Revealing Points of Attentional Interest: To Squint or Not to Squint?

Sujoy Kumar Chowdhury
MIT Media Lab
sujoy@sujoykumar.com

Jeremiah D. Still
San Jose State University
jeremiah.still@sjsu.edu

ABSTRACT
The human visual system can only attend to a limited number of regions within an interface at one time. Several methods for predicting the deployment of attention have been suggested. We describe an interest point method in which participants identify five "interesting" points in a display. Previous research establishes that this method can successfully predict attentional selection in complex scenes as reflected by eye movement data. Here we describe an experiment that examines the traditional squint test (blurring of display) and its ability to promote the detection of attentional interest. Results indicate that blurring an image disrupts the typical deployment of attention in an interface; therefore, squinting should not be used when attempting to detect interesting regions.

Keywords
interesting regions; squint test; search examination behavior

INTRODUCTION
The human visual system is limited by a working memory capacity restriction. In order for the cognitive system to function properly, attention must select a subset of the information available in the environment [5]. Attentional selection is driven by visual properties of the stimulus (bottom-up processing) and by goal-directed properties (top-down processing) [21]. Eye movements reflect the influence of both types of attention. In addition, it is generally accepted that eye movements are highly correlated with attention in that attention tends to select those things we move our eyes to. Because of this, eye movements can provide insight into attentional selection in complex visual displays like websites. In design, eye movements can be used as an evaluative tool - if a key region fails to attract the user's attention, design changes may be necessary.

After an interface is created, eye tracking systems enable designers to determine what aspects of an interface attract visual attention. But eye trackers are often criticized for being expensive [1] by those with limited resources, tedious to use [15] and for gradually losing calibration. Given these drawbacks, there has been an effort in industry and academia to find an alternative. Johansen and Hansen [8] compared the predictive power of eye tracking to that of a designer (the designer predicts where the user will look). Attentional selection was measured in this study by asking users to remember different visual aspects of a web page. Their results showed that users could remember 70% of the web elements they had seen (eye tracking results), while designers could only predict 46% of the web elements users look at. These findings highlight the limitation of expert predictions in capturing eye-movements.

If experts are unable to predict where users will look, perhaps computational models can be used instead. According to Still and Masciocchi (2010), the visual properties that drive attention within a web interface can be described by a pure, stimulus driven model – a saliency model. Saliency models predict unique regions within a scene based on low-level feature extraction [7]. However, ongoing research suggests that users' attention is driven by both top-down and bottom-up processing. Capturing the interaction between these processes is a difficult task often leaving designers to make a best guess about which regions will be attended.

Considering the inadequacy of saliency models and expert prediction for predicting user deployment of attention, one remaining alternative is to find a low-cost alternative to eye tracking. One such alternative has been to establish a relationship between eye movements and mouse movements [2, 13]. In recent years, a number of such usability tools have been developed that depend on mouse tracking. These tools can reliably report user actions, like mouse clicks, that have been associated with attentional selection (Crazy Egg, [4]). Instead of recording all mouse clicks, it has been suggested that the user's first five mouse clicks can be used as a proxy for attention if participants are instructed to select no more than five interesting points from a screenshot within five seconds (Five Second Test, [6]). The method we examine is similar to the Five Second Test.

Reference:

Chowdhury, S. K., & Still, J. D. (2014). *Revealing points of attentional interest: To squint or not to squint* (PoD Research Report No. 3). Retrived from Psychology of Design Laboratory website: http://psychofdesign.com/reports/Chowdhury_Still_2014.pdf

Figure 1. Example of an interest point plot (left) and interest point map with Gaussian intensity "blobs" (right).

Interesting Point Selection Method

Masciocchi, Mihalas, Parkhurst and Niebur [9] were the first to demonstrate experimentally that interest points selected by participants are associated with fixations within scenes. A large population (*n* = 1395) of participants from many different cultures were found to agree on what was "interesting" in the scenes (i.e., their interest points clustered). In addition, their selections correlated with the eye movements observed in a different group of participants. Masciocchi et al. showed that interest points were better predictors of eye movements than the saliency model. They suggested the better performance of interest points resulted from including both top-down and bottom-up processing. It is worth noting that the stimuli used in that study were scenes which are categorically different than interface screenshots. Importantly, subsequent investigations have demonstrated that interest points cluster and correlate with fixated locations within web page screenshots [10].

The Squint Test: Effect of blurring an interface

Our decision to contrast blurring with the interest point technique was prompted by the prevalent use of the 'squint test' in industry to evaluate if a design has succeeded in creating the intended experience for the user [12, 19, 20]. Squinting one eye with the other eye closed blurs the visual details of a display and allows the groupings of major structures to stand out – which can be an assessment of the display's "gestalt" [14]. A display's structures ought to be apparent under this blurred state; if not, it is unlikely to be perceived by users during actual use. Squint test can help manage effectively the subtle interrelationships of scale and contrast of a well-designed display [11]. Changing the perspective through squint test can uncover otherwise undetected issues related to visual hierarchy and relationships in an interface [3].

Several usability tools (e.g., Enhanced Restricted Focus Viewer [18], Stompernet Scrutinizer Foveal Gaze Simulator [17]) employ the blurring technique to simulate the traditional squint tests. When using these tools, a small portion of a blurred display is brought into focus when the computer mouse passes over it. This artificial constraint is introduced assuming that mouse-movement in the blurred interface would strongly correlate with natural eye movements in a corresponding clear interface. In the current experiment we explore the effect of blurring visual information in a webpage on the clustering of interest points. We want to see if squinting disrupts the natural guidance of attention in clear displays.

METHOD

Participants

The university institutional review board approved all experimental procedures. Twenty-six undergraduate volunteers (24 right handed, 14 female, all with normal or corrected-to-normal vision) were recruited to participate in exchange for course credit.

Stimuli and Apparatus

An Adobe Air cross-platform desktop application was created to experimentally implement the interest point method. Our application was installed on Dell Pentium 4 computers with 15 inch monitors (resolution 1024 × 768). Each participant viewed a total of 50 images. The images were screenshots containing mostly text (13), half picture and text (23), or mostly pictures (11). Four research assistants independently classified each screenshot. Only 3 of the screenshots were unclassifiable through group agreement. The opacity condition was manipulated by rendering half of the images as clear and the rest as blurred. The blurred condition was implemented by adjusting the horizontal and vertical blur parameters of the Adobe Flash blur filter by a value of five. The blur action was performed four times (configured through the quality parameter) to ensure the text was not legible. Participants were randomly

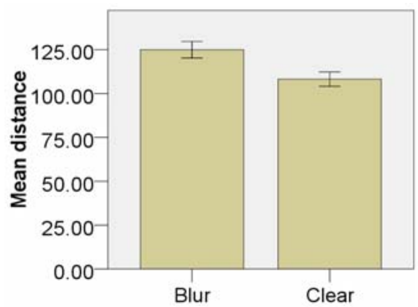

Figure 2. The mean of distances (pixels) from the average interest point.

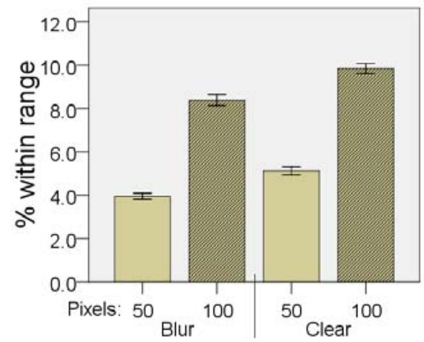

Figure 3. The solid and patterned bars represent percentage of interest-point-pairs which are separated by Euclidean distances not more than 50 pixels and 100 pixels respectively. The 100-pixel distances include the 50-pixel distances.

assigned to one of five preset random sequences of images. Responses were collected through a mouse and keyboard.

Procedure

Our experimental application asks test-participants to click on the five most interesting points in screenshots of interfaces with no time limit. Small red circles are overlaid on the images in real-time to provide visual feedback of clicked locations. These circles remain on screen until the user navigates to the next image. Participants completed a training sequence before the actual experiment to become familiar with the interface.

Data Analysis

In order to assess the effect of blurring on the selection of attention, we employed three different quantitative measurements.

The first quantitative measure determined the central tendency of the interest points. The central tendency is used to indicate how close the interest points are to one another. It was calculated by determining an average interest point for each clear or blur image by taking the arithmetic mean of the coordinates (x, y) of the five interest point selections. Then we calculated the Euclidian distances of this average interest point from the five corresponding interest points. The mean of the five distance values for each image was calculated – referred to as the Mean Distance Value (MDV). Finally, we determined the Grand Distance Value (GDV) for each participant under a particular condition by taking the mean of the 25 items presented in the same condition (e.g., clear or blurred). A high GDV indicates a low central tendency – interest points tend to be further apart - while a low GDV indicates a high central tendency – interest points tend to cluster together.

The second quantitative measure was interest point selection time. This was defined by the time elapsed between stimulus onset and the first interest point selection, or by the elapsed time between two successive selections of interest points.

The third quantitative measure examined differences in clustering to determine whether or not interest points are closer together in the blurred or the clear condition. To determine this, we first took the coordinate (x, y) of each interest point location, and determined its distance from every other interest point in the same image under the same condition. We then calculated the proportion of interest points that were separated by a given number of pixels: 50, 100, 200, 300 and 400 pixels. An additional constraint was that we considered interest points to be part of the same cluster if they were separated by less than 50 pixels (c.f. [9]).

RESULTS

All statistical tests used an alpha level of 0.01. Error bars in the figures represent the mean standard error.

Subjective Visualization

We generated an interest point plot for a screenshot using the interest-point selections across all users (see Figure 1: left image). Then we constructed an interest point map for the same screenshot by superimposing yellow Gaussian intensity "blobs" centered over each interest point (see Figure 1: right image).

Central Tendency

A paired-samples t-test revealed that the Grand Distance Values (GDV) for the blurred (M = 124.93, SD = 23.96) condition were significantly higher than those for the clear (M = 108.16, SD = 20.94) condition, $t(25) = 4.65$, $p < 0.001$. This indicates that interest points in blurred images are distributed farther apart than in clear images (see Figure 2).

Selection Time

A paired-samples t-test revealed that interest point selection times for blurred (M = 2348.62, SD = 2686.18) images were significantly slower than for clear (M = 1747.56, SD = 1646.68) images, $t(3249) = 11.48$, $p < 0.001$. This increased cognitive processing time for blurred images suggests that participants were required to complete more cognitive cycles to determine interest points than for clear images.

Clustering Strength

A paired-samples t-test revealed that there were significantly fewer interest point pairs located within 50 pixels of one another in the blurred (M = 3.95, SD = 1.01) condition than the clear (M = 5.13, SD = 1.32) condition, $t(49) = 6.04$, $p < 0.001$. A similar pattern was found when 100 pixels were used as the maximum distance parameter for a cluster. Significantly fewer interest point pairs were found within 100 pixels of one another in the blurred (M = 8.39, SD = 1.80) condition than in the clear (M = 9.85, SD = 1.68) condition, $t(49)= 4.92$, $p < 0.001$. These findings suggest that clustering strength is lower under blurred conditions (see Figure 3).

CONCLUSION

Our results provide converging evidence that interest point selections under squint conditions differ significantly from those made under normal conditions, therefore, the squint test should not be used in conjunction with the interest point method. Blurring a display disrupts the natural course of a user's attentional processes.

Squinting to view an image [12, 19, 20], or blurring an image [17, 18], makes text unreadable and removes other important high frequency spatial information used by the attentional system to select visually interesting locations. It is commonly recognized in the vision literature that high frequency contrast plays an important role in the programming of eye movements [21]. As the data demonstrate, removing high frequency information from the image results in "noisier" selection by the attentional system. Therefore, we recommend that only clear stimuli be used for the interest point method when the goal is to predict the deployment of visual attention.

Blurring an interface may be a useful technique for identifying global features, but not for identifying attentional interest.

REFERENCES

1. Aaltonen, A. Eye tracking in usability testing: Is it worthwhile? *Usability & Eye Tracking Workshop in CHI 1999*, ACM Press (1999).
2. Chen, M.C., Anderson, J.R., and Sohn, M.H. What can a mouse cursor tell us more?: Correlation of eye/mouse movements on web browsing. *Ext. Abstracts CHI*, ACM (2001), 281-282.
3. Cooper, A., Reimann, R., and Cronin, D. *About Face 3: The essentials of interaction design*. Wiley (2007).
4. Crazy Egg Click Tracker and Heat-map Builder. http://www.crazyegg.com.
5. Desimone, R. and Duncan, J. Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, (1995), 18: 193-222.
6. Five Second Test. http://www.fivesecondtest.com.
7. Itti, L., Koch, C. and Niebur, E. A model of saliency based fast visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (1998), 20(11): 1254-1259.
8. Johansen, S.A. and Hansen, J.P. Do we need eye trackers to tell where people look? *Ext. Abstracts CHI*, ACM (2006), 923-928.
9. Masciocchi, C.M., Mihalas, S., Parkhurst, D., and Niebur, E. Everyone knows what is interesting: Salient locations which should be fixated. *Journal of Vision*, (2009), 9(11): 1-22.
10. Masciocchi, C. M. and Still, J. D. Alternatives to eye tracking for predicting stimulus-driven attentional selection within interfaces. *Journal of Human-Computer Interaction*, (in press).
11. Mullet, K.E. Designing visual interfaces. *Conference companion on Human factors in computing systems common ground in CHI*, ACM (1996), 332-333.
12. Mullet, K. and Sano, D. *Designing visual interfaces: communication oriented techniques*. SunSoft Press (1995).
13. Rodden, K., Fu, X., Aula, A., and Spiro, I. Eye-mouse coordination patterns on web search results pages. *Ext. Abstracts CHI*, ACM (2008), 2997-3002.
14. Rosson, M.B. and Carroll, J.M. *Usability engineering: scenario-based development of human-computer interaction*. Morgan Kaufmann (2002).
15. Schnipke, S.K. and Todd, M.W. Trials and tribulations of using an eye-tracking system. *Ext. Abstracts CHI*, ACM (2000), 273-274.
16. Still, J. D. and Masciocchi, C. M. A saliency model predicts fixations in web interfaces. *5th International Workshop on Model-Driven Development of Advanced User Interfaces 2010*.
17. Stompernet Scrutinizer Foveal Gaze Simulator. http://about.stompernet.com/scrutinizer.
18. Tarasewich, P., Pomplun, M., Fillion, S., and Broberg, D. The enhanced restricted focus viewer. *International Journal of Human-Computer Interaction*, (2005), 19: 35-54.
19. Watzman, S. The art of the interface. *Conference summary CHI*, ACM (1998), 151-152.
20. Watzman, S. Visual design principles for usable interfaces. In Sears, A. and Jacko, J. A. (Eds.), *The Handbook of Human-Computer Interaction*. CRC Press (2007), 352-353.
21. Wolfe, J. M. Guided search 4.0: Current progress with a model of visual search. In W. Gray (Ed.), *Integrated Models of Cognitive Systems*, Oxford Press, (2007), 99-119.