Visual Attention in 3D Video Games

Magy Seif El-Nasr
Assistant Professor
Penn State University
College of Information Sciences and Technology
814-865-6166
magy@ist.psu.edu

Su Yan
Graduate Student
Penn State University
College of Information Sciences and Technology
syan@ist.psu.edu

ABSTRACT

Understanding players' visual attention patterns within an interactive 3D game environment is an important research area that can improve game level design and graphics. Several graphics techniques use a perception based rendering method to enhance graphics quality while achieving the fast rendering speed required for fast-paced 3D video games. Game designers can also enhance game play by adjusting the level design, texture and color choices, and objects' locations, if such decisions are informed by a study of players' visual attention patterns in 3D game environments. This paper seeks to address this issue. We present results showing different visual attention patterns that players exhibit in two different game types: action-adventure games and first person shooter games. In addition, analyzing visual attention patterns within a complex 3D game environment presents a new challenge because the environment is very complex with many rapidly changing conditions; the methods used in previous research cannot be used in such environments. In this paper, we will discuss our exploration seeking a new approach to analyze visual attention patterns within interactive 3D environments.

Categories and Subject Descriptors
H.5. [**Information Systems**]: Information Interfaces and Presentation (e.g., HCI). J.5 [**Computer Applications**]: Arts and Humanities.

General Terms
Design and Experimentation.

Keywords
Visual perception, games, visual attention.

ACE 06, June 14-16, 2006, Hollywood, California, USA.
Copyright 2006 ACM 1-59593-380-8 /06/0006 ...$5.00.

1. INTRODUCTION
Visual attention has long been an important topic in psychology and cognitive science. Visual attention is gaining more importance in graphics rendering. Recently, results from visual attention research [1-7] are being adopted by computer graphics research. Due to speed limitations, there has been a movement to use a perception-based rendering approach where the rendering process itself takes into account where the user is most likely looking [4]. Examples include trying to achieve real-time global illumination by concentrating the global illumination calculation on salient parts of the scene [8] or varying the level of detail of terrain rendering depending on salience of locations [9]. Video games have achieved a high degree of popularity because of such advances in computer graphics. These techniques are also important because they allowed adoption of game environments in health and training applications.

In addition, we believe that research on visual attention can further improve the design of game environments, by decreasing frustration and increasing engagement. Many non-gamers get lost in 3D game environments, or they don't pick up an important item because they don't notice it. A study of visual attention patterns in video games can be used to inform designers where to place items in a level or how to choose colors and use other visual tools to stimulate attention and eliminate these problems.

While the visual attention process has been extensively researched and is still under much research within psychology and cognitive science, most experiments within these fields employ simplified strictly controlled composition of 2D objects, e.g. combination of lines and blocks [10]. While some researchers have used photographs as well as simple 3D synthetic objects [11, 12], these environments are no match to game environments. 3D Video games employ complex 3D architectures, objects, and characters that move and interact in 3D space. While some of the concepts generated by previous visual attention research may transfer to video game environments, an experiment is needed to confirm this hypothesis.

In this paper, we seek to study the visual attention process within a video game environment. We focus on exploring whether visual attention within a 3D game follows the bottom-up or top-down visual attention theories. We also seek to explore the difference between visual attention patterns exhibited in two game genres:

first person shooters and action-adventure games. Our hypothesis is that different game genres will stimulate different visual attention patterns depending on the activities the player is engaged in. We also seek to identify saliency of the visual cues present in the environment.

Studying visual attention within a 3D environment as the one employed by video games is hard. While there are methods established that study visual search through the use of eye tracking techniques within interactive environments, most of these methods focus on web applications, which are 2D, slow-paced (i.e. no rapid movements), and static in terms of their interface (e.g., buttons mostly stay in the same place on the monitor). By comparison a game environment is highly complex using 3D graphics with many moving elements and almost no static interface.

In this paper, in addition to discussing some preliminary results showing visual attention patterns in 3D game environments, we will also present a new method by which eye-tracking data is analyzed within such complex 3D environments. We hope this method can stimulate researchers to begin to formalize methods for evaluating visual attention in 3D complex environments, and thus informing research in psychology and cognitive science as well as enhancing 3D graphics rendering and game level design.

Although the experimental results presented in this paper are exploratory since our subject and game pools are quite small, they still present several contributions. First, our results confirm previous visual attention research concerning saliency of stimuli in a bottom-up visual attention process [13-15]. Second, we present a method for studying and analyzing visual attention within a 3D game environment. Third, we present interesting results uncovering different visual attention patterns within two different game types: action-adventure games and first person shooters.

According to our results, both action-adventure games and first person shooter games tend to illicit a top-down visual search pattern. Visual stimuli are more salient when located near objects that fit player's top-down visual search goals. In terms of eye movements, players playing first person shooter games tend to concentrate their eyes on the center of the screen while in an action-adventure game players' gaze explored the entire screen.

In this paper, we first briefly introduce important visual attention theories that are relevant to our work. Then we describe the experiment design that we undertook in detail, describing the games chosen, the data collected, and the analysis process. We then discuss some limitations of the described experiments. We will conclude by describing the contributions and future research.

2. PREVIOUS WORK

Much research evidence are in favor of a two-component framework for visual attention: bottom-up and top-down [13-15]. The bottom-up approach identifies several features or properties that subconsciously attract attention, including movement, color, contrast, and brightness. For example, a red rose in a green bush will automatically attract attention because the rose is of a warmer color than the green bush and because of the contrast in color. On the other hand, the top-down mechanism is under voluntary control, and biases the observer towards selecting stimuli that are related to the observer's goal when perceiving a scene. Studies have shown that there is a strong correlation between a viewer's eye movements and the visual task that he is performing [16].

Several visual features were identified to have a major influence on the bottom-up visual attention process. These features include color, brightness, contrast, orientation, shape, size, lines, and motion. In our study, we concentrate our experiment on testing two features: color contrast and motion.

Color has long been accepted as a pre-attentive feature [17-22]. Search for a target color among homogeneous distracters is efficient as long as the target and distracter colors are not too similar [23, 24]. Because it is so effective, selection by color is common in the design of visual displays. There is evidence for a pre-attentive ability to detect a change in color. D'Zmura and Mangalick [25] found that subjects can search efficiently for a target undergoing a smooth color change when such color is different from (and often complementary to) that of the distracters.

When it comes to guiding attention towards more complex targets, motion is a very effective feature [26]. Moving targets pop-out of a field of stationary distracters [27]. Search for random or linear motion among stationary distracters is very efficient. However, search for the absence of motion is less efficient [26].

High-level top-down driven visual attention covers goal-directed factors; it is essentially a task-dependent visual attention model. Wolfe and Gancarz [28] produced a biologically inspired guided search model that controls the bottom-up features that are relevant to the current task by a top-down mechanism, through varying the weighting of the feature maps. However, it is unclear what the task relevant features are, particularly in the case of action recognition.

3. EXPERIMENT

In designing the experiment, we made three hypotheses, which can be described as follows:

- Bottom-up visual features affect players' perception of a 3D video game environment. We specifically focus on color and motion, and study their effectiveness in grabbing players' attention in a 3D environment.
- Since video games are highly goal oriented, we hypothesize that top-down visual features are more effective in attracting players' attention than bottom-up visual features.
- Eye movement patterns may differ among different game genres. We believe that eye movement patterns reveal the way that game players visually perceive the 3D environment. Since the pace and visual composition of game levels are different in different genres, we hypothesize that eye movement patterns among different game genres are also different.

3.1. Method

We used an ISCAN ETL-500 head-mounted eye tracker to collect players' visual attention patterns. The eye tracker receives a video signal input from a VCR, which is connected to an Xbox. The eye tracker superimposes a cursor indicating viewer's eye sight position onto the video signal. The output signal is then passed to a monitor for real-time observation and a DVD recorder for recording. The recorded data will be used for further study and analysis, as will be discussed in a later section.

3.2. Participants

Six participants were gathered from the Pennsylvania State University. The age of these participants ranged from twenty to thirty years old. We acknowledge that six is a small number to derive any conclusive results. The analysis process that we devised is very extensive requiring several months of study for only six participants. Therefore, gathering data for more than six participants will make this analysis process almost impossible to perform. This number can be justified, however, for a preliminary study in such a new area, but further experimentation is needed to make more conclusive contributions.

We asked the six participants to fill in a survey that assessed their gaming experience. We then divided them into three groups according to their gaming experiences, as follows:

- novice gamers: participants who never or rarely play video games;
- casual gamers: participants who do not play video games as often as core gamers, but are better in playing video games than novice gamers.
- core gamers: participants who play video games regularly and are good at video games;

Among the six participants, two were core gamers, two were novice gamers, and two were casual gamers. We also note that two of the six participants were girls; one is a novice gamer, while the other is a casual gamer.

3.3. Procedure

After gathering the gaming experience data, we asked the participants to put on the head mounted eye-tracker and play a fighting Xbox game, called *Soul Caliber 2* (shown in figure 1). They were asked to play the game for a period of time until they get accustomed to the equipment and the head mounted eye-tracker. We chose a fighting game for this preliminary stage because fighting games are different from the other games tested in the experiment, thus eliminating any bias that may be stimulated through the role of memory or pattern recognition. The time given to participants in this phase varied, e.g. *novice gamers* were given more time to get familiar with game controllers.

After the participants felt comfortable with the equipment and were ready to start the experiment (they were asked to inform us when comfortable and ready to start), we started to configure the eye-tracker system. The length of this process varied according to participant's different visual characteristics, including frequency of blinking or head movements during game play.

After these preparations were done, we asked participants to play two video games and recorded the video signal (game video superimposed by cursor positions) onto a DVD. We varied the order that the video games were played across the pool. Each participant was asked to play each game for ten minutes. The video games played were *Legacy of Kain Blood Omen II* (action-adventure game) and *Halo II* (first person shooter game).

Figure 1: A screenshot from *Soul Caliber 2*

4. RESULTS AND ANALYSIS

4.1. Method of Analysis

Because analyzing eye-tracking data in a complex 3D environment such as the ones used in video games is quite new, and not much work has been reported on it, there are no effective methodologies that we can follow. Therefore, we spent several months exploring different approaches until we arrived at one that was relatively successful given the nature of the environment and the data collected.

Our experiment differs from previous eye-tracking experiments in that we gather eye-tracking information within a complex 3D environment. Unlike previous eye-tracking experiments that gather eye-tracking data while participants are viewing static pictures, web pages, or documents, in a video game the visually perceived environment is complex and dynamic. The objects on the screen are changing and the content of the environment is also changing all the time. Therefore, existing eye-tracking analysis methods are not suitable for our purpose.

For our experiment, we first recorded the video signal, which is the video game signal superimposed by a cursor signal, on to DVD discs. The DVD signal is transformed into an AVI file for further quantitative analysis. We first watched the AVI file and annotate the video describing the context of play. We also cut the AVI file into several parts that are of interest to our experimental purpose. We perform this step for two reasons: first, the analysis of the eye tracking data should be guided by the game context. Second, experiments on video games contain a huge amount of data, which may not be all useful (e.g. a period that the player is looking at their game controller rather than the screen, or the period where no cursor appears on the screen because the subject is blinking or some other reasons). Watching the AVI files, annotating them, and cutting them into smaller files helped us (a) concentrate on the most informative part of the video, and (b) understand the player’s context.

After selecting the segments of interest, we input the video clips into a computer program that we coded to extract information from the key frames (2 key frames per second) of the video clip. The extracted information includes the coordinates of the avatar, the coordinates of the objects that we want to track, and the coordinates of the cursor. Our further analysis is based on these three categories of coordinate data.

Note that the coordinate extraction process is done manually by marking positions of the avatar, the objects, and the cursor. The reason is that the avatar and objects in the 3D environment are constantly changing orientations and scale as they move within the environment. Such changes present a major challenge to current automatic image analysis or pattern recognition techniques. We tried several existing algorithms and software for object tracing in video (e.g. *Adobe After Effects*), but found they are not effective for our purposes and decided to take a manual approach.

Using these coordinate data, we were able to make several deductions concerning the visual attention patterns exhibited by players in 3D environments. These deductions also take into account the annotated game context. The results presented in the rest of this paper are based on this method.

4.2. Evidence of bottom-up visual attention using low-level features

Through studying the videos and coordinate data extracted, we found that bottom-up visual features, including color contrast and movement, subconsciously trigger the visual attention process, thus verifying the bottom-up visual attention theory.

To back up this claim with some data, we would like to further discuss one case. We use one specific example that occurred while interacting with the action-adventure game *Legacy of Kain Blood Omen II*. In the game, the player controls an avatar in third person view.

In one video segment, the player wonders around in a big castle trying to find an exit. At this stage of the game, the player has successfully finished the first part of the game and has accumulated some experience. There is a tiny object that changes color to red and becomes brighter, and thus becomes more salient according to the bottom-up visual attention theory. The object is, in fact, important because it can improve the avatar's health, but its narrative function is trivial. This object, however, was not on the right path to an exit, and thus would distract participants from the exit route. Figure 2 shows several key frames extracted from the video clip that we used in the experiment.

As we analyzed the recorded video for the six participants, we found that only two out of the six subjects followed the path of the object. As we analyzed the video, we found that the reason they followed the path was not related to the object, because participants began to follow the path where the object resides, before the object of interest even appeared on the screen (as shown in figure 2 a-c). When the object appeared on the screen with the same color as the background, none of the two participants noticed it. They went straight to the door at the end of the path and did not pay attention to the object. But as the object began to change its color to bright red (figure 2d), both participants paid attention to it. This confirms the theory of saliency described by the bottom-up visual attention theory.

(a)

(b)

(c)

(d)

Figure 2: Screenshots from Legacy of Kain

By analyzing the eye-tracking data, we can clearly notice this attention shift. Figures 3 and 4 show the player's eye movement for 90 frames extracted from the video recordings. The cursor line shows the cursor's position on the screen. The avatar line denotes the avatar's position, and the object line denotes the object of interest's position.

The figure shows that, at the beginning, the player was paying attention to the avatar, controlling the avatar to the direction that he thought might lead to the exit. Then the object of interest appeared on the screen. At this time, the object was brown in color, similar to wall color (i.e. there was no contrast). The player did not notice it, he moved the avatar to the end of the hall and stood still, looking for an exit. At this period, the player did not pay attention to the avatar, but started looking around for an exit. The object of interest then began to turn to bright red. The player's attention shifted to the bright red colored object.

This attention shift is illustrated in figures 3 and 4. At the beginning, the figure shows the player was looking (the cursor line) at the avatar (the avatar line). Then at frame 11 the object of interest appeared with brown color. As it can be seen, the player's attention shifted from looking at his avatar to looking around, as shown the line of cursor (the cursor line) does not follow the line of avatar (the avatar line) any more. At frame 32, the object started to become bright red. This movement and change of color and brightness caused the player to shift his attention to the object; as we can see at frame 32, there is an obvious shift of the player's gaze (the cursor line) to positions near the object (the object line). At frame 77, after the player shifted his attention to the object, he moved the avatar to the object as indicated by the merge of the three lines.

4.3. Evidence of Top-down and Bottom-up Visual Attention

By analyzing the videos collected, we found that, since action-adventure games, such as *Legacy of Kain Blood Omen II*, are highly goal oriented, top-down visual features control players' attention more than bottom-up visual features.

Evidence of this claim is evident from several examples in the video tapped interactions. We will discuss one specific example. In one segment of *Legacy of Kain Blood Omen II*, the player found himself trapped and attempted to find an exit. All the participants in our experiment first paid attention to doors, no matter if the door was opened or closed. They tried the doors; then they shifted their attention to something that looks like an exit, e.g. the fireplace. Figure 5 shows the scene where the bright wall is the right exit point. The designer intended the wall to be noticeable, hence gave it a very bright color. However, only one subject from the six subjects noticed the wall in first sight. This incident happened around three times within the ten minute segment of the interaction. In the second and third incident, only two out of the six subjects noticed the wall. Most players after several trials were able to finally grasp that the wall was the object presenting an exit. However, the brightness and color did not impact this decision for the participants.

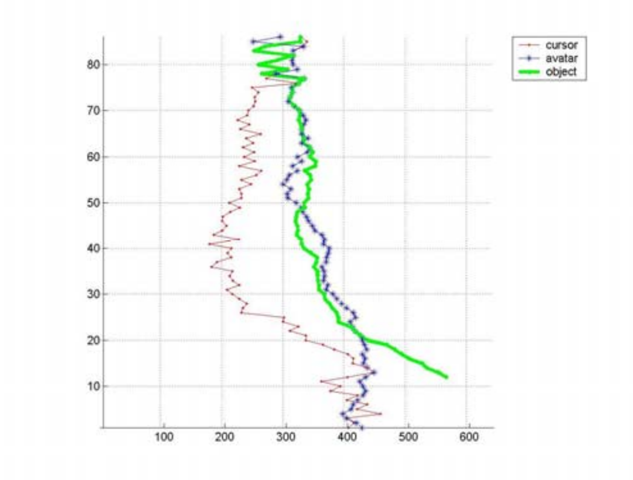

Figure 3: 2D view (y axis: frame 1 to frame 90; x axis: long length of the screen size, 640*480)

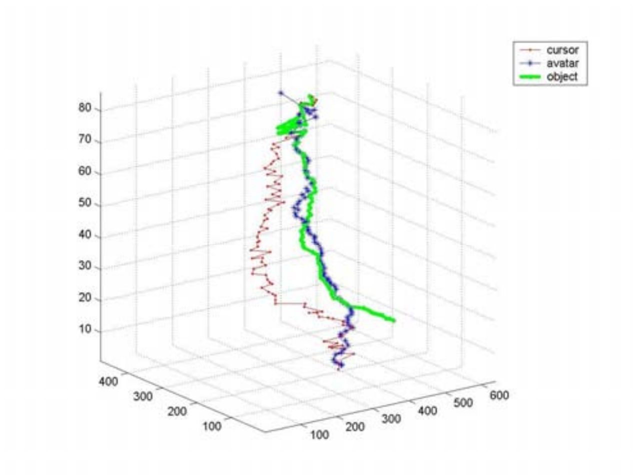

Figure 4: 3D view (640*480*90 frames)

Analysis of the recorded data showed that participants eventually tried the wall because it was positioned right next to a door. This suggests that, if game level designers want to make some important object more noticeable to the player, they have to choose positions for these objects such that the position would match a goal oriented visual search pattern. For example, if player's goal is to find an exit, then doors, windows, or stairs will be important, and thus objects that need to be quickly noticed should be placed near these structures. Game designers may consider using this finding to balance difficulties and playability of action-adventure video games. Note that we acknowledge that game designers may have made these decisions to intentionally deviate participants from obvious routes, thus making the experience more challenging. We suggest that designers should be informed of these results, and use them in their designs as they see fit.

Figure 5: Screenshot of the exit

4.4. Eye-movement patterns

We compare visual attention patterns between two games: *Legacy of Kain Blood Omen II* (action-adventure game) and *Halo II* (first person shooter game). Eye-tracking data shows that in the first person shooter game, players paid attention only to the center of the screen, where the cross of their gun was located. Most of the time, they shifted their eye-sight only to read the information that appeared on the corner of the screen. Figure 6 shows the eye movement pattern in *Halo II*. The figure is plotted using collected eye movement data of 350 consecutive frames from the game *Halo II*. From this figure, we can see that the player's eye sight is always located on the center part of the screen. The range of the eye sight is relatively small, and is around the cross.

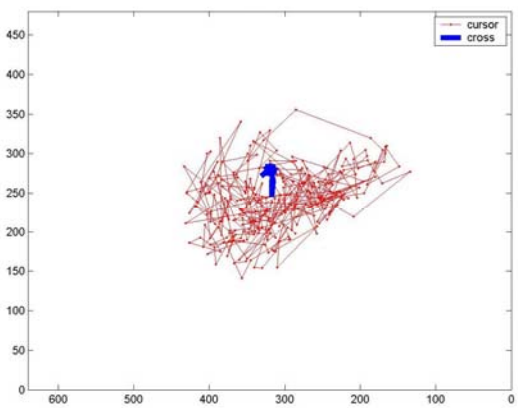

Figure 6. Eye movement pattern in *Halo II*

On the contrary, in the action-adventure games, the player's visual pattern is quite different; it covers more area of the whole screen. Figure 7 shows eye movement patterns in *Legacy of Kain Blood Omen II*. The figure is drawn by using 350 consecutive frames from the action-adventure game *Legacy of Kain Blood Omen II*. The range of eye movements is larger than in *Halo II*. This is a reasonable result, since action-adventure games are relatively slow paced, and most of the time, the avatar is safe so the player does not have to pay attention to them. Rather, the player needs to spend their attention to the important objects that are important for their goals. Figure 8 compares the eye gaze patterns of *Halo II* and *Legacy of Kain Blood Omen II*. We can see that *Legacy of Kain Blood Omen II* has much larger range of eye movements than *Halo II*.

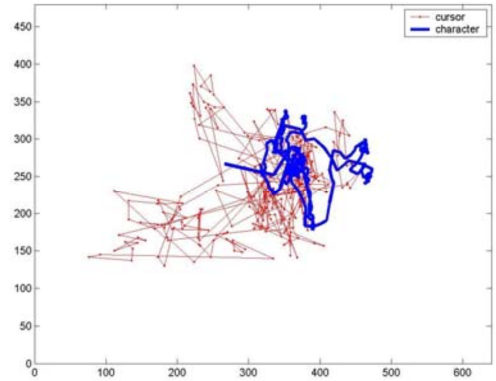

Figure 7: Eye movement pattern in *Legacy of Kain Blood Omen II*

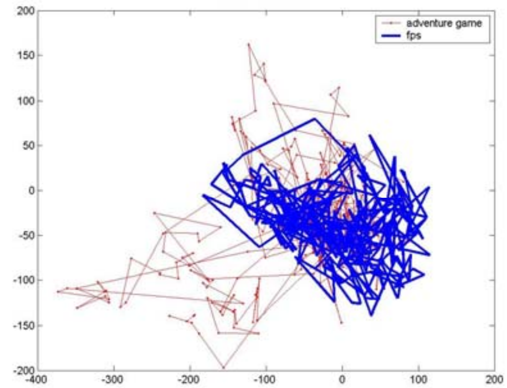

Figure 8: comparison between visual patterns exhibited in *Halo II* and *Legacy of Kain Blood Omen II*

5. LIMITATIONS

There are many limitations to the experiment and analysis discussed. First, the participant pool is quite small. A sample of six participants is very small to make any kind of conclusive remarks. However, we believe we showed enough evidence to warrant more experimentation, and to make the claim that such results can be used to inform game level design and graphics rendering. Second, we used only one game for each game genre. In order to make a general remark on the visual patterns exhibited in any one genre we need a more representative sample of games. We keep this for future research. Third, we know that 20% of boys are green/red color blind. We did not test for this condition. In the next stages of this experiment, this condition will be accounted for. In addition, the results collected and analyzed did not take into account camera movement or perspective which may influence the deductions and interpretations of the results as well as the methods used. In future research, we aim to investigate camera movement, field of view, as well as camera perspective (first vs. third person).

6. CONCLUSION

In conclusion, we have presented several results showing the existence of both bottom-up and top-down visual attention patterns within 3D games. We also showed results comparing visual patterns in *Halo II* (first person shooter) and *Legacy of Kain Blood Omen II* (action-adventure game). These results can be applied and used to inform game level design and graphics rendering. We believe that more experimentation is needed to formulate a model of visual attention in 3D games to enhance game level design and rendering. We have presented evidence to warrant more experimentation. We have also presented a new analysis method for analyzing eye tracking data in a 3D game environment.

7. REFERENCES

[1] V. Sundstedt, A. Chalmers, K. Cater, and K. Debattista, "Top-down visual attention for efficient rendering of task related scenes," presented at Vision, Modeling, and Visualization, Stanford, 2004.

[2] K. Cater, A. Chalmers, and G. Ward, "Detail to attention: Exploiting visual tasks for selective rendering.," presented at Eurographics Symposium on Rendering, 2003.

[3] K. Cater, A. Chalmers, and P. Ledda, "Selective quality rendering by exploiting human inattentional blindness: Looking but not seeing," presented at Symposium on Virtual Reality Software and Technology, 2002.

[4] J. Haber, Myszkowski, H. Yamauchi, and H. Seidel, "Preceptually guided corrective splatting," *Computer Graphics Forum*, vol. 20, 2001.

[5] K. Myszkowski, T. Tawara, H. Akamine, and H. Seidel, "Preception-guided global illumination solution for animation rendering," presented at Siggraph, 2001.

[6] H. Yee, S. Pattanik, and D. P. Greenberg, "Spatiotemporal sensitivity and visual attention for efficient rendering of dynamic environments," *ACM Transaction on Graphics*, vol. 20, pp. 39-65, 2001.

[7] R. Dumont, F. Pellacini, and J. Ferwerda, "Perceptually-driven decision theory for interactive realistic rendering," *ACM Transaction on Graphics*, vol. 22, 2003.

[8] K. Myszkowski, "Perception-based global illumination, rendering, and animation techniques," presented at Conference on Computer Graphics, 2002.

[9] B. Turner, "Real-time dynamic level of detail terrain rendering with ROAM," in *Gamasutra*, vol. April 2003, 2003.

[10] C. T. Weidemann, D. E. Huber, and R. M. Shiffrin, "Confusion and compensation in visual perception: Effects of spatiotemporal proximity and selective attention," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 31, pp. 40-61, 2005.

[11] D. J. Simons and C. F. Chabris, "Gorillas in our midst: Sustained in attentional blindness for dynamic events," *Perception*, vol. 28, pp. 1059-1074, 1999.

[12] C. O'Sullivan, "Collisions and attention," *ACM Transaction on Applied Perception*, vol. 2, pp. 309-321, 2005.

[13] A. Treisman, "A feature-integration theory of attention," vol. 12, pp. 97-136, 1980.

[14] W. James, "A saliency-based search mechanism for overt and covert shifts of visual attention," *Journal of psychology*, 1890.

[15] J. R. Bergen and B. Julesz, "Parallel versus serial processing in rapid pattern discrimination," *Nature*, vol. 303, 1983.

[16] A. L. Yarabus, "Eye Movements During Perception of Complex Objects," in *Eye Movements and Vision*, 1967, pp. 171-196.

[17] C. Bundesen and L. F. Pedersen, "Color Segregation and Visual Search," *Perception and Psychophysics*, vol. 33, pp. 487-493, 1983.

[18] R. C. Carter, "Visual Search with Color," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 8, pp. 127-136, 1982.

[19] E. W. Farmer and R. M. Taylor, "Visual Search Through Color Displays: Effects of Target-background Similarity and Backgroun Uniformity," *Preception and Psychophysics*, vol. 27, pp. 267-272, 1980.

[20] B. F. Green and L. K. Anderson, "Color Coding in a Visual Search Task," *Journal of Experimental Psychology*, vol. 51, pp. 19-24, 1956.

[21] S. L. Smith, "Color Coding and Visual Search," *Jorunal of Experimental Psychology*, vol. 64, pp. 434-440, 1962.

[22] L. G. Williams, "The Effect of Target Specification on Objects Fixated During Visual Search," *Preception and Psychophysics*, vol. 1, pp. 315-318, 1966.

[23] A. L. Nagy, R. R. Sanchez, and T. C. Hughes, "Visual Search for Color Differences with Foveal and Peripheral Vision," *Journal of Optical Society of America - A*, vol. 7, pp. 1995-2001, 1990.

[24] A. L. Nagy and R. R. Sanchez, "Critical Color Differences Determined with a Visual Search Task," *Journal of Optical Society of America - A*, vol. 7, pp. 1209-1217, 1990.

[25] M. D'Zmura and A. Managalick, "Detection of Contrary Chromatic Change," *Journal of Optical Society of America - A*, vol. 11, pp. 543-546, 1993.

[26] J. M. Wolfe, "Visual Attention," in *Seeing*, D. V. K. K, Ed., 2 ed. ed, 2000, pp. 335-386.

[27] M. Dick, S. Ullman, and D. Sagi, "Parallel and Serial Processes in Motion Detection," *Science*, vol. 237, pp. 400-402, 1987.

[28] J. M. Wolfe and G. Gancarz, "Guided Search 3.0.," in *Basic and Clinical Applications of Vision Science*: Kluwer Academic Publishers, Dordrecht, 1996, pp. 189-192.