Selective Quality Rendering by Exploiting Human Inattentional Blindness: Looking but not Seeing

Kirsten Cater
Alan Chalmers and Patrick Ledda
University of Bristol
Merchant Venturers, Building Woodland Road,
Bristol, BS8 1UB, UK
Tel: +44(0)1179545112
cater@cs.bris.ac.uk

ABSTRACT

There are two major influences on human visual attention: bottom-up and top-down processing. Bottom-up processing is the automatic direction of gaze to lively or colourful objects as determined by low-level vision. In contrast, top-down processing is consciously directed attention in the pursuit of predetermined goals or tasks. Previous work in perception-based rendering has exploited bottom-up visual attention to control detail (and therefore time) spent on rendering parts of a scene. In this paper, we demonstrate the principle of Inattentional Blindness, a major side effect of top-down processing, where portions of the scene unrelated to the specific task go unnoticed. In our experiment, we showed a pair of animations rendered at different quality levels to 160 subjects, and then asked if they noticed a change. We instructed half the subjects to simply watch our animation, while the other half performed a specific task during the animation.

When parts of the scene, outside the focus of this task, were rendered at lower quality, almost none of the task-directed subjects noticed, whereas the difference was clearly visible to the control group. Our results clearly show that top-down visual processing can be exploited to reduce rendering times substantially without compromising perceived visual quality in interactive tasks.

Categories and Subject Descriptors

I.3.3 [**Computer Graphics**]: Picture/Image Generation - Viewing Algorithms; I.3.7 [**Computer Graphics**]: Three-Dimensional Graphics and Realism - Animation; I.4.8 [**Image Processing and Computer Vision**]: Scene Analysis -Time-varying imagery.

General Terms

Experimentation, Human Factors and Theory.

Keywords

Inattentional Blindness, Human Visual Perception, Task Related Realistic Rendering, Interactive Rendering of Dynamic Scenes and Image Quality.

1. INTRODUCTION

One major goal of virtual reality and computer graphics is achieving real-time frame rates at the highest perceived rendering quality. Improvements in basic rendering hardware and algorithms have produced some remarkable results, but we are still unable to render highly realistic imagery in anything close to real-time. From the early days of flight simulation, researchers have studied what parts of a scene or image are most likely to be noticed in an inter-active setting. Most of this research has attempted to exploit gaps in low-level visual processing, similar to JPEG and other image compression schemes [1].

In practice, the perception of a virtual environment depends on the user and the task that he/she is currently performing in that environment. Another way to produce perceptually high-quality images in reasonable times is to exploit existing limitations in the actual human visual system.

Figure 1: An example of a top-down visual attention process, counting pencils in a mug. *This figure is reproduced in colour on page 211.*

Visual attention is the process by which humans select a portion of the available visual information for localisation, identification and understanding of objects in the environment. It allows our visual system to process visual input preferentially by shifting attention about an image, giving more attention to salient locations and less attention to unimportant regions. When attention is not focused onto items in a scene they can literally go unnoticed. Therefore, ***Inattentional Blindness is the failure of the human to see unattended items in a scene*** [2]. The scan path of the eye is thus very strongly affected by this visual attention.

There are two general processes, called *bottom-up* and *top-down*, which determine where humans locate their visual attention [3]. The bottom-up process is purely stimulus driven, for example, a candle burning in a dark room, a red ball amongst a large number of blue balls, or the lips and eyes of a human face as they are the most mobile and expressive elements of the face. In all these cases, the visual stimulus captures attention automatically without volitional control. The top-down process, on the other hand, is directed by a voluntary control process that focuses attention on one or more objects that are relevant to the observer's goal when studying the scene. Such goals may include looking for street signs, searching for a target in a computer game, or counting the number of pencils in a mug, shown in Figure 1. In this case, the conspicuous objects in a scene that would normally attract the viewer's attention may be deliberately ignored if they are irrelevant to the task at hand. This is called *Inattentional Blindness*. It is precisely this top-down processing of the human visual system while performing a task that we may exploit in our virtual environments.

After reviewing visual attention models and previous research we will describe our experiment, where certain subjects were asked to perform a simple visual task of counting pencils in a sequence of virtual mugs as they move by, whilst other subjects were asked to simply watch the animations. We discuss the different responses between these two groups in the results section and conclude with some ideas for future research.

2. VISUAL ATTENTION

Visual scenes typically contain many more objects than can ever be recognized or remembered in a single glance. Some kind of sequential selection of objects for detailed processing is essential if humans are to cope with this wealth of information. Coded into the primary levels of human vision is a powerful means of accomplishing this selection, namely the retina [4]. Spatial acuity is highest right at the centre of the retina, the fovea, and then falls off rapidly toward the periphery. The receptors in the fovea, consisting exclusively of colour-sensitive cones, are much more densely packed than those in the periphery. As a result, spatial and chromatic resolutions are much higher in the small central fovea than elsewhere.

However, the visual angle covered by the fovea is only approximately 2 degrees, about the size of eight letters on a typical page of text or the size of your thumbnail held at arm's length. If detailed information is needed from many different areas of the visual environment, it can only be obtained by redirecting the eye so that the relevant objects fall sequentially on the fovea.

The study of saccadic exploration of complex images was pioneered by the Russian psychologist Yarbus [5]. By using relatively crude equipment he was able to record the fixations and saccades observers made while viewing natural objects and scenes. Yarbus demonstrated that humans do not scan a scene in a raster-like fashion, but rather the eyes jump to foveate a new point of interest in the scene. These are known as saccades. Specifically he studied saccadic records made by observers studying an image after they were given a particular task. Each observer then views the scene with that particular question or task in mind. This seeking information of a specific kind has a significant affect on the eye-gaze pattern. To illustrate this Yarbus instructed several observers to answer a number of different questions concerning the depicted situation in Repin's picture 'An Unexpected Visitor' [5]. This resulted in substantially different patterns, each one once again being easily construable as a sampling of those picture objects that were most informative for the answering of the question, as shown in Figure 2.

Figure 2a: Repin's picture 'An Unexpected Visitor'. *This figure is reproduced in colour on page 211.*

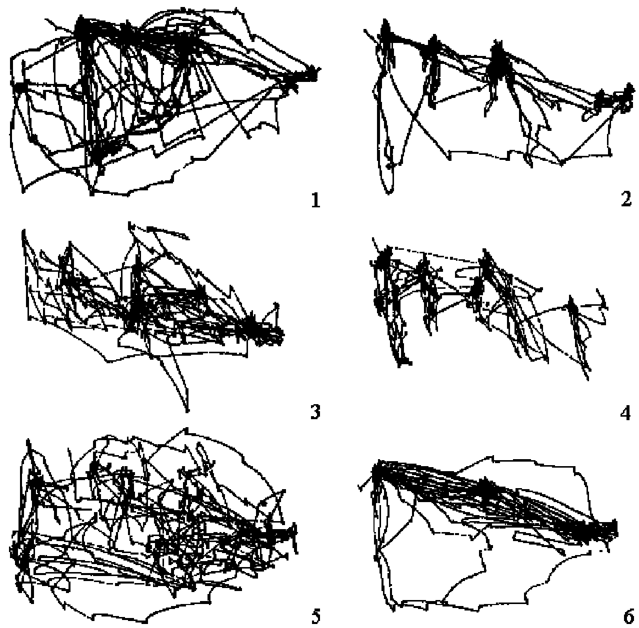


Figure 2b: Effects of task on eye movements. Repin's picture was examined by subjects with different instructions; 1. Free viewing, 2. Judge their ages, 3. Guess what they had been doing before the unexpected visitor's arrival, 4. Remember the clothes worn by the people, 5. Remember the position of the people and objects in the room & 6. Estimate how long the unexpected visitor had been away from the family [5].

What Yarbus, and subsequently many others, showed was that, while performing a task, once an initial eye saccade has found the appropriate object to locate it on the fovea, the eye subsequently performs a ***smooth pursuit movement*** to keep the object in foveal vision6. Because the image of a successfully tracked object is nearly stationary on the retina, pursuit movements enable the visual system to extract maximum spatial information from the image of the moving object itself. Untracked objects, both stationary and moving objects with different velocities and directions to the target object, are experienced as smeared and unclear because of their motion on the retina. To experience this Palmer [6] suggests a simple example: *place your finger on this page and move it fairly quickly from one side of the page to another. As soon as you track your moving finger, the letters and words appear so blurred you are unable to read them, but your finger is clear.* Even when you stop moving your finger, only the words located within the visual angle of your fovea become sharp and thus readable.

2.1 Related work

The pioneering research that has been performed in the application of visual perception to computer graphics has exploited the bottom-up visual attention process. This work has included using knowledge of the human visual system to improve the quality of the displayed image, for example [7,8,9,10,11,12]. Other research has investigated how complex detail in the models can be reduced without any reduction in the viewer's perception of the models, for example [13,14,15] and, Maciel and Shirley's visual navigation system that uses texture mapped primitives to represent clusters of objects to maintain high and approximately constant frame rates [16]. In addition, saliency models have been developed to simulate where people focus their attention in images and peripheral vision studies have been carried out to control level of detail in display devices. Of this previous research, peripheral vision studies and saliency models are the closest to our current work.

2.2 Peripheral vision

As with our approach, this research builds on previous vision work that indicates the human eye only processes detailed information from a relatively small part of the visual field. Watson et al [17] proposed a paradigm for the design of systems that manage level of detail in virtual environments; they performed a user study to evaluate the effectiveness of high detail insets used in head-mounted displays. Subjects performed a search task with different display types, each of these display types was a combination of two independent variables: peripheral resolution and the size of the high level of detail inset. The high detail inset they used was rectangular and was always presented at the fine level of resolution. The level of peripheral resolution was varied at three possible levels: fine resolution 320x240, medium resolution 192x144 and coarse resolution 64x48. There were three inset sizes; the large inset size was half the complete display's height and width, the small inset size was 30% of the complete display's height and width, the final size was no inset at all. Their results showed observers found their search targets faster and more accurately for the high resolution no inset, however it was not significantly better than the high-resolution inset displays with either medium or coarse peripheral resolutions. Thus peripheral level of detail degradation can be a useful compromise to achieve desired frame rates. Watson et al. are now working on measuring and predicting visual fidelity for simplifying polygonal models [18], McConkie and Loschky [19], Loschky and McConkie [20] and Loschky et al [21] had observers examining complex scenes with an eye-linked multiple resolution display, which produces high visual resolution only in the region to which the eyes are directed. Image resolution and details outside this 'window' of high resolution are decreased. This significantly reduces bandwidth as many interactive single-user image display applications have prohibitively large bandwidth requirements. Their recent study measured viewers' image quality judgements and their eye movement parameters, and found that photographic images filtered with a window radius of 4.1 degrees produced results statistically indistinguishable from that of a full high-resolution display.

This approach does, however, encounter the problem of keeping up with updating the multi-resolutional display after an eye movement without disturbing the visual processing. The work has shown that the image needs to be updated after an eye saccade within 5 milliseconds of a fixation otherwise the observer will detect the low resolution.

These high update rates were only achievable by using an extremely high temporal resolution eye tracker and by pre-storing all possible multi-resolutional images that were to be used.

2.3 Saliency models

Saliency models determine what is visually important within the whole scene. Yee [22] and Yee et al [23] presented a method to accelerate global illumination computation in pre-rendered animations by using a model of visual attention to locate regions of interest in a scene and to modulate spatiotemporal sensitivity. They create a spatiotemporal error tolerance map24, constructed from data based on velocity dependent contrast sensitivity, and a saliency map25 for each frame in the animation. The saliency map is obtained by combining the conspicuity maps of intensity, colour, orientation and motion. This then creates an image where bright areas denote greater saliency, i.e. where attention is more likely to be drawn. An Aleph map is then created by combining the spatiotemporal error tolerance map with the saliency map. The resulting Aleph map is then used as a guide to indicate where less rendering effort should be spent in computing the lighting solution and thus significantly reduce the overall computational time to produce their animations.

3. OUR APPROACH

The methodology we describe here is similar to that proposed by Yee [22] and McConkie and Loschky [19], with the crucial difference that we directly exploit the *top-down visual attention process* rather than the bottom-up process.

3.1 Overview

As Yarbus showed [5], the choice of task is important in helping us predict the eye-gaze pattern of the viewer. It is precisely this knowledge of the expected eye-gaze pattern that will allow us to reduce the rendered quality of objects outside the area of interest without affecting the viewer's overall perception of the quality of the rendering.

The task chosen for each user was to count the number of pencils that appeared in a mug on a table in a room as they moved on a fixed path through four such rooms. To count the pencils the user needs to perform a smooth pursuit eye movement tracking the mug in one room until they have successfully counted the number

Figure 3: Close up of the mug showing the pencils and paintbrushes. *This figure is reproduced in colour on page 211.*

of pencils in that mug and then perform an eye saccade to the mug in the next room. To further complicate the task and thus retain the viewer's attention each mug also contained a number of spurious paintbrushes, shown in Figure 3.

3.2 The psychophysical experiments

The study involved three rendered animations of an identical fly through of four rooms, the only difference being the quality to which the individual animations had been rendered.

Figure 4 (a) shows the high quality rendered scene, while (b) shows the same scene rendered at a significantly lower quality, with a much reduced computational time.

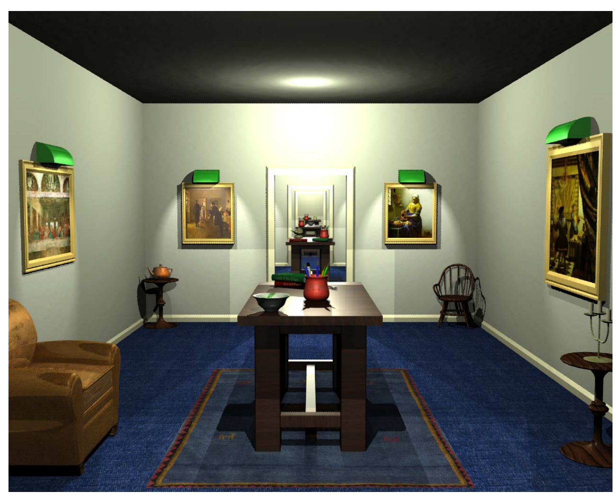
Each frame for the high quality animation took on average 18 minutes 53 seconds to render in Alias Wavefront Maya on an Intel Pentium 4 1GHz Processor, while the frames for the low quality animation were each rendered on average in only 3 minutes 21 seconds.

Low quality (LQ) in this experiment was classified by rendering the entire animation at custom low quality with no anti-aliasing, no motion-blur and one ambient bounce (the maximum number of times a light ray can be reflected). In comparison, high quality (HQ) was classified by rendering the entire animation at highest quality anti-aliasing, which computes each frame in two passes, looking for colour contrasts within pixels and in surrounding pixels, such as highlights. The first pass is the high quality computation the second pass looks for colour contrast in the results of the first pass. In the regions where colour contrast is high, for example in regions containing highlights, more shading samples are taken. The ambient bounce value was set to six. The circle quality (CQ) animation was created by using the low quality frames with high quality rendering substituted in the visual angle of the fovea (2 degrees) centered around the pencils, shown by the green circle in Figure 5. The high quality is blended to the low quality at 4.1 degrees visual angle (the red circle in Figure 5) McConkie et al [19].

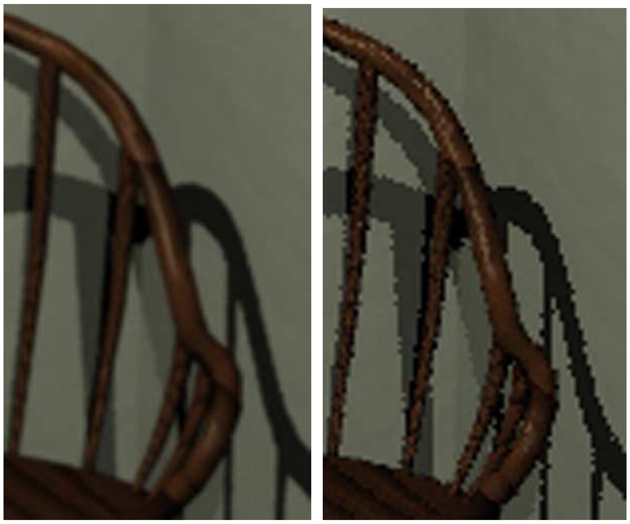
In the final experiment a total of 160 subjects were studied. Each subject saw two animations of 35 seconds, displayed at 15 frames

Figure 4: (a) High Quality (HQ) image (Frame 26 in the animation). *This figure is reproduced in colour on page 211.*

(b) Low Quality (LQ) image (Frame 26 in the animation). *This figure is reproduced in colour on page 211.*



(c) Selectively rendered (CQ) image with two Circles of high Quality over the first and second mugs (Frame 26 in the animation). *This figure is reproduced in colour on page 211.*



(d) Close up of High Quality rendered chair and the Low Quality version of the same chair. *This figure is reproduced in colour on page 212.*

Figure 5: Visual angle covered by the fovea for mugs in the first two rooms at 2 degrees (green circles) and 4.1 degrees (red circles). *This figure is reproduced in colour on page 212.*

per second. Table 1 describes the conditions tested with 32 subjects per condition. Fifty percent of the subjects were asked to count the pencils in the mug while the remaining fifty percent were simply asked to watch the animations. To minimise experimental bias the choice of condition to be run was randomised and for each, 16 were run in the morning and 16 in the afternoon. Subjects had a variety of experience with computer graphics and all exhibited at least average corrected vision in testing.

Before beginning the experiment the subjects read a sheet of instructions on the procedure of the particular task they were to perform. After the participant had read the instructions they were asked to clarify that they had understood the task. They then rested their head on a chin rest that was located 60cm away from a 17-inch monitor. The chin rest was located so that their eye level was approximately level with the centre of the screen. The animations were displayed at a resolution of 1280 x 1024.

Table 1: Conditions Tested. The two animations shown for the experiments were thus: (1) HQ+HQ, (2) HQ+LQ, (3) LQ+HQ, (4) HQ+CQ, and (5) CQ+HQ

Acronym	Description
HQ	**High Quality:** Entire animation rendered at the highest quality.
LQ	**Low Quality:** Entire animation rendered at a low quality with no anti-aliasing.
CQ	**Circle Quality:** Low Quality Picture with high quality rendering in the visual angle of the fovea (2 degrees) centered around the pencils, shown by the green circle in figure 5. The high quality is blended to the low quality at 4.1 degrees visual angle (the red circle in Figure 5) McConkie et al [19].

To ensure that the viewers focused their attention immediately on the first mug and thus did not have to look around the scene to find it, a countdown was shown to prepare them that the animation was about to start followed immediately by a black image with a white mug giving the location of the first mug. They were shown the second animation immediately afterwards.

On completion of the experiment each participant was asked to fill in a detailed questionnaire. This questionnaire asked for some personal details including age, occupation, sex and level of computer graphics knowledge. The participants were then asked detailed questions about the objects in the rooms, their colour, location and quality of rendering. These objects were selected so that questions were asked about objects both near the foveal visual angle (located about the mug with pencils) and in the periphery. They were specifically asked not to guess, but rather state 'don't remember' when they had failed to notice some details.

3.3 Results

Figure 6 shows the overall results of the experiment. Obviously the participants did not notice any difference in the rendering quality between the two HQ animations (they were the same). Of interest is the fact that, apart from one case in the CQ+HQ experiment, the viewers performing the task consistently failed to notice any difference between the high quality rendered animation and the low quality animation where the area around the mug was rendered to a high quality. Surprisingly 25% of the viewers in the HQ+LQ condition and 18% in the LQ+HQ case were so engaged in the task that they completely failed to notice any difference in the quality between these very different qualities of animation.

Figures 7 (a) and (b) show that having performed the task of counting the pencils, the vast majority of participants were simply unable to recall the correct colour of the mug (90%) which was in the foveal angle and even less the correct colour of the carpet (95%) which was outside this angle. The Inattentional Blindness was even higher for less obvious objects, especially those outside the foveal angle.

Overall the participants who simply watched the animations were able to recall far more detail of the scenes, although the generic

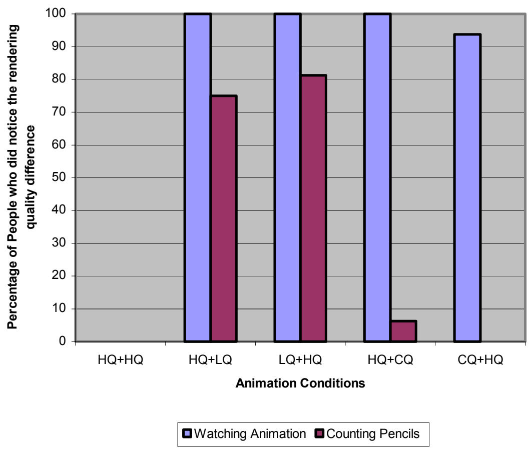


Figure 6: Experimental results for the two tasks: Counting the pencils and simply watching the animations.

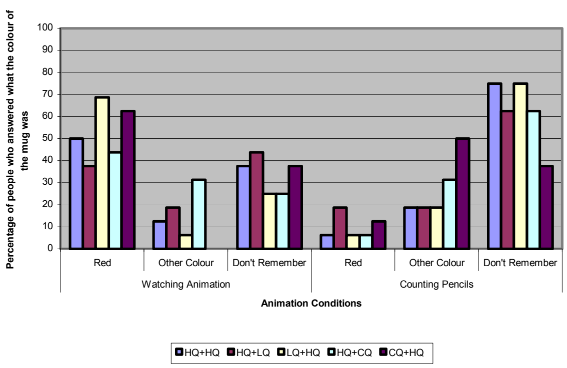


Figure 7(a): How observant were the participants: Colour of the mug.

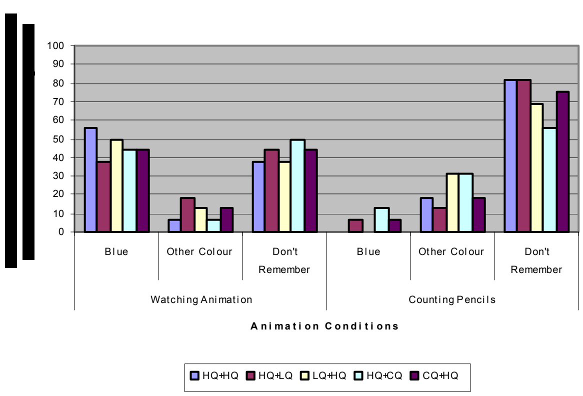


Figure 7(b): How observant were the participants: Colour of the carpet.

nature of the task given to them precluded a number from recalling such details as the colour of specific objects, for example 47.5% could not recall the correct colour of the mug and 53.8% the correct colour of the carpet.

3.4 Analysis

Since the response of the observers was binary. The appropriate method of statistical analysis was the Chi-square test ($X2$) for significance. Standard linear regression models or analysis of variance (ANOVA) are only valid on continuous data from normal distributions and therefore were not appropriate. By performing pair-wise comparisons of all the other animations to the animations HQ+HQ we could determine whether the results were statistically significant.

When simply watching the animation, the test statistics for all the pair-wise comparisons were statistically significant. The result for the pair-wise comparison of HQ+HQ and HQ+CQ was $X2 = 28.125$, df = 1, $p < 0.005$ (a p value of 0.05 or less denotes a statistically significant result). In comparison when counting the pencils the test statistics were significant for the pair-wise comparisons HQ+HQ with HQ+LQ ($X2 = 32$, df = 1, $p < 0.005$). However, for the comparisons HQ+HQ with CQ+HQ the results were statistically insignificant, and thus the null hypothesis was retained ($X2 = 0.125$, df = 1, $p > 0.1$). From this we can conclude that when the observers were counting the pencils the HQ+HQ animations and the HQ+CQ animations produced the same result, i.e. the observers thought that they were seeing the same animation twice, with no alteration in rendering quality.

3.5 Verification with an eye-tracker

To make certain that the attention of the observer was being captured by the task, counting pencils, the experiment was briefly repeated with the Eyelink Eyetracking System developed by SR Research Ltd. and manufactured by SensoMotoric Instruments. Figure 8 shows an example of the scan path of one of the observers whilst performing the counting pencils task for 2 seconds. While all the observers had slightly different scan paths none of their eye scans left the green box (Figure 8). Figure 9 shows an example of the scan path of one of the observers who was simply watching the animation for 2 seconds.

These results demonstrate that Inattentional Blindness may in fact be exploited to significantly reduce the rendered quality of a large portion of the scene without having ***any*** affect on the viewer's perception of the scene.

Figure 8: An eye scan for an observer counting the pencils. The green crosses are fixation points and the red lines are the saccades. *This figure is reproduced in colour on page 212.*

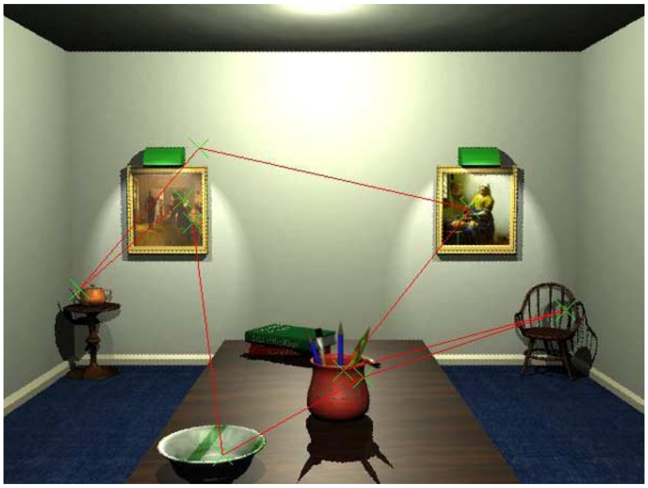


Figure 9: An eye scan for an observer who was simply watching the animation. *This figure is reproduced in colour on page 212.*

Figure 10: A two second scan path for an observer counting teapots. *This figure is reproduced in colour on page 212.*

4. CONCLUSIONS AND FUTURE WORK

For virtual reality applications in which the task is known *a priori* the computational savings made by exploiting Inattentional Blindness can be dramatic.

Our approach works by identifying the area of user fixation determined by the task being performed, rendering this to a high quality and exploiting the Inattentional Blindness inherent in the Human Visual System to render the rest of the scene at a significantly lower quality. Our results show that while engaged in the task, users consistently failed to notice the quality difference, and even objects, within the scene.

Although we are primarily interested in high-quality image synthesis, this technique may also be applied to other areas of computer graphics, for example geometry level-of-detail selection, video telephony and video compression.

There are many aspects of this methodology which require further research:

Visual attention. It is already known from visual psychological researchers such as Yarbus [5], Itti and Koch [25] and Yantis [26] that the visual system is highly sensitive to features such as edges, abrupt changes in colour and sudden movements. A further understanding of the complex interaction between the bottom-up and top-down visual attention processes of the human visual system will enable us to ***combine*** the important saliency models of Yee with our Inattentional Blindness approach to determine more precisely the order in which people may attend to objects in a scene. Such knowledge can provide a detailed priority queue for selective rendering, providing the best perceptibly high-quality images within the time constraints of the interactive display system. Such a priority rendering queue also offers exciting possibilities for efficient task scheduling within any parallel implementation of our methodology.

Selective rendering. The animations for our experiments were pre-computed at the low and high qualities using Alias Wavefront Maya. In future this rendering will be interactive using a customised raytracer, with the choice of high or low quality rendering, for example the number of directions sampled per intersection etc, being determined dynamically from the rendering priority queue.

Peripheral vision. Foveal information is clear and fully chromatic, whereas peripheral information is blurry and colour-weak to a degree that depends on the distance from the fovea. Thus we would like to do more work on decreasing the amount of time spent rendering good colour in the periphery.

Type of task. The task undertaken is crucial in determining the eye-gaze patterns of users studying the images. Future work will involve verifying whether or not the type of task effects the observer's perception of the selective rendering quality. Figure 10 shows the scan path of an observer performing a pilot study to answer this question. The observers were asked to count the number of teapots in a series of images, i.e. scanning the whole of each image, and then comment on whether they had seen any rendering change on each image.

Multisensory experiences. The introduction of sound and motion within the virtual environments may further increase the level of Inattentional Blindness and the related Change Blindness [27, 28].

ACKNOWLEDGEMENTS

This research is funded by the Engineering and Physical Sciences Research Council (Award No: 00301786). We are indebted to Greg Ward for all his advice, and to Peter Longhurst and James Edwards for their help in setting up the experiments. We would also like thank everyone who attended the experiment; we would not have got any results if it were not for them.

REFERENCES

1. BOLIN MR. AND MEYER GW. 1998 A Perceptually Based Adaptive Sampling Algorithm, In *Proceedings of SIGGRAPH 1998*, ACM Press / ACM SIGGRAPH, New York. M. Cohen, Ed., Computer Graphics Proceedings, Annual Conference Series, ACM, 299-309
2. MACK A. AND ROCK I. 1998 *Inattentional Blindness*, Massachusetts Institute of Technology Press.
3. JAMES W. 1890 *Principles of Psychology*, New York: Holt.
4. DOWLING JE. 1987 The retina: An approachable part of the brain, Cambridge: Belknap.
5. YARBUS AL.1967 Eye movements during perception of complex objects. In L. A. Riggs, Ed., *Eye Movements and Vision*, Plenum Press, New York, Chapter VII, 171-196.

6. PALMER SE. 1999 Vision Science- Photons to Phenomenology. Massachusetts Institute of Technology Press.

7. FERWERDA J., PATTANAIK SN, SHIRLEY P. AND GREENBERG DP. 1996 A Model of Visual Adaptation for Realistic Image Synthesis, In *Proceedings of SIGGRAPH 1996*, ACM Press / ACM SIGGRAPH, New York. H. Rushmeier, Ed., Computer Graphics Proceedings, Annual Conference Series, ACM, 249-258

8. GREENBERG DP, TORRANCE KE, SHIRLEY P., ARVO J., FERWERDA J., PATTANAIK SN. LAFORTUNE E., WALTER B., FOO S-C. AND TRUMBORE B. 1997 A Framework for Realistic Image Synthesis, In *Proceedings of SIGGRAPH 1997* (Special Session), ACM Press / ACM SIGGRAPH, New York. T. Whitted, Ed., Computer Graphics Proceedings, Annual Conference Series, ACM, 477-494.

9. MCNAMARA A., CHALMERS AG, TROSCIANKO T., AND GILCHRIST I. 2000 Comparing Real and Synthetic Scenes using Human Judgements of Lightness. In B Peroche and H Rushmeier (eds), *12th Eurographics Workshop on Rendering,* 207-219.

10. MYSZKOWSKI K., TAWARA T., AKAMINE H. AND SEIDEL H-P. 2001 Perception-Guided Global Illumination Solution for Animation Rendering. In *Proceedings of SIGGRAPH 2001*, ACM Press / ACM SIGGRAPH, New York. E. Fiume, Ed., Computer Graphics Proceedings, Annual ConferenceSeries, ACM, 221-230.

11. PATTANAIK SN, FERWERDA J, FAIRCHILD MD, AND GREENBERG DP. 1998 A Multiscale Model of Adaptation and Spatial Vision for Realistic Image Display, in Proceedings *of SIGGRAPH 1998*, ACM Press / ACM SIGGRAPH, New York. M. Cohen, Ed., Computer Graphics Proceedings, Annual Conference Series, ACM, 287-298.

12. RAMASUBRAMANIAN M., PATTANAIK SN, GREENBERG DP. 1999 A Perceptually Based Physical Error Metric for Realistic Image Synthesis, In *Proceedings of SIGGRAPH 1999*, ACM Press / ACM SIGGRAPH, New York. Computer A. Rockwood, Ed., Graphics Proceedings, Annual Conference Series, ACM, 73-82.

13. LUEBKE D. AND HALLEN B. 2001 Perceptually driven simplification for interactive rendering, In *Proceedings of 12th Eurographics Workshop on Rendering*, 221-223.

14. LUEBKE D., REDDY M., WATSON B., COHEN J. AND VARSHNEY A. 2001. Advanced Issues in Level of Detail. *Course #41 at SIGGRAPH 2001*. Los Angeles, CA. August 12-17, 2001.

15. REDDY M. 1997. Perceptually Modulated Level of Detail for Virtual Environments. Ph.D. Thesis (CST- 134-97), University of Edinburgh.

16. MACIEL PWC. AND SHIRLEY P. 1995 Visual Navigation of Large Environments Using Textured Clusters, In *Proceedings of Symposium on Interactive 3D Graphics*, 95-102.

17. WATSON, B., FRIEDMAN, A. ANDMCGAFFEY, A. 1997 An evaluation of Level of Detail Degradation in Head-Mounted Display Peripheries. *Presence*, 6, 6, 630-637.

18. WATSON B., FRIEDMAN A. AND MCGAFFEY A 2001 Measuring and Predicting Visual Fidelity, *In Proceedings of SIGGRAPH 2001*, ACM Press / ACM SIGGRAPH, New York. E. Fiume, Ed., Computer Graphics Proceedings, Annual Conference Series, ACM, 213.220.

19. MCCONKIE GW. AND LOSCHKY LC. 1997 Human Performance with a Gaze-Linked Multi-Resolutional Display. ARL Federated Laboratory Advanced Displays and Interactive Displays Consortium, Advanced Displays and Interactive Displays First Annual Symposium, 25-34.

20. LOSCHKY LC. AND MCCONKIE GW. 1999 Gaze Contingent Displays: Maximizing Display Bandwidth Efficiency. ARL Federated Laboratory Advanced Displays and Interactive Displays Consortium, Advanced Displays and Interactive Displays Third Annual Symposium, 79-83.

21. LOSCHKY LC, MCCONKIE GW, YANG, J AND MILLER ME. 2001 Perceptual Effects of a Gaze-Contingent Multi-Resolution Display Based on a Model of Visual Sensitivity. In the ARL *Federated Laboratory 5th Annual Symposium - ADID Consortium Proceedings*, 53-58.

22. YEE H. 2000 Spatiotemporal sensitivity and visual attention for efficient rendering of dynamic environments, MSc. Thesis, Program of Computer Graphics, Cornell University.

23. YEE H., PATTANAIK, S., AND GREENBERG, DP. 2001 Spatiotemporal sensitivity and Visual Attention for efficient rendering of dynamic Environments, In *ACM Transactions on Computer Graphics*, vol. 20, no 1, 39-65.

24. DALY S. 1998 Engineering observations from spatiovelocity and spatiotemporal visual models. In I*S&T/SPIE Conference on Human Vision and Electronic Imaging III*, SPIE vol. 3299, 152-166.

25. ITTI L. AND KOCH C. 2000 A saliency-based search mechanism for overt and covert shifts of visual attention, In *Vision Research*, vol. 40, no 10-12, 1489-1506.

26. YANTIS S. 1996 Attentional capture in vision. In A. Kramer, M. Coles and G. Logan (eds), Converging operations in the study of selective visual attention, 45-76, American Psychological Association.

27. CATER K., CHALMERS AG. AND DALTON C. 2001 Change blindness with varying rendering fidelity: looking but not seeing, Sketch *SIGGRAPH 2001, Conference Abstracts and Applications.*

28. O.REGAN JK, DEUBEL H., CLARK JJ. AND RENSINK RA. 2000 Picture changes during blinks: looking without seeing and seeing without looking. *Visual Cognition.*7, 1, 191-212.