Towards a Flexible, Reusable Model for Predicting Eye Movements During Visual Search of Computer Screens: Investigating the Effects of Grouping

Tim Halverson (thalvers@cs.uoregon.edu)
Department of Computer and Information Science
1202 University of Oregon, Eugene, OR 97403-1202 USA

Introduction

Visual search is an integral and complicated component of most human-computer interaction (HCI) tasks. Further, interface design guidelines exist that affect visual search without direct or sufficient empirical support. For example, Mullet and Sano (1995) recommend using background color to visually organize layouts instead of bounding boxes (rectangles that enclose a group of items), even though little or no evidence exists to suggest that one aids visual tasks more than the other. Generally, experienced interface designers have proposed and refined such guidelines, so the guidelines are assumed to provide meaningful guidance. Visual search research can provide generalizable support and additional insight into the guidelines.

Computational cognitive models have been used to better understand and predict how design decisions affect human-computer visual interaction. For example, the effects of icon complexity and spacing (Everett & Byrne, 2004), hierarchical organization (Hornof, 2004), and the relevance of link labels (Brumby & Howes, 2004) have been investigated to better understand how these factors affect visual search. However, most existing cognitive models of visual search account for only one or two visual factors. Considering the large variety of visual layouts users can encounter, many strategies can be employed during visual search of human-computer interfaces.

This dissertation will investigate the effects of grouping, both visually and semantically, on visual search strategies. Specifically, this research will focus on the visual search of menu-like layouts. This research is a significant step in the direction of (a) extending the range of visual layouts for which cognitive modeling can be used to predict visual search, (b) building theory of visual search for HCI, and (c) informing guidelines for the design of computer interfaces.

Visual and Semantic Grouping

One set of interface design guidelines that have little or no direct empirical support is the use of Gestalt principles of perceptual organization. Clearly some types of visual grouping do affect visual search. For example, if people know the color of the target, grouping items in the display by color has a significant affect on visual search (Bundesen & Pedersen, 1983). An easily identifiable target is found more quickly and accurately when non-targets are grouped by similarity (Gilmore, Tobias, & Royer, 1985). But it is not as clear how other types of grouping affect visual search. Faraday's web critiquing tool (2000) predicts that people will search items grouped by common regions (defined by background color) before moving on to another region. However, subsequent research by Faraday found that proximity has a stronger affect on eye movements than does common region. More research is required to determine the effects of common region on visual search.

While visual properties clearly affect visual search, the semantic content of the objects (or groups of objects) searched can also have a substantial impact. For example, Brumby and Howes (2004) investigated the effects of the semantic distance between target words and distractors relative to a goal in menu search. They found that people tend to search fewer items when distractor menu items are less similar to the goal and when the target is more similar to the goal. Further, people tend to revisit a smaller and smaller subset of menu items as visual search progresses. An ACT-R model was constructed to explain the findings. The human data were best predicted by a strategy that adapts visual search behavior based on the semantic assessment of previously perceived menu items. Still, little research has addressed the effects of semantic cohesion between items *grouped together* in visual interfaces.

The goal of this dissertation is to investigate the effects of visual and semantic grouping on visual search. This will be addressed both with human subject experimentation and cognitive modeling. Progress has been made on both fronts.

Current Progress

The majority of the work completed so far falls into two categories: refining a "minimal model" of visual search, and collecting and analyzing data from an experiment designed to investigate the effects of perceptual and semantic grouping.

A "minimal model" of visual search was built using the EPIC cognitive architecture (Halverson & Hornof, to appear). This modeling endeavor refined and rounded out previously reported cognitive modeling and eye tracking analysis. The new model accounts for a variety of eye movement data, from fixation duration to the most common scanpaths. The model does so primarily by employing three straightforward characteristics, motivated by eye movement data and previous research: (a) Eye movements tend to go to nearby objects, (b) fixated objects are not always identified, and (c) eye movements start after the fixated objects are identified. This minimal model accounts for the observed visual search behavior better than a previous model of the same task that was not informed by eye movement data.

An experiment has been conducted to investigate the effects of common region and semantic cohesion on the

visual search of groups of words in visual layouts. A total of six structured layouts were used. Three variables were manipulated: the semantic cohesion of groups of words, the presence of group labels, and the use of background color. All layouts contained eight groups randomly arranged in a 5x3 grid; each group contained five words. Groups of words were semantically related (e.g. shirt, pants, socks, jacket, and skirt) or randomly grouped. Groups were labeled (e.g. clothing) or not. Background color divided the groups into four regions or not. When the common regions were used in the semantically-grouped layouts, groups in the same region were semantically related (e.g. clothing and cosmetics). The only levels of variables not combined were randomly grouped words and labels.

Search time and eye movement data were analyzed using a mixed-model ANOVA. All results discussed are from reaction time; similar trends were observed in the eye movement data. In short, two factors overwhelmed the others: semantics and proximity. There was a main effect of layout ($p < .0001$). The contrast analysis of within-group semantics was the only significant effect ($p < .0001$); the participants searched layouts with semantically-cohesive groups faster. Additionally, people tended to move to nearby words, as indicated by a significant effect of the distance between where search started and the target, ($p < .0001$). Interestingly, there was no effect of the group labeling. People exhibited similar behavior whether labels were present or not. The scanpaths of the participants seem to indicate that, in the semantically cohesive layouts, people would effectively use any word in a group as a "substitute" for a group label.

The data from the experiment will work well for validating and identifying areas for improvement in our current model of visual search. The model relies heavily on proximity for predictions. The data observed in the above experiment suggests that our model will have some predictive power here. However, the model does not currently consider semantic content. A challenge will be to find what the current model predicts well and to what extent it must be modified to account for the effects of semantics.

Future Challenges

A few challenges remain for this line of research. These include validating our current model of visual search and extending the model to predict the effects of semantic grouping on visual search.

The experiment reported above provides data that will be used to validate and improve our current model of visual search. The visual structure and content of the layouts are substantially different from the layouts for which the model currently predicts visual search. Perhaps most importantly, the new task relies heavily on the semantic content in the layout, which the current model does not account for.

Predicting the effects of semantics on visual search will be a significant challenge for two reasons. EPIC does not currently include any theory of semantics or retrieval of semantic information from long-term memory. Therefore, choosing and integrating theories of semantic perception into EPIC that are consistent with the theory already instantiated in EPIC will be a challenge. Previous research, like that of Brumby and Howes (2004) will provide guidance for the current research and possibly provide an opportunity for integrating visual search models across cognitive architectures. Additionally, the integration of an automatic semantic information tool is needed to provide *a priori* predictions of semantic similarity for the model. Investigation of the appropriate tools for determining semantic information (e.g. like the tool used in Budiu, Pirolli, & Fleetwood, 2006) is underway.

Conclusion

Visual search is an important component of HCI. Computational cognitive modeling is an effective means of expanding the theory in HCI and ultimately will provide a means of predicting user visual search behavior to aid in the evaluation of user interfaces. This research is investigating the effects of visual and semantic grouping of visual layouts to extend the breadth of visual layouts for which cognitive modeling can be used to predict visual search behavior, to build theory of visual search for HCI, and to inform guidelines for the design of computer interfaces.

Acknowledgments

I would like to thank my advisor, Dr. Anthony J. Hornof. This research is supported by the National Science Foundation and the Office of Naval Research.

References

Brumby, D. P., & Howes, A. (2004). Good Enough but I'll Just Check: Web-page Search as Attentional Refocusing. Proceedings of the 6th International Conference on Cognitive Modeling, Pittsburg, PA.

Budiu, R., Pirolli, P., & Fleetwood, M. (2006). Navigation in Degree of Interest Trees. Proceedings of the Advanced Visual Interfaces Conference, Venice, Italy.

Bundesen, C., & Pedersen, L. F. (1983). Color segregation and visual search. Perception & Psychophysics, 33(5), 487-493.

Everett, S. P., & Byrne, M. D. (2004). Unintended Effects: Varying Icon Spacing Changes Users' Visual Search Strategy. Proceedings of the Conference on Human Factors in Computing Systems, Vienna, Austria.

Faraday, P. (2000). Visually Critiquing Web Pages. 6th Conference on Human Factors & the Web, from http://www.tri.sbc.com/hfweb/faraday/FARADAY.HTM

Gilmore, G. C., Tobias, T. R., & Royer, F. L. (1985). Aging and similarity grouping in visual search. Journal of Gerontology, 40(5), 586-592.

Halverson, T., & Hornof, A. J. (to appear). A minimal model of for predicting visual search in human-computer interaction. Conference on Human Factors in Computing Systems, San Jose, CA.

Hornof, A. J. (2004). Cognitive Strategies for the Visual Search of Hierarchical Computer Displays. Human-Computer Interaction, 19(3), 183-223.

Mullet, K., & Sano, D. (1995). Designing Visual Interfaces: Communication Oriented Techniques. Englewood Cliffs, New Jersey: Prentice Hall PTR.