited together to form the final display (Barkans 1997). In this way, only changing image layers have to be modified or re-rendered.

3D warping techniques in computer graphics have also been looked into by a number of other researchers. Mark et al. (1997) have experimented on what they termed, post-rendering 3D warping, on adjacent frames in order to avoid re-rendering entire frames by exploiting frame-to-frame coherence. The purpose of this warping was to increase the frame rate of a graphics system. This was done by warping the images of existing frames in order to extrapolate for new viewpoints of future frames. Thus, less rendering had to be performed for the derived frames resulting in an increase in the rendering frame rate. They have also suggested that the priority rendering technique for the ARP system, and also the Talisman architecture, would benefit from the implementation of 3D warping techniques.

In light of the fact that the human visual system can only perceive a limited amount of detail, perceptually orientated graphics techniques have been designed to optimize a system's performance. For example, Level of Detail (LOD) management techniques attempt to remove or reduce less perceptible details from the computer graphics (Reddy 1997). Watson et al. (1997) have experimented on the effects of degrading the peripheral detail of a scene in a HMD virtual reality system's display with respect to user performance. Visual perception experiments using visual attention models have also been conducted to predict where a user will look in a scene and to selectively concentrate computational effort on those sections of the scene (Chalmers et al. 2003).

4 Psychophysics Experiment

The aim of this experiment was to measure the threshold where a human can perceive the distortions caused by region warping. Other goals of the experiment were to determine whether different display devices and/or region warping methods might affect the human perception of these distortions. Psychophysical methods of testing were therefore employed for the experiment. Psychophysics is a branch of psychology that deals with the measurement of perception. It is the scientific study of the relation between stimulus and sensation (Gescheider 1985).

4.1 Method

The method used for the experiment was a variation of the method of limits known as an adaptive staircase Two Alternative Forced Choice (2AFC) method. An adaptive procedure means that the stimulus of a trial is determined by the preceding stimuli and responses (Levitt 1971). The advantage of using this procedure is that trials will be concentrated around the area of interest, i.e. near the threshold, and therefore presents an efficient method for measuring threshold.

The staircase 2AFC technique has been widely used in a variety of different fields to measure threshold of perception. Experiments using similar methodologies have also been conducted in the field of virtual reality and computer graphics by other researchers, for example, to measure perception of latency in virtual reality (Regan et al. 1999, Ellis et al. 2004, Mania et al. 2004), and in perception of modulated Level-of-Detail (LOD) in computer graphics (Reddy 1997).

In a 2AFC discrimination task, two stimuli are presented to the participant - a standard (S) and a comparison (C) (Ulrich and Miller 2004). The participant is then required to choose which of these two alternatives, contained (or did not contain) the signal. The comparison stimulus presented for each trial varies in signal strength, and the number of different strength values is usually 5 or 7 (Farell and Pelli 1999).

The forced-choice method was chosen for the experiment over yes/no or same/different methods, as these other methods contain an internal subjective criterion (Farell and Pelli 1999). In a yes/no or same/different experiment, for each trial the participant would be asked whether they could discern a difference between the two stimuli or whether the stimuli were the same or different. This might introduce a certain bias to the experiment, for example, the participant's responses might be influenced if he/she knew the purpose of the experiment beforehand. The 2AFC method attempts to eliminate this bias, by forcing the participant to choose between the two alternatives.

The staircase or up-down method used for the experiment followed a 2-Down 1-Up (2D-1U) approach, where two correct responses reduced the signal strength whereas an incorrect response increased or augmented the signal strength. This method gives a 70.7% correct response threshold (Levitt 1971). Descending and ascending staircases were used in the experiment. A descending staircase is one which starts at a high/strong signal strength, and the signal strength is typically reduced depending on the participant's responses. An ascending staircase starts at a low/weak signal strength, and the signal strength is typically increased based on the participant's responses (Gescheider 1985). An example of ascending and descending staircases can be seen in figure 6, in the results and discussions section below.

4.2 Design

For this experiment, the standard stimulus was the scene rendered using normal rendering while the comparison stimulus was the scene rendered using the region warping technique, and containing varying levels of distortion. For each trial, the two scenes, scene A and scene B, were presented sequentially one after the other. This is known as temporal-forced choice (Gescheider 1985). The order of the two scenes (i.e. the standard and comparison stimulus) presented to the participants was pair-wise randomized. Participants were then forced to choose

which of the scenes appeared better, in other words, which scene did not contain distortions (the standard stimulus). Participants used a 2 button mouse to input their forced-choice feedback.

There were 2 parts to the experiment. Each part required the participant to carry out the experimental tasks on a different display device. The two display devices used were: a 17 inch computer monitor and a Head Mounted Display (HMD). The HMD used for the experiments was a Virtual Research V8 HMD, which has a 640×480 resolution and a 60Hz refresh cycle. The other display device was a 17 inch computer monitor set to a resolution of 640×480 and a 60Hz refresh cycle, in order to match the HMD's resolution and refresh rate. In view of the fact that the other device used was a computer monitor, the HMD was used in monoscopic mode. The wide angle lens optical distortion of the V8 HMD was corrected using a technique proposed by Watson and Hodges (1995).

A pair of staircases (one ascending, one descending) was used for each region warping technique (i.e. linear and quadratic), giving a total of four staircases. Both ascending and descending staircases were used in this experiment in order to comprehensively test the participants' responses to the full range of distortion levels. All four staircases were randomly interleaven based on a pre-randomized script. This was done so that participants would not know which warping method was being presented during each trial and also to minimize the chances of the participant anticipating or predicting the signal strength of the next trial. For example, if the participant realized that he/she was on an ascending or descending staircase he/she would be able to anticipate whether successive trials would increase or reduce in signal strength. Figure 4 depicts the overall experimental design. The entire experiment was automated on the computer.
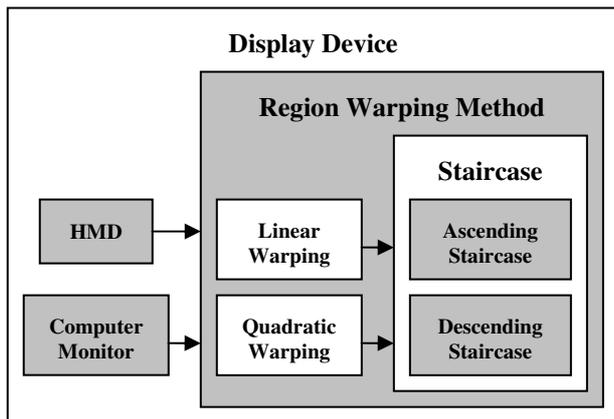

Figure 4: Overview of experimental design.

Fixed-step staircases with 7 signal strengths (distortion levels) were used. The signal strengths corresponded to the level of region warping distortions. Each staircase would end after 6 reversals or a maximum of 25 trials, whichever came first. A reversal refers to a change in the direction of a staircase. In a 2D-1U staircase approach, a reversal would mean 2 correct responses after 1 (or more)

incorrect responses or an incorrect response after 2 (or more) correct responses. An example of this is shown in figure 6, where the ascending staircase ended after 6 reversals while the descending staircase ended after 25 trials. This meant that each part of the experiment would finish after a maximum of 100 trials. It is important to note that unlike experimental simulations that can be done purely on a computer, in experiments involving humans one must design the experiment to avoid exhausting the participants.

The V8 HMD has a 60^0 diagonal Field of View (FOV), which gives a 48^0 horizontal FOV and a 36^0 vertical FOV. Since the HMD has a 640×480 resolution, this equates to 4.5 arc minutes per pixel (Mania et al. 2004). This means that a θ_t value of 4.5′ or 0.075^0 corresponds to 1 pixel. The 7 levels of distortion used in the experiment were obtained by varying θ_t (refer to θ_t in the translational validity period equation shown in figure 1) in step sizes of ½ a pixel. Table 1 shows the relations between distortion levels, θ_t, and the number of pixels.

Distortion Level	Value of θ_t (arcmin)	Pixels
1	2.25′	½
2	4.5′	1
3	6.75′	1½
4	9.0′	2
5	11.25′	2½
6	13.5′	3
7	15.75′	3½

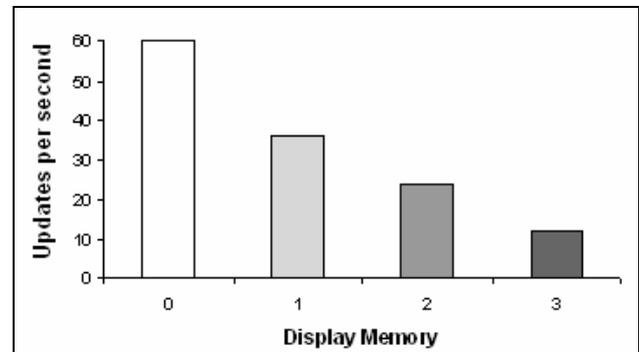**Table 1: Value of θ_t for each signal strength.**

Figure 5: Display memory update rates for the scene using a θ_t value of 4.5′.

By using these θ_t values, a display memory would be forced to update if the image in that display memory was no longer valid to within θ_t from the user's viewpoint. A view of the virtual environment used in the experiment can be seen in figure 10. Region sizes in the virtual environment were 10 square metres. The scenes were rendered using a simulation of region priority rendering. Each scene in the experiment had the camera or viewport translating along a fixed path. This was to ensure that all participants would view the exact same sequence of frames. The worst case scenario was chosen for the fixed path. In other words, the path chosen would give the maximum distortions with the viewport looking in the direction where the distortions would be most apparent. The fixed path included both horizontal and vertical

movements, with the viewport translating through the scene at a constant velocity of 1.5 meters per second (a fast walking speed). Figure 5 illustrates the update rates used for a θ_t value of 4.5′ at this translational speed (refer to figure 2 for locations of regions with respect to the user). The display memories had different update rates for other values of θ_t. Nevertheless, display memory 0 was always kept at 60 updates per second.

Each scene was 3 seconds in duration. The scenes had to be short as two scenes were presented sequentially to the participants per trial. This meant that the participant had to remember what he/she saw of the two scenes before deciding which of the scenes appeared to look better. This was also done in the interest of keeping the experiment short, so that the participants would not lose concentration during the experiment.

No anti-aliasing techniques were used in the rendering of the virtual environment. Anti-aliasing is used in computer graphics in order to smoothen out aliasing artefacts such as jagged edges. Anti-aliasing also has the effect of smoothening out the pixel transitions of moving objects. For this reason, it is expected that region warping distortions will be less obvious if anti-aliasing were to be used. However, this was left for future experiments. Note that hardware anti-aliasing is part of the design of the ARP system.

4.3 Procedure

A total of 16 volunteers (12-male and 4-female) participated in the experiment. Participants were aged between 18-41, and had normal vision or corrected normal vision. All participants were experienced computer users.

Participants were informed as to the purpose of the experiment, which was to determine whether or not they could tell the difference between the different methods of rendering. They were told that even though the same virtual environment was used for all the trials, the two scenes presented in each trial were different, and that for each trial they were to choose which of the scenes looked better. Participants were instructed to guess if they could not tell which of the two scenes was better.

Before the experiment, participants were shown an example of region warping distortions during a test run of the experiment. A different virtual environment was used for the example, so that it would not influence the actual experiment. The purpose of the test run was to familiarize the participants with the experiment, what they were required to do, as well as what they were expected to look for during the experiment. However, even though participants were shown region warping distortions on the virtual environment used for the test run, participants were not instructed as to what criterion to adopt when looking for distortions in the virtual environment employed for the actual experiment and therefore had to judge for themselves which of the two scenes contained the distortions.

Participants were encouraged to take a rest break in between the two parts of the experiment, i.e. when switching from one display device to the next. This was to ensure that the participants would remain attentive throughout the experiment. Some participants also took rest breaks at various occasions in between trials. The duration of the whole experiment was approximately 30-40 minutes per participant. In order to balance the experiment, half the participants used the HMD first then the computer monitor, while the other half used the computer monitor first followed by the HMD. This was done to balance any 'learning effects', i.e. participants would more likely perform better during the second part of the experiment as he/she would be more familiar with the experimental tasks and would know where to focus their attention in the virtual environment.

After the experiment, participants were required to provide some personal computer usage information as well as to answer some questions relating to the experiment by filling in a questionnaire.

5   Results and Discussions

Of the experimental data collected from the 16 participants, only 15 sets were used in the analysis. The experimental results of one of the participants who consistently answered incorrectly even at the highest distortion level were not considered in the analysis. Prior to this experiment, only one of the participants had ever used a HMD.

In the experimental results, a *higher* threshold meant that the participant was *less* likely to be able to correctly perceive a difference in the scenes presented during each trial of the experiment, whereas a *lower* threshold meant that the participant was *more* likely to be able to correctly differentiate between the scenes. From observations of the individual participant responses, it was clear that region warping distortions of level 5, corresponding to 2½ pixel distortions, and above were generally perceptible by the participants.
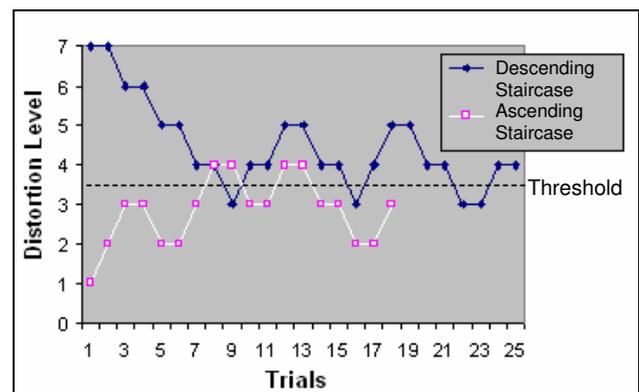
**Figure 6: A plot showing the ascending and descending staircases for one of the participants.**

From the experimental data, graphs of the ascending and descending staircases were plotted for the different display devices and the different region warping methods. This gives a total of 4 graphs for each individual participant. Figure 6 shows an example of this. The mid-run estimation method was used to determine the threshold for the individual participants. Mid-run estimations are the average of the mid-points of staircase

runs. A staircase run is an ascending or descending sequence in the staircase. This means that a staircase reversal would mark the end of a staircase run and the start of another run. The mid-run method of threshold estimation is an efficient way of estimating threshold and the precision of the mid-run estimates has been found to be excellent for small experiments of less than 30 trials (Levitt 1971). In the analysis, the first staircase run for each staircase was ignored in order to give the staircases a chance to stabilize.

Display Device	Warping Method	Mean	Standard Deviation
Monitor	Linear	2.7409	0.84552
	Quadratic	2.6910	1.18052
HMD	Linear	3.1789	1.07441
	Quadratic	3.2867	0.75123

Table 2: Means and standard deviations of participants' thresholds.

Table 2 shows the means and standard deviations of the participants' thresholds for both display devices and both methods of region warping. It can be seen from the table that there is not much difference in average threshold for the different warping methods for each display device respectively. However there is a clear difference in average threshold between the different display devices. The HMD has a higher mean threshold than the computer monitor, which indicates that participants found distortions harder to see on the HMD as compared to the monitor.
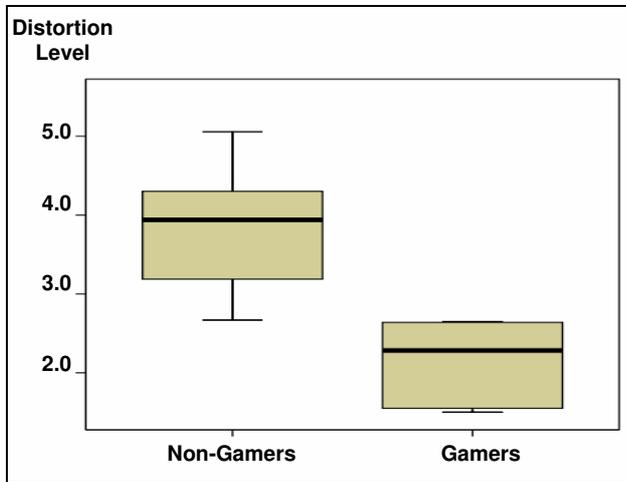
Figure 7: Box plot depicting the difference in threshold between gamers and non-gamers for the case of linear warping viewed using the HMD.

Upon closer examination of the thresholds, it appeared that in general participants who indicated on the questionnaire that they frequently played computer games (this includes console gaming platforms like the X-Box, Play Station, and etc.) had a lower threshold compared to participants who answered to the contrary. A total of 6 participants indicated that they were gamers. Results of a t-test only reported a statistically significant difference in threshold for the HMD and linear warping case [$t(13) = 4.93, p<0.01$], the box plot in figure 7 depicts this difference. Nevertheless, differences in average threshold

levels between gamer and non-gamers can clearly be seen in figure 8. This result is not surprising as gamers are more attuned to interactive graphics.
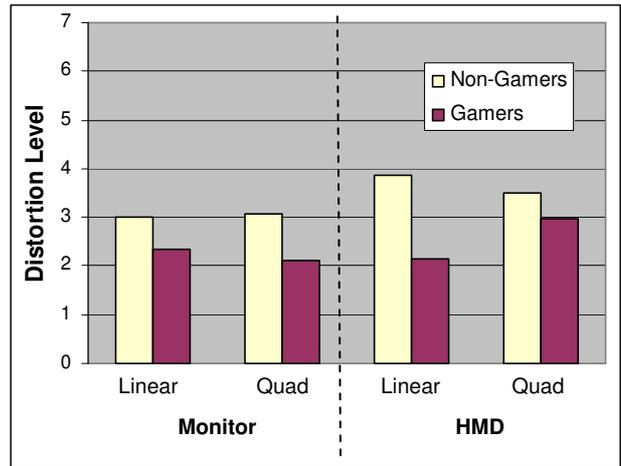
Figure 8: Graph comparing mean threshold between gamers and non-gamers for different display devices and different warping techniques.

In order to further investigate the threshold differences between gamers and non-gamers, group thresholds for the gamers and non-gamers were calculated. Note that mid-point estimations could not be performed for group thresholds, as the staircases could only be plotted for individual participants. Therefore group thresholds were obtained by first calculating the probability of correct responses at each distortion level for the group, and then fitting a cumulative normal ogive to the groups' experimental results. As the cumulative normal ogive is an exponential function, a log transform was first applied to the average probability of correct responses for the two group (to 2 minus the average probability of correct responses). A linear regression was then performed to obtain a fitted psychophysical function. This was performed for all cases involving the different warping methods and the different display devices.
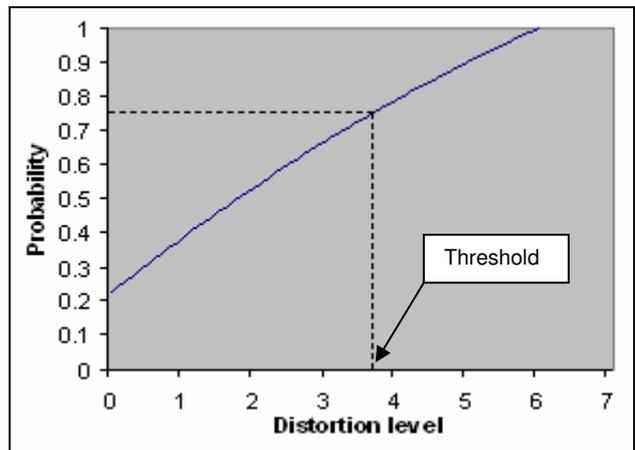
Figure 9: Group thresholds obtained from the 0.75 probability of the psychophysical function.

From fitted psychophysical functions the 0.75 probability was used to obtain the group threshold. For the 2AFC method, a probability of 0.5 signifies random guessing while a probability of 1.0 would mean 100% correct

responses. 0.75 is typically used to obtain the threshold in 2AFC experiments (Ulrich and Miller 2004). Figure 9 illustrates the 75% threshold obtained from a fitted psychophysical function. Table 3 gives the results of the group thresholds.

Gamer	Monitor		HMD	
	Linear	*Quad.*	*Linear*	*Quad.*
	3.0558	2.3341	2.4372	3.6815
Non-Gamer	Monitor		HMD	
	Linear	*Quad.*	*Linear*	*Quad.*
	3.6886	3.8115	4.4975	4.3570

Table 3: Group thresholds between the gamers and non-gamers for the different cases.

From the table it can be seen that there is a clear difference in threshold between the gamers and non-gamers. This agrees with observations of the average individual thresholds between gamers and non-gamers presented earlier. For the non-gamers, the difference in threshold between the display devices also indicates that distortions on the HMD are not as perceivable as on the computer monitor, and that there is no apparent difference between warping methods. This also agrees with analysis presented earlier. However for the participants who were gamers, this was not the case. The results suggest that the gamers perceived less distortion in linear region warping on the computer monitor. But on the other hand, the gamers perceived less distortion in quadratic region warping on the HMD.

At this stage it is not clear as to why the difference in warping technique on the different display devices for the gamers is so apparent. Green and Bavelier (2003) have also found that habitual gamers have a perceptually modified visual attention compared to non-gamers. One possible explanation might be due to the size of the display area of the display device. Plumert et al. (2004) have observed that user perception of distance judgment in virtual environments is better on large screen immersive displays rather than virtual reality systems involving HMDs. Results of another study by Willemsen et al. (2004) that experimented with the use of HMDs and 'mock HMDs' [replica shell of a HMD that restricts a person's Field-of-View (FOV)], also implies that there is an apparent compression of perceived virtual spaces when using HMDs. Wu et al. (2004) have also previously conducted experiments with regards to human distance perception, by using a pair of opaque goggles with a monocular rectangular aperture to limit the horizontal or vertical FOVs of human subjects, and have reported similar findings.

These conclusions by other researchers also suggest that a user's sense of perception is affected by the use of HMDs. This is possibly why the gamers perceived warping distortions differently on the HMD as opposed to the monitor, and also why the non-gamers had higher thresholds on the HMD as compared with the computer monitor. Another reason might be because when set to a resolution of 640×480 in full screen, the area covered by 1 pixel will appear much larger on a 17 inch monitor

compared to the HMD. In order to further investigate this, we are currently setting up visual perception experiments using two different HMDs with exactly the same resolution but with different FOVs, in order to affirm whether a HMD's FOV makes a difference to perception of region warping distortions.

In the questionnaire, participants were asked to circle a section or sections of the scene where they focused their attention the most. It was found that most participants focused their attention around one of two major sections in the virtual environment. Some participants circled both sections. These two sections are shown in figure 10 below and close ups of the two sections are provided in figures 11 and 12 respectively.

Figure 10: Sections of the virtual environment where participants focused their attention the most.

Figure 11: Close up of trees section.

Figure 12: Close up of mountain-sand section.

Chalmers et al. (2003) explains this focus of attention in virtual environments as being 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. Therefore 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 virtual environment. This gives rise to the possibility of concentrating distortions away from where the user might focus his/her attention or to place eye catching objects in the scene in order to draw the user's attention away from the region warping distortions.

Participants were also asked to choose whether the distortions were more noticeable during horizontal movement, vertical movement or whether direction of movement did not make a difference. It was interesting to note that participants who circled the trees section, typically indicated that distortions were more apparent during horizontal movement, whereas participants who circled the mountain-sand section typically selected vertical movement. When asked to clarify their answers, participants who circled the trees section said that they concentrated on the jerkiness in the movement of the trees during horizontal movement, and participants who circled the mountain-sand section said that they focused on the transition rate of the light and dark pixels between the mountain and the sand.

6   Conclusions and Future Work

The results of this experiment suggest that in general region warping distortions are less perceptible on a HMD compared to on a computer monitor. These findings concur with various other studies conducted by other researchers that conclude that the use of a HMD does affect a user's sense of perception. Distortions of over 2 pixels were generally perceptible by all the participants, whereas distortions below 2 pixels were less perceptible and distortions of 1 pixel and below were very much less perceptible.

Analysis of the experimental data also showed that there was a difference in visual perception between participants who frequently played computer/electronic games and participants who were non-gamers. It was also observed that a user will normally focus his/her attention on certain sections of a virtual environment and pay less attention to the remainder of the scene.

This work is relevant in a wide variety of fields in computer graphics research and applications, ranging from the design of interactive animation hardware like the Talisman to various perceptually based rendering techniques as well as other perception studies involving graphics and HMDs.

This experimentation is designed to enable us to concentrate our computing power on processes that improve the user experience. To that end we are now setting up visual perception experiments involving HMDs with different FOVs in order to further investigate human perception of region warping distortions. Considerations in planning for future experiments include employing anti-aliasing in the rendering of the virtual environment, as well as the use of different virtual environments with various scene characteristics and complexities. The impact of stereoscopy, different image resolutions and user-initiated interaction also has to be investigated. A discussion of a number of other factors that have to be considered when designing human visual perception experiments have previously been outlined in Chow et al. (2005c).

8   References

Barkans, A.C. (1997): High Quality Rendering using the Talisman Architecture. *Proc. of the ACM SIGGRAPH/EUROGRAPHICS Workshop on Graphics Hardware*, Los Angeles, California, 79-88.

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

Chow, Y.W., Pose, R. and Regan, M. (2005a): Large Object Segmentation with Region Priority Rendering. *Proc. of the 28th Australasian Computer Science Conference 2005 (ACSC 2005)*, Newcastle, NSW, Australia, 19-28.

Chow, Y.W., Pose, R. and Regan, M. (2005b): Region Warping in a Virtual Reality System with Priority Rendering. *Proc. of the 2nd IADIS International Conference on Applied Computing*, Algarve, Portugal, 451-458.

Chow, Y.W., Pose, R. and Regan, M. (2005c): Design Issues in Human Visual Perception Experiments on Region Warping. *Proc. of the 2nd IADIS International Conference on Applied Computing*, Algarve, Portugal, 210-217.

Ellis, S.R., Mania, K., Adelstein, B.D. and Hill, M. (2004): Generalizeability of Latency Detection in a variety of Virtual Environments. *Proc. Of the Human Factors and Ergonomics Society 48th Annual Meeting*. New Orleans, Louisiana, 2632-.

Farell, B. and Pelli, D.G. (1999): Psychophysical Methods, or how to Measure a Threshold, and why. In

Vision Research: A Practical Guide to Laboratory Methods. 129-136. CARPENTER, R.H.S. and ROBSON, J.G. (eds). Oxford, Oxford University Press.

Gescheider, G.A. (1985): *Psychophysics: Method, Theory and Application, 2nd Edition*. Hillsdale, New Jersey, Lawrence Erlbaum Associates.

Green, C.S. and Bavelier, D. (2003): Action Video Game Modifies Visual Selective Attention. *Nature*, **423**(29 May):534-537.

Levitt, H. (1971): Transformed Up-Down Method, *Journal of the Acoustical Society of America*, **49**(2): 467-477.

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. *ACM International Conference Proceeding Series, Proc. of the 1st Symposium on Applied Perception in Graphics and Visualization*. Los Angeles, California, **73**:39-47.

Mark, W.R., McMillan, L., and Bishop, G. (1997): Post-Rendering 3D Warping. *Proc. of the 1997 Symposium on Interactive 3D Graphics*, Providence, Rhode Island, 7-16.

Meehan, M., Razzaque, S., Whitton, M.C. and Brooks, F.P. (2003): Effect of Latency on Presence in Stressful Virtual Environments. *Proc. of IEEE Virtual Reality 2003*, Los Angeles, California, 141-148.

NVIDIA (2005): GPU Programming Guide version 2.4.0, NVIDIA Corporation, http://download.nvidia.com/developer/GPU_Programming_Guide/GPU_Programming_Guide.pdf. Accessed 5 August 2005.

Paul Bourke's Personal Pages: Texture Library. http://astronomy.swin.edu.au/~pbourke/texture/ Accessed 1 August 2005.

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

Pose, R. and Regan, M. (1994): Techniques for Reducing Latency with Architectural Support. *Proc. of East-West International Conference on Multimedia, Hypermedia and Virtual Reality*, Moscow, Russia, 153-160.

Reddy, M. (1997): Perceptually Modulated Level of Detail for Virtual Environments. Ph.D. thesis. Dept. of Computer Science, University of Edinburgh, UK.

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

Regan, M. and Pose, R. (1994): Priority Rendering with a Virtual Reality Address Recalculation Pipeline. *Proc. ACM SIGGRAPH '94, in Computer Graphics, Annual Conference Series*, Orlando, Florida, 155-162.

Regan, M.J.P., Miller, G.S.P., Rubin, S.M., and Kogelnik, C. (1999): A Real-Time Low-Latency Hardware Light-Field Renderer. *ACM SIGGRAPH '99, Proc. of the 26th Annual Conference on Computer Graphics,* Los Angeles, California, 287-290.

Torburg, J. and Kajiya, J.T. (1996): Talisman: Commodity Realtime 3D Graphics for the PC. *ACM SIGGRAPH '96, in Computer Graphics, Annual Conference Series*, New Orleans, Louisiana, 353-363.

Ulrich, R. and Miller, J. (2004): Threshold Estimation in Two-Alternative Forced-Choice Tasks: The Spearman-Karber Method. *Perception and Psychophysics*. **66**(3):517-533.

Watson, B.A., and Hodges, L.F. (1995): Using Texture Maps to Correct for Optical Distortion in Head-Mounted Displays. *Proc. of the IEEE Virtual Reality Annual Symposium (VRAIS '95)*, Research Triangle Park, 172-178.

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*. **4**(4):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. *Proc. of the 1st Symposium on Applied Perception in Graphics and Visualization (APGV)*, Los Angeles, California, 35-58.

Wu, B., Ooi, T.L. and He, Z.J. (2004): Perceiving distance accurately by a directional process of integrating ground information. *Nature*, **428**(4 March):73-77.