

Game Design Guided by Visual Attention

Li Jie James J Clark

Centre of Intelligent Machines
McGill University

Abstract. Visual attention plays a critical role in game playing. A better understanding of the allocation of visual attention can benefit the design of game scenarios. In this paper we propose to design games in different difficulty levels based on the estimation of attention. We use eye movement information in an analysis of attention. Eye fixation and pursuit conditions are considered separately. We find that the image complexity at eye fixation points and the relative position between pursuit direction and disturbance during pursuit eye movements are significant for attention allocation. This result is applied to the design of an interactive 2D game having two difficulty levels. Player response speeds and scores are compared for each difficulty level of the game.

game design, visual attention, eye fixation, eye pursuit

1 Introduction

It is well known that computer games have become a very popular and widespread form of entertainment. Now recent studies have shown that computer games are no longer only for entertaining. Some research suggests that games may increase player's attention capacities [1]. Work done by Green and Bavelier found that frequent game players score higher on some attention tasks [2]. Special computer games are also used to help children train their brains to pay attention [3, 4] and seniors to improve their memory and attention. It is becoming promising to apply games to health therapy and training.

Undoubtedly visual attention plays a critical role during game playing. Losing games often occur due to player's late noticing or even complete unawareness of important items. A better understanding of visual attention can further improve the design of game scenarios. Studies of visual attention can inform game designers how to attract player's attention (to make games easier) and how to avoid player's attention (to make games harder) under various situations. In this way, games can be designed for different levels of players or specific usages. This brings a new view for game design.

Most current visual attention models ignore top-down influence and focus only on bottom-up information in the processing of natural scenes. The saliency model of attention proposed by Koch, Itti and colleagues used bottom-up information of scenes to indicate the possible locations of attention [5]. It is, however,

well known that scene context and task strongly influence attention [6–10, 14]. It is even more important to consider top-down factors of attention when estimating attention allocation during games, because attention is more task-relevant and goal-driven for game playing than for free-viewing of scenes.

In this paper, we propose a new way to estimate task-dependent attention allocation based on eye movement information (specifically detecting fixation and pursuit conditions). Combining such eye movement information with image information at eye fixation points, we also provide a simple way to estimate to which extent attention is engaged at a fixation point. Based on the estimation of attention allocation, we show how an interactive computer game can be designed to exhibit different difficulty levels. The game is used to verify our hypotheses concerning attention and also shows the feasibility of applying these ideas to game design.

2 Motivation

Research on eye movements show that eye movements reflect human thought processes; this led Yarbus to suggest that an observer’s thought may be followed to some extent from records of his eye movements [11]. This motivates us to take eye movement information into estimating attention during game play, and to infer top-down influences. Task dependence is a very complex issue, as there are a large number of cognitive tasks that a person can be engaged in, even in an activity as focused as game playing. To simplify the problem we will consider just two rather general classes of task - those that involve having the eye fixed at a specific location and those in which the eye is moving (visual pursuit). As such, we will examine the allocation of attention in each of these two conditions.

Tracking of moving objects with the eye is needed in many different computer games. Conversely, many game situations require the eye be fixed at a given spot, in order to extract information from that spot or to wait for an expected event. Thus, it is necessary, and rewarding, to consider attention based on the specific motion of the eye. Recent studies [12, 13] have shown that during pursuit, the allocation of attention in space is found to lead the pursuit direction (i.e. in the direction of object motion). The distance by which attention leads the eye is observed to increase as the pursuit velocity increases. Thus, if we know that the eye is undergoing a pursuit motion we can predict that the attention is most likely to be allocated at some distance ahead of the motion.

In the case of eye fixation, we consider two important aspects. The first is consideration of where the gaze tends to be directed (the fixation point) and the second is consideration of how long the gaze remains at the fixation point (fixation duration). There has been much work done regarding the location of fixation points [15–17]. But relatively little research has been aimed at understanding fixation duration during scene viewing. The prediction of the distribution of attention can be markedly different based on different prediction of fixation duration. In this paper, we will explore a way of relating fixation duration to scene complexity. The issue of fixation duration is, in our view, the problem of how

sticky attention is "glued" to a fixation point once the fixation point is chosen. The problem has been studied [18–22] and combined into models of reading [23]. But for computational attention models of scenes, it has been neglected. The saliency model of attention relates fixation duration to the saliency value at fixation points [5], incorporating features such as scene luminance [21] and contrast [22]. Fixation duration has also been found to be longer when viewing face images [24] and color photographs [20]. All these studies suggest that fixation duration is influenced by the processing of information presented at the current fixation point. But saliency only tells part of the story. Not only bottom-up but also top-down attentional factors are involved in determining fixation duration. The saliency value may give clues of where attention or the gaze is directed at the beginning of viewing a scene. But once fixation begins, for duration of the fixation, we believe that the processing of the information of image at the fixation point is more important than the raw saliency at the fixation location. In particular, we take as a simple model, that the fixation duration is related to the complexity of the image or scene at the fixation point, as a higher complexity implies that the brain requires a more detailed processing of the visual input to make sense of the scene. For the purposes of computer game design, we hypothesize that local measurements of fixation point scene complexity can be used to predict the stickiness of attention, which will affect the speed at which attention can be shifted away to new targets, and hence will affect reaction times.

Based on these two task-dependent aspects of the attention related to eye movements (i.e. fixation duration and attention during pursuit) we carried out a set of experiments. These experiments tested the effect of (in the fixation case) image complexity on reaction time, and (in the pursuit case) the relative location of targets on reaction time. The results obtained during the experiments were then applied to an interactive 2D computer game to make the game easier or harder through controlling the position of game elements based on our attention models. Consistent results were shown. The idea of designing games, taking into account of attention characteristics during different tasks was demonstrated.

3 Methods

3.1 Experiment procedure and apparatus

The purpose of the first experiment was to detect if image complexity at the fixation point affected the disengagement of attention during fixation and pursuit conditions.

The visual stimuli were generated by a computer and displayed on a screen of a 1280*1024 pixel monitor at a distance of 18 inches (1 degree of visual angle corresponded to approx 30 pixels). Five hundred color images (640*480) from different scene categories (landscape, building, city scene, indoor scene, animal, etc) were used as background. Figure 1 shows some sample images from the 500 images that were shown as background images. At the beginning of each trial, a fixation point of size (0.26*0.26 deg) appeared at the center of the monitor. The background of the monitor was set to be black at this time. After subjects

centered their fixations on the fixation point, and felt ready for the experiment, they initiated trials with a key-press. Once subjects triggered the trials, a background image was shown, centered on the display. One small green square of size (0.26*0.26 deg) started to move either rightward or leftward at a speed of 1.4 deg per sec from the center of the background. Subjects were instructed to either pursue the moving square or freely view images. Approximately 2 seconds after the start of every trial, one square object of size (0.65*0.65 deg), either in green or in blue, appeared at random positions inside of images. Subjects were required to react to the color of the square object as soon as possible by pressing corresponding mouse buttons.



Fig. 1. Examples of images used as background for the experiment. Each column represents one category of images. From left to Right: city scene, landscape, building, animal, and interior scene.

Six subjects (five males, one female) participated in the experiment. The subjects were all recent graduate students. Data were collected after informed consent was obtained. Experimental sessions lasted for approximately one and a half hours with mandatory three-minute rest periods occurring after approximately every five minutes of data collection. Subjects were given practice in performing the task before collecting experimental data. Subjects had control over when to start a trial through the pressing of a keyboard button. They were seated approximately 18 inches away from the display, and a chin rest was used to minimize head movements. An eye tracker (ISCAN RK-726PCI) was used to record the subjects left eye positions at 240HZ during experiments. Subjects used both eyes to conduct the experiments.

3.2 Data analysis of eye positions

At the beginning of every experimental session, we calibrated the eye tracker by having subjects look at a five-point calibration display. Data analysis was carried

out on every single trial. By visual inspection of the individual recordings, trials with blinks before the appearance of the flashed square object, missed pursuits, or missed executions were excluded from further analysis. For free viewing trials, trials were excluded as well if the flashed square object appeared within the same area as previous eye fixation locations. Eye position data were smoothed with a median filter (9 samples).

3.3 Entropy of images

To quantify how much information contained at a fixation point, we used the entropy of the image data in a neighborhood around fixation points. We will refer to this as the information quantity.

A segmentation of the image at the fixation point was first computed [25]. It is a process to classify each image pixel to one of the image parts, and reduces noise. After segmentation, the color pixel data were converted into graylevel. Given the human visual systems "preference" for luminance information, the graylevel values were obtained by the luminance of original images. Entropy was therefore computed for the images based on the luminance values (Y) of the segmented images. Y was obtained according to the following formula. $Y = 0.299 R + 0.587 G + 0.114 B$.

4 Results

4.1 For the free viewing condition

The effects of local image entropy in a neighborhood about the fixation area and the global image entropy over the whole image were both checked. Figure 2 shows the mean reaction time (to the flashed square object) as a function of the local image entropy at fixation area and image entropy of whole image separately. As expected, we observed the tendency that reaction time increased as the image entropy at local fixation area increased. But for the global image entropy over the whole image, no such tendency was observed. This indicates that the local entropy at the fixation area is potentially one factor related to the time of disengagement of attention at fixation points.

4.2 For the pursuit condition

The same factors were checked in the pursuit condition. In this case, however, no consistent relationship was found between the image entropy at the eye position and the reaction time. Results are shown in Figure 3. For the case when the flash appeared in front of the pursuit direction, the image entropy of the whole image showed a tendency to affect reaction time. However, this was not found for the case when the flash appeared in the opposite direction as pursuit.

Following the experiments described in [12, 13], we also checked the subject's reaction time for fixation and pursuit situations separately. For pursuit case, the

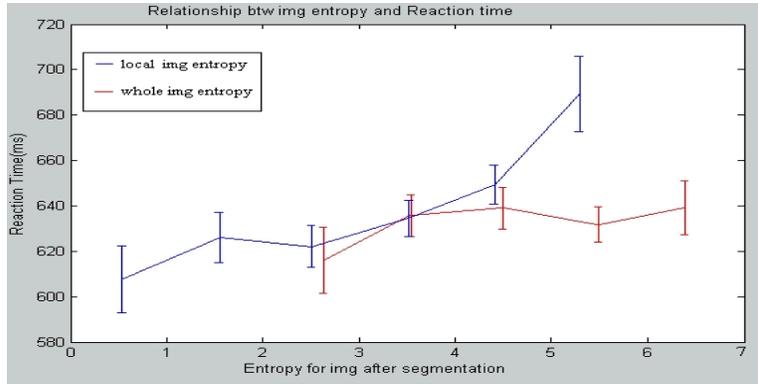


Fig. 2. Relationship between image entropy and reaction time during eye fixation. Error bars represent standard errors of the mean.

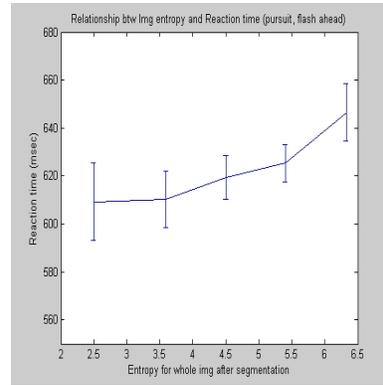
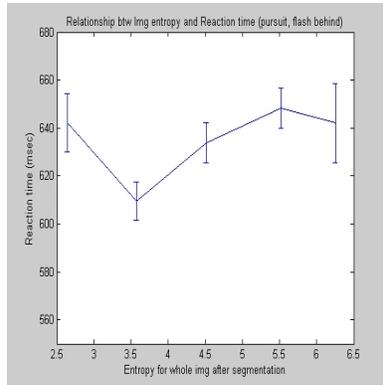
reaction time was compared under two different conditions. They are where the flashed object appears ahead of the pursuit direction and where the flashed object appears in the wake of pursuit direction. The results for the comparison of reaction time in these two conditions is shown in Figure 4. From the figure, we can observe that the reaction times for two pursuit cases are significantly different. Reaction times are close to each other for the case of fixation and the case of flash behind the pursuit direction, and significantly larger than when the flash occurs in the direction of pursuit. This result is consistent with what was found in [12, 13], where attention was observed to be biased towards the pursuit direction. The bias offset is related to pursuit velocities. The faster the pursuit velocity is, the further ahead attention tends to bias.

We see from these results that the image content at fixation points is more important in affecting attention allocation for fixation than for pursuit. In the pursuit condition, the relative position between the pursuit direction and the visual disturbance is more important than image content.

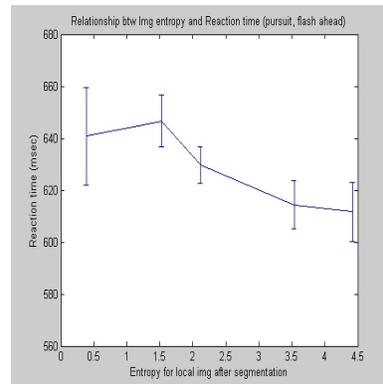
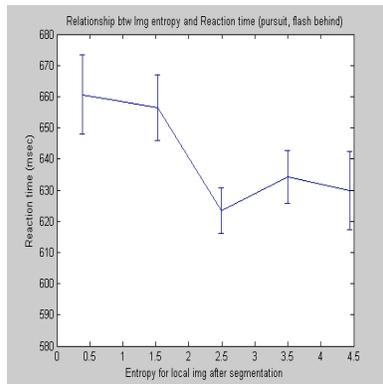
5 Game application

The results obtained from the previous experiment were applied to an interactive pc game. The game was designed as a normal pc shooter game, except that during the game, eye position information was recorded and analyzed. The eye tracker information was used to determine whether the eye was fixated or was engaged in visual pursuit. This was then used to adjust the strategy employed to present game elements.

The game was designed to exhibit two difficulty levels, hard and easy. For the hard level, enemies were designed to appear at the location of high attentional cost. For the easy levels, enemies were designed to appear at the location with



(a) Relationship between image entropy and reaction time. The flash appears behind the pursuit direction. (b) Relationship between image entropy and reaction time. The flash appears ahead of the pursuit direction.



(c) Relationship between local image entropy and reaction time. The flash appears behind the pursuit direction. (d) Relationship between local image entropy and reaction time. The flash appears ahead of the pursuit direction.

Fig. 3. Relationship between image entropy and reaction time during eye pursuit.

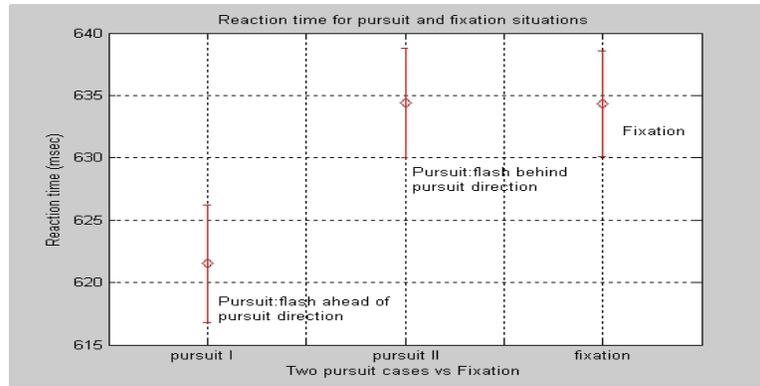


Fig. 4. Comparison of reaction time under the pursuit and fixation conditions.

attentional benefit. For the hard level, enemies were displayed when the eye was fixated at a location of image with high entropy, or were presented at a location in the direction opposite to pursuit. For the easy level, enemies were displayed when the eye was fixated at a location of low image entropy, or if the eye was engaged in visual pursuit, the enemies were displayed at a location in the same direction as pursuit.

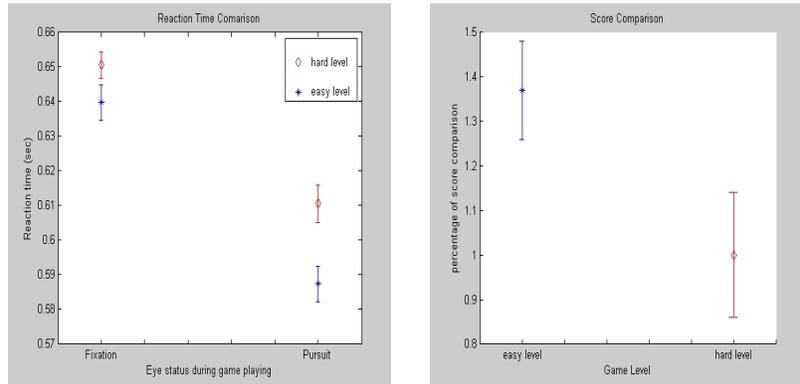
The game was designed as follows. Four types of enemies appeared during games. They were all shooters, either with bullets, bombs, fireballs, or missiles. Three shooters were displayed right from the beginning of the games. These moved horizontally and shot during moving. They stopped moving temporarily if shot 20 times by players. The player's score was increased by shooting shooters or decreased by being shot. In addition to these three shooters, other enemies would appear continuously. The type of enemies and their locations depended on the eye movement information during playing. When a bomb appeared, it would explode shortly after its appearance. To be protected from the blast, players needed to hide from it. Fireballs would always appear at the right or left bottom corner of the background. They would move horizontally either leftward or rightward, depending on their initial location. To be protected from a fireball, players needed to jump up from the ground. Scores would be deducted if caught by a bomb blast or a fireball. The missile enemy type was used to trigger eye pursuit in the player. This type of enemy flew horizontally at a speed of 4 deg per sec. Players could gain higher scores by pursuing the missile and correctly responding to a number (from 0 to 9) displayed on the missile. The number would appear randomly during the flight of the missile. Players were also allowed to shoot and move horizontally, except while hiding and jumping. The game background was set to images selected from different categories (landscape, city scene, building, animal, etc). Each image background lasted approximately 10 seconds with random numbers of appearance of enemies. A screenshot of the game during play is shown in Figure 5.



Fig. 5. One screenshot for the game. The green figures are the three shooters. The blue figure represents players. The red circles show the fixation areas with low and high entropies (low entropy corresponding to easy level, high entropy corresponding to hard level).

The same experimental environment was used in testing the game as in the previous experiment. Five subjects (three females, two males) played this game. Each session of the game lasted approximately six minutes. In total, each subject played the game for approximately three hours. Game levels were alternated randomly for each player without notifying them.

Reaction times and scores were analyzed after the experiment. Figure 6 shows the results. We observed longer reaction times for the hard game and shorter reaction time for the easy game. Statistical significance testing using a Ttest shows significant differences between the reaction times. Also the scores for the two levels of games were significantly different. Higher scores were observed for the easy game and lower scores were observed for the hard game. These indicate that the games were successfully designed into two difficulty levels based on the consideration of attention allocation.



(a) Reaction time for hard and easy levels of the games. (b) Score comparison for hard and easy levels of the games. Mean score for hard level game is taken to be 1.

Fig. 6. Comparison of RT and Scores for hard and easy games.

6 Discussion and Conclusion

Attention plays a critical role during game playing. A better understanding of attention allocation during different tasks will benefit game design. Based on a model of attention allocation, we can make a game harder by placing important game-relevant items in regions with less attention or we can make a game easier by placing important game-relevant items in regions with more attention. We applied this strategy to an interactive computer game. The test result shows that subjects responded significantly differently to items placed at different attention allocation areas. As expected, for eyes fixated in areas with lower local image

entropy, reaction to peripheral targets tends to be faster. Attentional benefit was also associated with items appearing ahead of eye pursuit movements.

Both fixation and pursuit eye motion patterns appear during game playing. Because of the different attention allocation strategies in these two conditions, only considering one type of eye movement during game design is not complete. Our test results separating different eye movement types show the significant differences of reaction time for each type of eye movement situation.

The main contribution of this study is its novel consideration of eye movement types: fixation and pursuit into game design. Although the same idea can be applied to saccade eye movement as well, our consideration is only limited to fixation and pursuit currently, because saccades occur much less than fixation and pursuit, and also once saccades occur, they only last a few tens or hundreds of milliseconds. Results acquired from the experiment were applied to an interactive 2D computer game. Consistent results were shown, which validated the previous psychophysical studies and showed the feasibility of the design idea. Our results suggest that computer game design can benefit from the study of attention, taking into account of different types of eye movements.

Ethics Approval Disclaimer The research presented in this paper involved psychophysical experimentation with human subjects. Prior to carrying out such experimentation, details of the procedures, techniques, and equipment involved in it were approved by the Ethics Review Committee of the Faculty of Education at McGill University.

References

1. Eriksen, B. A. and Eriksen, C. W. Effects of noise letters upon the identification of a target letter in nonsearch task. *Percept. Psychophys.* 16, 143V149 (1974).
2. Green, C.S and Bavelier, D. Action video game modifies visual selective attention. *Nature* 423, 534V537 (2003).
3. Rueda, M.R. and Rothbart, M. K. and McCandliss, B. D. and Saccomanno, L. and Posner, M.I. From The Cover: Training, maturation, and genetic influences on the development of executive attention. *PNAS* 2005 102: 14931-14936 (2005).
4. Dye, M.W.G. and Bavelier, D. Playing video games enhances visual attention in children. *Journal of Vision.* Vol. 4, Number 11, Abstract 40, pp. 40a (2004).
5. Itti, L and Koch, C. A saliency-based search mechanism for overt and covert shift of visual attention. *Vision Research*, 40(10-12):1489-1506, (2000).
6. Wolfe, J. and Cave, K. and Franzel, S. Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 15:419-433, (1989).
7. Oliva, A. et al. Top-down control of visual attention in object detection. *IEEE proceedings of the International Conference on Image processing (Vol. I)*, IEEE, pp. 253-256 (2003).
8. Land, M.F. and Hayhoe, M. In what ways do eye movements contribute to everyday activities? *Vision Research*, 41, 3559-3565 (2001).

9. Turano, K.A. et al. Oculomotor strategies for the direction of gaze tested with a real-world activity. *Vision Research*, 43, 333-346 (2003).
10. Henderson, J.M. et al. Effects of semantic consistency on eye movements during scene viewing. *J. Exp. Psychol. Hum. Percept. Perform.* 25, 210-228. (1999).
11. Yarbus, A.F. *Eye Movements and Vision*, New York, Plenum Press. (1967).
12. van Donkelaar, P. and Drew, A.S. The allocation of attention during smooth pursuit eye movements. *Prog. Brain Res.* 140: 267-277. (2002)
13. Jie, L. and Clark, J.J. Microsaccadic Eye movements during ocular pursuit. *Vision Sciences Society Annual Meeting (VSS)*, Vol. 5, Number 8, Abstract, pp. 697a (2005).
14. Pomplun, M. Saccadic selectivity in complex visual search displays. *Vision Research*, 46(12):1886-1900, June (2005).
15. Stiefelhagen, R. and Zhu, J. Head orientation and gaze direction in meetings. *Conference on Human Factors in Computing Systems*. Pp. 858-859. (2002).
16. Kayser, C. and Nielsen, K.J. and Logothetis, N.K. Fixations in natural scenes: interaction of image structure and image content. *Vision Res.* 46(16): 2535-2545. (2006).
17. Rajashekar, U. and Cormack, L.K. and Bovik, A.C. Image features that draw fixations. *Proc. IEEE Int. Conf. Image Proc.*, vol. 3, pp. 313-316, Sept, (2003).
18. Rayner, K. Eye movements in reading and information processing: 20 years of research. *Psychol. Bull.* 124, 372-422. (1998).
19. Henderson, J.M. et al. Effects of semantic consistency on eye movements during scene viewing. *J. Exp. Psychol. Hum. Percept. Perform.* 25, 210-228. (1999).
20. Henderson, J.M. and Hollingworth, A. Eye movements during scene viewing: an overview. In *Eye Guidance in Reading and Scene Perception* (Underwood, G., ed.), pp. 269-283. (1998).
21. Loftus, G.R. Picture perception: effects of luminance on available information and information-extraction rate. *J. Exp. Psychol. Gen.* 114, 342-356. (1985).
22. Loftus, G.R. et al. Effects of visual degradation on eye-fixation durations, perceptual processing, and long-term visual memory. In *Eye Movements and Visual Cognition: Scene Perception and Reading* (Rayner, K., ed.), pp. 203-226, Springer. (1992).
23. Reichle, E.D. et al. Toward a model of eye movement control in reading. *Psychol. Rev.* 105, 125-157. (1998).
24. Guo, K. and Mahmoodi, S. and Robertson, R. G. and Young, M. P. Longer fixation duration while viewing face images. *Experimental Brain Research*. Vol. 171. Number 1, (2006).
25. Felzenszwalb, P.F. and Huttenlocher, D. P. Efficient Graph-Based Images Segmentation. *International Journal of Computer Vision*, Vol. 59, Number 2, Sept, (2004).