

## Directing Attention in Open Scenes

Stephen Hughes and Michael Lewis  
University of Pittsburgh  
Pittsburgh, PA 15260  
{shughes, ml}@mail.sis.pitt.edu

Attention direction is particularly difficult in open scenes such as detailed aerial photographs, wide angle/multiple video feeds, or high-resolution digital imagery. Unlike conventional displays that have single parameters with definable set points, open scenes lack easy criteria for directing attention. The present study investigates the effects of varying levels of highlighting for correctly and incorrectly designated targets. The results suggest that linking the intensity of highlighting to degree of confidence may be an effective display strategy to maximize efforts to draw attention to objects in an open scene.

### INTRODUCTION

Human attention is generally the most limited resource in any human-machine system. From annunciator systems in nuclear power plants to highlighted links on webpages, the prime task of the designer is directing the user's attention to the right places.

According to feature integration theory, important symbols can be designed to "pop-out" from a field of distractor objects due to pre-attentive processing. (Triesman 1986). This theory has spawned numerous follow-up studies to attempt to isolate features that are pre-attentive (Ware 2000). Unlike artificial displays that have single parameters with definable set points, open scenes lack easy criteria for directing attention. Capturing Attention is particularly difficult in open scenes such as detailed aerial photographs, wide angle/multiple video feeds, or high-resolution digital imagery.

Moreover, open scenes tend to have more ambiguity since information is compiled from sensors rather than synthetically fabricated. If uncertain data from automatic target recognition (ATR) or collateral information from a source such as spatially correlated radio emissions is used, attention direction should be tentative rather than misleadingly absolute.

Superimposed icons can be used for directing attention in open scenes; symbology for identified targets on a radar display, for example. Targets that exceed the threshold appear on the display as icons while putative below threshold targets remain invisible.

While better than leaving the viewer adrift, this approach is misleading in two ways: 1) icons may indicate unreal targets and 2) real targets may be missed if evidence does not exceed the threshold. At different points along the ROC curve very different pictures of the scene may be conveyed. At the conservative end the scene may appear devoid of icons while at the other extreme it swarms with targets. To present either of these views to an analyst or some one performing a real time control task could be misleading. Iconic presentations also limit the ability of a human observer to bring more complex inferences to bear on target identification. If, for example, the observer knew that tanks tended to cluster at river crossings as a ford was being prepared, multiple uncertain target identifications on a river bank could be interpreted as strong evidence for the presence of tanks. This value added scene processing is largely lost in an iconic presentation.

Another common strategy is to superimpose annotations such as arrows or circles to indicate the presence of targets in a scene. Photographs showing probable missile sites from the Cuban missile crisis provide a good example of this sort of attention direction. Unlike icons, annotations avoid obscuring the evidence to which they draw attention. With access to the raw imagery viewers can make their own judgments about targets. Annotations, however, suffer from the same problem of dichotomizing marked targets from an unmarked scene where the evidence may be more equivocal (Figure 1).

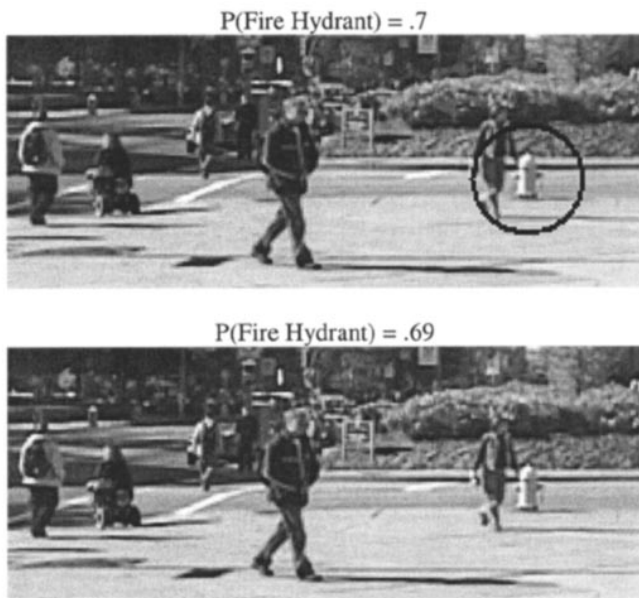


Figure 1: Annotations are dichotomous

Additionally, the marks, themselves, can add complexity and confusion to a scene and may obscure nearby or overlapping features (Figure 2).



Figure 2: Overlapping Annotations

Highlighting, varying the brightness of the target relative to the background, is a third commonly used strategy. Highlighting is often used in the popular media to designate an individual in a group photograph. Highlighting has the advantage of obscuring neither the target nor nearby regions of the image. Because highlighting can be varied continuously it allows attention to be directed by degree rather than requiring a dichotomous decision. Zhai et al (1997) experimented with these types of techniques for training people to locate controls on complex GUI interfaces. As described above however, GUI interfaces offer a highly controlled display. Despite its advantages and widespread use for open scenes, we have not been able to locate any systematic study of highlighting techniques or parameters influencing its effectiveness.

## Continuous Scale Feature Highlighting

The present study investigates the effects of varying levels of highlighting for correctly and incorrectly designated targets. The purpose of the study was to investigate the feasibility of using multiple levels of highlighting to differentially direct attention in open scenes. We envision applications such as real time situation assessment from hyperspectral imagery or medical diagnoses based on high resolution scans in which there are many putative targets of varying certainty and priority. By blending this inferential information with actual imagery we hope attention might be directed more efficiently through the scene. Our initial study is limited to investigating identification latency for single targets within a scene to evaluate the effects of highlighting level under controlled conditions.

## METHOD

### Participants

Participants were 20 undergraduate students enrolled in Information Science courses. Participants received extra credit for taking part in this study.

### Design

Subjects were asked to locate a target in computer presented scenes. Each trial contained either an orange circle annotation, one of three levels of highlighting (Figure 3), or a no highlight control. There were 24 trials in each condition for a total of 120 trials per subject. On 20% of the trials the target was not present. On 20% of the trials on which a target was present the designation was misaligned. The treatment for each image, as well as the presentation order, was randomized to control for order effects.

### Apparatus

Scenes were digital images taken using a 1.6 megapixel camera and presented on an XGA resolution 21" monitor. The highlight was obtained by separating a circular region of from the rest of the image. The size of the region varied slightly to entirely surround the target object. The brightness of the region was increased, while the remaining background was decreased to achieve a desired level of contrast. The three levels of highlighting correspond to contrasts of: low = 15, mod = 30, and high = 45. The annotation condition surrounded the target area with a 2-pixel wide orange border.

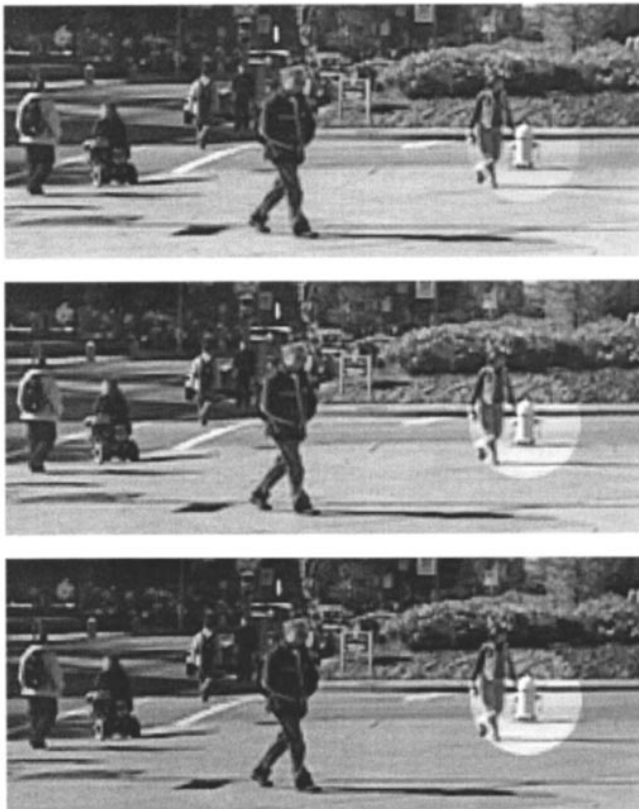


Figure 3: Degrees of highlighting

**RESULTS**

An initial signal detection analysis of the data was performed to assess the sensitivity of the participants. A response was recorded as a “hit” if the viewer clicked within highlight radius of the target object. False alarms occurred when the viewer failed to indicate that the target object was not present. No significant differences in performance were found between individuals or between the levels of highlighting. The average detectability score ( $d'$ ) was 2.64. From this we can conclude that the participants followed the instructions to seek the targets as opposed to blindly clicking on the location of the highlights and annotations. Regardless of the high accuracy rate of the participants, some errors were committed. However, given their uniformity and since there was no way to understand the reasons for these mistakes, these trials were removed from subsequent analyses.

A significant difference ( $t(509) = 23.13 p < .01$ ) between trials in which a target was present and those in which it was absent was found for response times across levels of highlighting (Figure 4). This result was expected because on target-absent trials the participant must search the entire scene to verify that no target is present while in target-present trials the search terminates as soon as the target is detected.

**Procedure**

A brief instructional session introduced the participants to the task of searching pictures for target objects. They were told that the target might be either present or absent but multiple targets would not be presented. They were shown examples of annotation and highlighting and told that while designations were usually accurate they could be in error on some trials. Participants were instructed to search the scene as rapidly as they could and respond by either clicking on the target object’s location in the image or pressing a button indicating that the object was not present. Each trial was preceded by a text message designating the target for the following trial. After a mandatory two second delay, the participant then clicked on a button to initiate the trial. Data were collected to track the location indicated by the viewer as well as the response time between the display of the image and the identification by the participant.

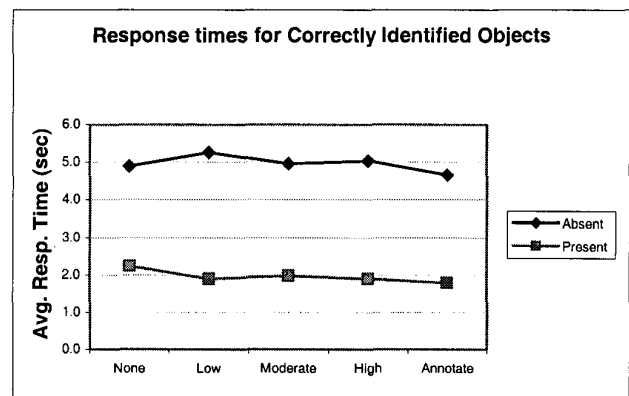


Figure 4: Serial exhaustive search effect

The interaction between level of highlighting and accuracy of designation was the central focus of this research. As Figure 5 shows, prominent highlighting leads to shorter search times when the highlight is aligned with the target and longer search times when it is not. The result agrees with the intuition that the degree of highlighting affects the overall conspicuity of the affected region and the greater the attraction to a highlight containing the target, the shorter the search. Conversely, when a nontarget is highlighted with high intensity it draws unwarranted attention and interferes with the search for the unhighlighted target.

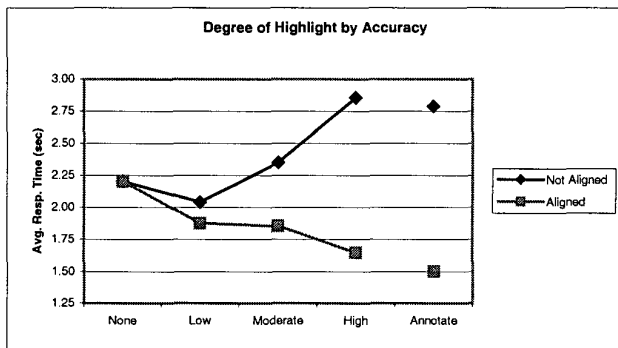


Figure 5: Accuracy by Highlight Interaction

## DISCUSSION

The benefits of using a high-contrast highlight are clear, however the results above indicate that a steep penalty must be paid if non-targets are highlighted. To explore the cost of this distraction, we calculated a weighted average for the response time under different probabilities of accurate highlights. Figure 6 shows the results of that analysis. As the figure shows, there is a substantial benefit to pronounced highlighting at a high probability of accuracy (100%); clearly if one has identified a target without doubt, the response time can be optimized by using a highlight with a substantial contrast. The appropriate level of highlighting rapidly falls off to neutral at 80% accuracy. Interestingly, it would appear to be preferable to have *no* highlight than a strong highlight if the confidence drops below a certain level.

As discussed in the introduction, the use of annotations or overlaid icons has drawbacks that make them undesirable in certain conditions. We observe however, that the strong highlight was roughly equivalent to the annotation condition for the purposes of drawing attention. This suggests that using a strong highlight might be able to effectively replace the use of annotation, surmounting its shortcomings.

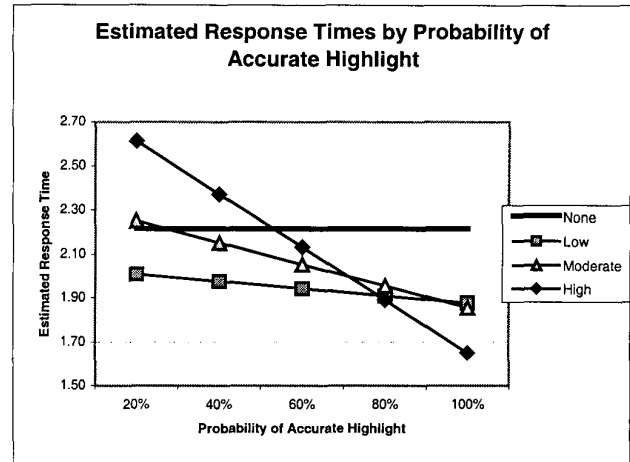


Figure 6: Expected Variations

Although we raised the idea of using highlights to simultaneously mark multiple targets, this study directly address this issue. Further work is planned to assess how multiple highlights will factor into a viewer's ability to identify targets.

## ACKNOWLEDGEMENTS

This research was supported by AFOSR contract number F49640-01-1-0542.

## REFERENCES

- Triesman, A. (1986). Features and Objects in Visual Processing. *Scientific American*: 114-125.
- Ware, C. (2000). *Information Visualization: Perception for Design*. San Francisco, Morgan Kaufmann.
- Zhai, S., J. Wright, T. Selker and S. Kelin (1997). *Graphical Means of Directing Users' Attention in the Visual Interface*. *Interact '97*. 59-66