Visual Attention for Significantly Influencing the Perception of Virtual Environments

Jasminka Hasic*
Sarajevo School of Science and Technology
Bosnia and Herzegovina

Alan Chalmers†
University of Bristol
United Kingdom

Abstract

The Human Visual System (HVS) is a key part of the rendering pipeline. The human eye is only capable of sensing image detail in a 2° foveal region, relying on rapid eye movements, or saccades, to jump between points of interest. These points of interest are prioritised based on the saliency of the objects in the scene or the task the user is performing. These "glimpses" of a scene are then assembled by the HVS into a coherent, but inevitably imperfect, visual perception of the environment. In this process, much detail, which the HVS deems unimportant, may literally go unnoticed.

In this paper we use knowledge of the HVS to influence what our attention is attracted to in computer graphics imagery, and thus what we actually perceive in those images. We influence the affinity of subjects towards an object based on the complexity of the context that object is put into. The images are rendered using the Radiance lighting simulation system. In this way, we are able to significantly influence users' preferences in an e-commerce application. Detailed psychophysical studies are used to validate our approach.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism-Colour, shading, shadowing, and texture—;

Keywords: Visual perception, attention, saliency, e-commerce

1 Introduction

Computer graphics imagery is increasingly playing a key role in e-commerce applications. High quality rendered images and virtual/augmented reality environments are now regularly used to help buyers select the product of their choice. One such example is choosing furniture [ARIS 2003]. This computer graphics imagery is ultimately seen by the Human Visual System (HVS) and while the HVS is good, it is not perfect, and is very much influenced by the scene being considered [James 1890]. The HVS does not process an image sequentially in a raster-like fashion, but rather our eyes rapidly "jump" through rapid eye movements, known as saccades, between features of interest in the scene. Our visual attention is a coordinated action involving conscious and unconscious processes in the brain. This coordinated action enables us to, rapidly and efficiently, find and focus on relevant information within a potentially cluttered scene. While the quality of the rendered image of a product itself is of course important, the choice of context also plays a significant role in capturing the HVS's attention.

In this paper, we investigate the perception of objects based on the context they are put in. Our target application is the sale of cars using both static images and dynamic web-based environments. The subjects considered were young adults, in particular students of the University of Sarajevo and Sarajevo School of Science and Technology. A detailed study, involving 114 subjects is carried out to determine their strong preferences for choice of car colour. We understand that car colour preferences vary across different cultures and age groups. Throughout the project, we kept the variance in cultural characteristics and age of subjects at a minimum.

We then demonstrate that this preference can be significantly influenced by the choice of context and other perceptual cues in a rendered scene, in particular a dynamic background object.

The rest of the paper is structured as follows. Section 2 details previous work in computer graphics which has exploited the human visual system. Section 3 describes the experiments that were conducted. The results of these experiments are given in section 4, and finally, in section 5, conclusions are drawn and avenues for future work presented.

2 Background

In 1890, James described the two general visual attention processes, termed *bottom-up* and *top-down*, which determine where humans locate their visual attention [James 1890]. The bottom-up process is automatic, without volitional control, and is purely stimulus driven. Our eyes are automatically attracted to, for example, any movement, especially in our peripheral vision, certain colours, the size, shape, brightness, edges and orientation of objects. This is evolutionary; the movement may be a predator lurking in the bushes, and our ancestors needed to be able to easily discern at a distance the red ripe fruit amongst the green of the trees. Top-down processes, on the other hand, are task dependent causing the HVS to focus on only those parts of a scene which are necessary for the user's current task, for example looking for street signs, or targets in a game [Yarbus 1967]. When attention calls for concentrating on an object or task, the viewer will often fail to perceive an unexpected object, even if it appears at fixation. This is known as Inattentional Blindness [Mack and Rock 1998]

Models of the human visual system have been used in

*e-mail: jh@ssst.edu.ba
†e-mail: alan@compsci.bristol.ac.uk

 ISBN 978-80-86943-98-5

computer graphics to investigate the perceptual quality of images [Daly 1993; Myszkowski 1998; Bolin and Meyer 1998], improve the quality of rendered images [Ferwerda et al. 1996; McNamara et al. 2000; Myszkowski et al. 2001; Pattanaik et al. 1998; Ramasubramanian et al. 1999], or to reduce the complexity of models without any perceptual loss of quality, such as [Luebke and Hallen 2001; Reddy 1997; Watson et al. 2001]. A good overview of perceptual rendering can be found in [O'Sullivan et al. 2004]. More recently, researchers have used detailed models of the bottom-up (known as Saliency Maps [Itti and Koch 2000]) and top-down (known as Task Maps [Cater et al. 2003]) visual attention processes to significantly reduce the computation of high-fidelity global illumination calculations [Yee et al. 2001; Haber et al. 2001; Cater et al. 2003; Sundstedt et al. 2004].

In this paper we use the saliency of objects to direct visual attention to different parts of a scene and thereby influence the manner in which a viewer perceives the scene.

3 Model and the Experiment

3.1 Model and Preference Defining Property

The environment chosen for our experiments consists of a car modeled in Maya and rendered using the Radiance lighting simulation system, as shown in Figure 1.

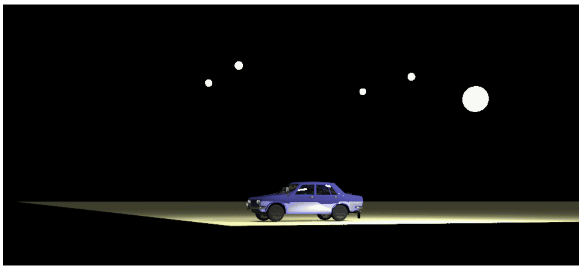

Figure 1: A radiance 3D scene

In order to carry out the experiments, we first need to establish a property of the car for which we have significant evidence that when altered will change the subjects' preference of the whole car. In [Hasic and Chalmers 2006] we showed that given the same car, if we just vary the car colour, the subjects, in general, prefer a blue car over the purple one. In this work 32 (16 male and 16 female) subjects were given the task to rank the chosen 6 car colours according to which car they would prefer to buy (1 - the most preferred car, 6 - the least preferred car). The background used in this experiment was neutral (medium grey). All car colours, Figure 2, were presented on one screen at once.

The table below summarizes statistics obtained for both male and female subjects. The number next to the car colour represents the average ranking of the particular car colour.

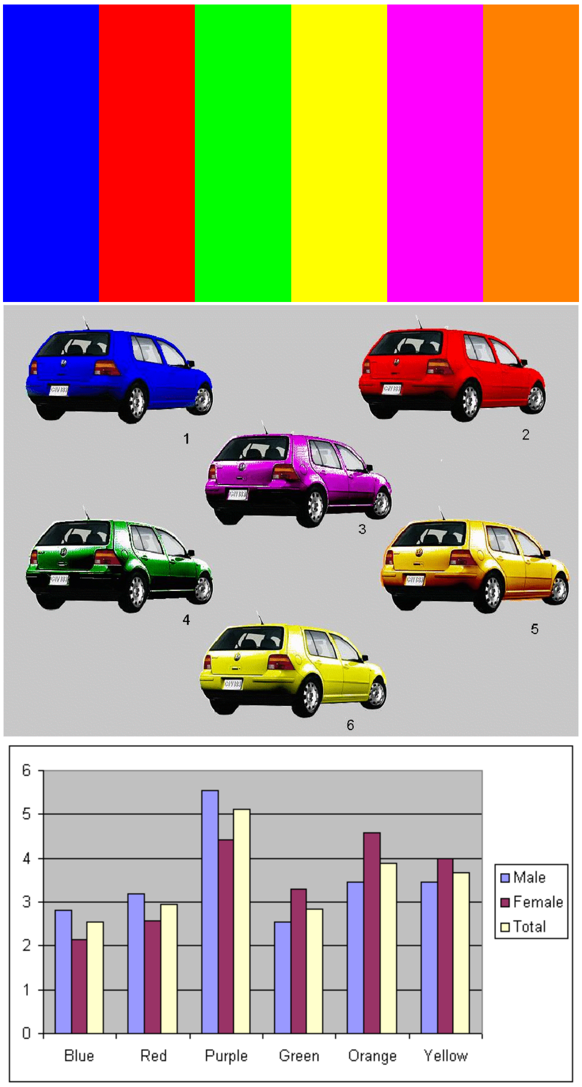

Figure 2: The different car colours shown and the subjects' preference

Female		Male		Total	
Colour	**Rank**	**Colour**	**Rank**	**Colour**	**Rank**
Blue	2.14	Green	2.55	Blue	2.56
Red	2.57	Blue	2.82	Green	2.83
Green	3.29	Red	3.18	Red	2.94
Yellow	4.00	Yellow	3.45	Yellow	3.67
Purple	4.43	Orange	3.45	Orange	3.89
Orange	4.57	Purple	5.55	Purple	5.11

This experiment established the blue car colour as most preferred and the purple car colour as least preferred one. To verify these results, another set of 16 subjects was asked to rank just those two car colours. The following preferences were obtained (see Figure 3):

Preferred Blue	Preferred Purple
86%	14%

The results demonstrate a clear difference in preference between those two car colours for our demographic. The blue and the purple car colours were used to represent preferred and non-preferred objects in the experiments conducted for

 ISBN 978-80-86943-98-5

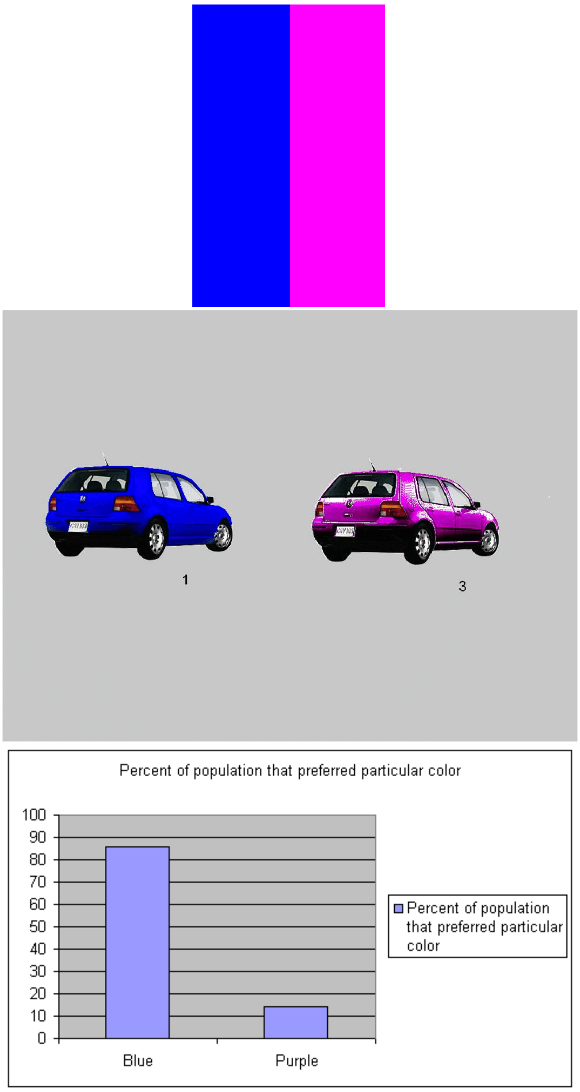

Figure 3: The best and the worst ranked car colour

this paper.

3.2 Experimental Setup

Two identical displays, with equivalent settings, placed next to each other were used in our experiments. 62 males students with normal, or corrected to normal vision, were the subjects for the experiments that followed. A pilot study was performed with 10 participants to determine the time required to comfortably complete the given task. Based on this pilot study the subject's response time was limited to 10 seconds. Most subjects responded well before this time limit. Due to the low number of female subjects in the overall student population used for this experiment, we chose to exclude answers from female students.

Each subject was seated in front of the two displays and had the same view of the both displays. On the two displays, a pair of images was presented and each subject was asked one question (not all subjects were asked the same question). Only one subject was present in the room where the experiment was conducted. Upon completion of the experiment, he was asked to keep the details of the experiment confidential and not to share any information about the experiment with other subjects since their experimental performance could be influenced by such information. The position of images in the two displays, within each pair, on the two displays was randomized to avoid bias.

4 Results

4.1 Comparison of objects on identical backgrounds

The results achieved in [Hasic and Chalmers 2006] about car preference based on the colour used 2D imagery. Since in this paper we are using a different car model and a global illumination 3D environment we had to verify that these earlier results would be consistent under the new conditions. The previous 2D imagery experiment tested the preference of car based on the car colour if the cars were placed on a neutral background (medium grey). In the first part of the experiment we verified that subjects would still prefer blue cars over purple ones if presented to them in the same context (on the same background). The two images were presented concurrently to one subject at a time and the subject was asked to answer the following question:

"Which car do you prefer?"

4.1.1 Car preference on neutral background

Each of the 10 subjects in this group was shown the pair of images in figure 4.

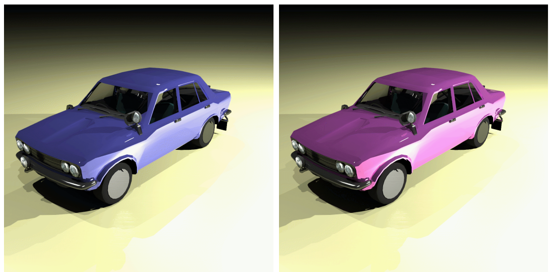

Figure 4: Car on neutral background

The following preferences were obtained from this group of subjects:

Preferred Blue	Preferred Purple
70%	30%

As we can see, this experiment confirmed that subjects do indeed prefer the blue new car model over the purple one on the neutral background.

4.1.2 Car preference on colourful background

The next group of 9 subjects was shown the pair of images in figure 5. This time the two cars were placed in a showroom-like setting. On the showroom walls colourful pictures were hung.

 ISBN 978-80-86943-98-5

Figure 5: Car on complex background

The results for this experiment were:

Preferred Blue	Preferred Purple
78%	22%

Even though the background of the two cars was changed from neutral to a more complex one, the subjects still highly preferred the blue car over the purple one.

4.1.3 Car preference on dynamic background

A final group of 9 subjects, in this part of the experiment, was shown the pair of images in figure 7. The background of these two cars was the same showroom-like background from the previous experiment, but with a dynamic addition. An animated button was added to the two pictures that were hanging in the show room. The face of the button had a label which read "CLICK HERE". As time went by, the face of the button opened in a window-like fashion and a new message was revealed reading "TO FIND OUT MORE" (see figure 6).

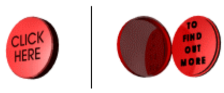

Figure 6: 2 frames of the dynamic button

Figure 7: Car on complex-dynamic background

The preferences for this part of the experiment were:

Preferred Blue	Preferred Purple
88%	12%

As can be seen from this subsection's experiment, if two objects are presented on the identical background and the subjects are given a task (in this case the question: "Which car do you prefer?") their visual behaviour is consistent with a top-down process and what happens outside the zone of interest is not visually important.

4.2 Comparison of different backgrounds

Now that we clearly established that the blue car is preferred over the purple one we needed to devise an experiment in which we could manipulate the environment so that subjects would no longer prefer the blue car over the purple one. We did so by using different contexts for the two objects. The hypothesis was that if the less preferred car was put in a more complex environment then it is likely that, as some subjects might prefer more complex images, just because they are more complex, they would thus choose the less preferred car. We performed this experiment in two stages.
In the first stage the subjects were asked to answer the question:

1. Which image do you prefer?

In the second stage the subjects were asked:

2. Which car do you prefer?

A different set of subjects was used for the two stages of the experiment. We thus wished to investigate whether, by introducing both bottom-up and top-down visual processes in images we could significantly influence subjects' preferences of objects. We used two pairs of images for this experiment. The first pair consisted of the blue car on neutral background and the purple car on the colourful background as in Figure 8. The second pair consisted again of the blue car on the neutral background and the purple car on colourful-dynamic background as shown in Figure 9. The hypothesis we are considering is: Even when subjects are given a task within an image (such as in question 2), if the context is sufficiently salient, the subject's preference could be significantly influenced.

Figure 8: Car on complex background

Figure 9: Car on complex-dynamic background

4.3 Question: Which image do you prefer?

These are the results of 16(8+8) subjects' preferences when presented a pair of images in figure 8.

 ISBN 978-80-86943-98-5

1. Comparison of static and complex backgrounds

Preferred Blue	Preferred Purple
37%	63%

The following are the results of subjects' preferences when presented a pair of images in figure 9.

2. Comparison of static and complex-dynamic backgrounds

Preferred Blue	Preferred Purple
25%	75%

As expected, when there was no clear task given within the image, the subjects preferred the more complex image dispate the fact the car that appears on this image is the car less preferred. During this question, subjects' perception worked in bottom-up fashion and most of the subjects were attracted to the more complex image.

4.4 Question: Which car do you prefer?

The following are the results of 18(9+9)subjects' preferences when presented a pair of images in figure 8.

1. Comparison of static and complex backgrounds

Preferred Blue	Preferred Purple
78%	22%

The following are the results of subjects' preferences when presented a pair of images in figure 9.

2. Comparison of static and complex-dynamic backgrounds

Preferred Blue	Preferred Purple
44%	56%

As we can see from the results above, the subjects responded in the top-down fashion with the first pair (figure 8) of images (their responses were task driven). Even though we introduced salient objects (the pictures) in the background, they were not salient enough to draw attention away from the task object (car). However, the experiment with the second pair images (figure 9) shows different results. With even more saliency added to the background (an opening window with a different text when the window is opened and closed) we were able to influence the subjects' preference of the task object itself (car). This suggests that we have successfully drawn attention away from the task region (car). Our test group in this case slightly preferred the purple car that before, in all being equal circumstances, was far less preferred than the blue car.

5 Conclusions and Future Work

In this paper we wished to explore how a viewer's preference of objects such as cars can be influenced by the overall appearance of realistic virtual environment. In particular we wanted to investigate whether the choice of the complexity of the background was sufficient to significantly influence the perception of the object of interest in the image.

Our results have shown that by only changing the background of objects to a complex and colourful one, we cannot significantly influence the perceptual preference of subjects. However, what we did discover is that other parameters of the image, in particular the presence of a highly salient animation in the background can significantly influence subjects' perception of the virtual environment.

So, it is possible to create an environment where, even though the subjects are performing a task in a specific region of the image, their overall perception of a targeted object can still influenced by what is happening outside of that specific region.

This is suggesting that the top-down process can be influenced in a bottom-up fashion.

This result has important implications for the design of virtual and augmented reality environments.

In future work, we will consider the effect that even more complex virtual environments may have on the perception of e-commerce objects and investigate whether we can use such salient cues to attract a viewer to consider "purchasing" additional objects in the virtual environment.

References

ARIS. 2003. Augmented reality image synthesis through illumination reconstruction and its integration in interactive and shared mobile ar-systems for e-(motion)-commerce applications proposal number: Ist-2000-28707. In *http://aris-ist.intranet.gr/*.

BOETTCHAR, L. C. 2005. No. one car color in america. In *http://channels.netscape.com/autos/package.jsp?name=fte/carcolors/carcolors*.

BOLIN, M., AND MEYER, G. 1998. A perceptually based adaptive sampling algorithm. In *SIGGRAPH 98 Conference Proceedings*, 299–310.

CATER, K., CHALMERS, A., AND WARD, G. 2003. Detail to attention: Exploiting visual tasks for selective rendering. In *Proceedings of the Eurographics Symposium on Rendering*, 270–280.

DALY, S. 1993. The visible differences predictor: An algorithm for the assessment of image fidelity. In *Digital Images and Human Vision*, MIT Press, Cambridge, Massachusetts, 179–206.

FERWERDA, J. A., PATTANAIK, S. N., SHIRLEY, P., AND GREENBERG, D. P. 1996. A model of visual adaptation for realistic image synthesis. In *Computer Graphics (SIGGRAPH '96 Proceedings)*, 249–258.

HABER, J., MYSZKOWSKI, K., YAMAUCHI, H., AND SEIDEL, H.-P. 2001. Perceptually guided corrective splatting. In *Computer Graphics Forum*, vol. 20, 142–152.

HASIC, J., AND CHALMERS, A. 2006. Visual attention for influencing the perception of virtual environment. In *SCCG2006 Proceedings (a short paper)*.

ITTI, L., AND KOCH, C. 2000. A saliency-based search mechanism for overt and covert shifts of visual attention. In *Vision Research*, vol. 40, 1489–1506.

JAMES, W. 1890. A saliency-based search mechanism for overt and covert shifts of visual attention. In *Principles of Psychology, New York: Holt*.

LUEBKE, D., AND HALLEN, B. 2001. Perceptually driven simplification for interactive rendering. In *Proceeding of 13h Eurographics Workshop on Rendering*, 221–223.

MACK, A., AND ROCK, I. 1998. Inattentional blindness. In *Proceedings of Symposium on Interactive 3D Graphics, Massachusetts Institute of Technology Press*.

MCNAMARA, A., CHALMERS, A. G., TROSCIANKO, T., AND GILCHRIST, I. 2000. Comparing real and synthetic scenes using human judgements of lightness. In *Proceeding of 12th Eurographics Workshop on Rendering*.

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.*, 221–230.

MYSZKOWSKI, K. 1998. The visible difference predictor: Application to global illuminarion problems. In *Proceeding of Eurographics Workshop on Rendering, G. Drettakis and N. Max Eds.*, 223–236.

O'SULLIVAN, C., HOWLETT, S., MORVAN, Y., MCDONNELL, R., AND O'CONNOR, K. 2004. Perceptually adaptive graphics. In *STAR, Eurographics*, 141–164.

PATTANAIK, S. N., FERWERDA, J. A., FAIRCHILD, M. D., AND GREENBERG, D. P. 1998. A multiscale model of adaptation and spatial vision for realistic image display. In *Computer Graphics (SIGGRAPH '98 Proceedings)*, 287–298.

RAMASUBRAMANIAN, M., PATTANAIK, S. N., AND GREENBERG, D. P. 1999. A perceptually based physical error metric for realistic image synthesis. In *Computer Graphics (SIGGRAPH '99 Proceedings)*, 73–82.

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

SUNDSTEDT, V., CHALMERS, A., CATER, K., AND DEBATTISTA, K. 2004. Top-down visual attention for efficient rendering of task related scenes. In *VMV 2004 - Vision, Modelling and Visualization*.

WATSON, B., FRIEDMAN, A., AND MCGAFFEY, A. 2001. Measuring and predicting visual fidelity. In *Proceedings of SIGGRAPH 2001*, 213–220.

YARBUS, A. 1967. Eye movements during perception of complex objects. In *Eye Movements and Vision*, 171–196.

YEE, H., PATTANAIK, S., AND GREENBERG, D. 2001. Spatiotemporal sensitivity and visual attention for efficient rendering of dynamic environments. In *ACM Transactions on Computer Graphics*, vol. 20, 39–65.

 ISBN 978-80-86943-98-5