Int. J. Human-Computer Studies 80 (2015) 45–55

Contents lists available at ScienceDirect

Int. J. Human-Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs

The adaptation of visual search to utility, ecology and design ☆

Yuan-Chi Tseng [a,*], Andrew Howes [b]

[a] Department of Industrial Design, National Cheng Kung University, 1 University Road, Tainan 701, Taiwan
[b] School of Computer Science, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom

ARTICLE INFO

Article history:
Received 15 April 2014
Received in revised form
20 March 2015
Accepted 30 March 2015
Available online 8 April 2015

Keywords:
Image search
Strategic adaptation
Utility maximization
Ecology
Information design
Eye movements

ABSTRACT

An important question for Human–Computer Interaction is to understand the visual search strategies that people use to scan the results of a search engine and find the information relevant to their current task. Design proposals that support this task include space-filling thumbnails, faceted browsers, and textually enhanced thumbnails, amongst others. We argue that understanding the trade-offs in this space might be informed by a deep understanding of the visual search strategies that people choose given the constraints imposed by the natural ecology of images on the web, the human visual system, and the task demands. In the current paper we report, and empirically evaluate, a computational model of the strategies that people choose in response to these constraints. The model builds on previous insights concerning the human visual system and the adaptive nature of visual search. The results show that strategic parameters, including the number of features to look for, the evaluation-stopping rule, the gaze duration and the number of fixations are explained by the proposed computational model.

1. Introduction

The question of how to design interfaces so as to facilitate search for information on the web is an important challenge for Human–Computer Interaction researchers and designers (Cutrell and Guan, 2007; Klöckner et al. 2004; Rele and Duchowski, 2005; Russell-Rose and Tate, 2013; Tseng and Howes, 2008). For example, many alternative designs for the standard list of search engine results have been proposed. They include Space-filling thumbnails (Cockburn et al. 2006), Tabular interface (Resnick et al. 2001), Faceted category interfaces (Yee et al. 2003), and Textually-enhanced thumbnails (Woodruff et al. 2001), amongst others (Bederson, 2001; de Bruijn and Spence, 2000; Fertig et al. 1996; Öquist and Goldstein; Snavely et al. 2006; Walter et al. 2006). The number of proposals is, in part, a reflection of the scale of the design space, and in part, a reflection of how pervasively search engines are used for a range of everyday tasks.

Identifying which of the potentially hundreds of interesting points in this space is best might be informed by empirical usability testing that directly contrasts one design to another. However, while an empirical basis to any design work is essential, such an approach, used exclusively, carries the danger of leaving mysterious the underlying interactive processes that lead to one advantage, or the other, and may be unlikely to lead to rapid convergence on good designs. For example, some designs will be better or worse in different circumstances and explaining the differences demands theory. An alternative approach is to rely on design guidelines. For example, findings concerning the function of human mind, using conventional laboratory tasks, have implied user-centered design guidelines for more efficient use (Shneiderman, 1992). However, it is possible that guidelines may have similar problems to pairwise usability tests if they encourage a somewhat shallow understanding of the task and the constraints imposed by the design. This is one of a set of known problems with design guidelines (e.g. see Introduction in Johnson, 2010).

Empirical studies and guidelines can be complemented with cognitive modeling, with the potential advantage of developing a deeper understanding of how and why one design is better than another. Many cognitive models of visual search have been proposed. For example, integrated ACT-R and EPIC models of the cognitive, motor, and perceptual processing required to achieve visual search tasks provide one approach (Anderson et al. 2004; Halverson and Hornof, 2011; Kieras and Meyer, 1997; Meyer and Kieras, 1997a, 1997b). One key idea to emerge from this literature is that it is difficult to ascertain whether one design is better than another unless a detailed analysis is conducted of how the user's strategy changes with the task demands (Charman and Howes, 2003; Eng et al., 2006; Howes et al. 2009; Howes et al. 2007; Kieras and Meyer, 2000; Payne and Howes, 2013; Payne et al. 2001; Pirolli, 2007). The purpose of the current article is therefore to report and test a computational model of visual search in which the search strategy is adjusted to the constraints imposed by (a) interface design, (b) the human visual system and (c) the

☆ This paper has been recommended for acceptance by Henrik Iskov Christensen.
* Corresponding author. Tel.: 886 6 275 7575 × 54335.
E-mail addresses: yctseng@mail.ncku.edu.tw (Y.-C. Tseng), HowesA@bham.ac.uk (A. Howes).

http://dx.doi.org/10.1016/j.ijhcs.2015.03.005
1071-5819/

priorities of the user, particularly priorities concerning time costs relative to the quality of the acquired results. The model explains how choices that people make about, for example, gaze duration, number of fixations, and which images to look at and which to select, are a consequence of ecological distributions of relevance, the diminishing acuity of the visual system with eccentricity from the fovea, and the priorities of the user.

In the following section we review the background to the problem addressed in the current article (Section 2). Subsequently, we described an experimental task environment that models the naturalistic task environment of images on the web (Section 3) and then a model of human behavior in this environment (Section 4), followed by an empirical investigation (Section 5). The results are reported in Section 6 and discussed in Section 7.

2. Background

2.1. The effect of design on strategy choice

In an effort to understand how to build better interfaces, researchers in HCI have suggested that the details of interface design affect visual search and overt attention strategies (e.g., Everett and Byrne, 2004; Halverson and Hornof, 2004, 2011; Pirolli et al. 2003; Tseng and Howes, 2008). Everett and Byrne (2004), for example, showed that a small difference of 1.6 degrees of visual angle between items can result in participants either fixating on an icon or not. Similarly, Halverson and Hornof (2004) provided evidence that low density, task-meaningless large font words could lead participants to use fewer and shorter fixations and so shorter overall search time than when given high density and small words. Presumably, when items are more closely packed together then more use can be made of peripheral vision. However, the pattern of findings is complex. In contrast to the previous results, Pirolli et al. (2003), for example, found that participants used more but shorter fixations when using a Hyperbolic browser (please see Lamping and Rao, 1996) than when using a standard browser, especially in areas of the Hyperbolic browser in which small size and low information scent items were grouped closely together.

There is further evidence showing that visual search strategy is adapted to the demands imposed by task environments, particularly the density of items on the display (Bertera and Rayner, 2000; Näsänen et al. ,2001; Ojanpää et al. 2002; Vlaskamp and Hooge, 2006; Vlaskamp et al. ,2005). For example, Ojanpää et al. (2002) found that decreased spacing in a vertical list of words (common Finnish verbs, nouns and adjectives) resulted in longer but fewer fixations. Longer fixations enable more information to be gathered from fovea and peripheral vision, although longer fixations can only be effective if the information is available within the perceptual span. Vlaskamp et al. (2005) found that the fixation duration, number of fixations, and search time increased dramatically with decreasing item spacing, as the range of spacing became smaller than 1.5° visual angle. On the other hand, their data showed that at wide spacing range between 1.5° and 7.1° fixation duration, number of fixations, and search time increased slightly as the spacing increased. Bertera and Rayner (2000) found that as the item spacing increased, the number of fixations and fixation duration also increased. These results indicate that people need to manage the trade-off between the increased information gain of longer fixations and the effort and time cost of holding a fixation.

The spacing of fixations is also known to change during the course of a search (Over et al. ,2007; Rao et al. ,2002). Over et al. (2007) found that fixation duration increased and the amplitude of saccade decreased gradually as search progressed. They called this a coarse-to-fine strategy. Rao et al. (2002) used a coarse-to-fine matching mechanism to model the skipping saccades because it could increase the probability of an early match. In contrast, Brumby and Howes (2008) found a fine-to-coarse search strategy. People increased saccade amplitude once they had found, but not committed to, a highly relevant target.

Although allowing a high degree of experimental control, many tasks used in vision science lack ecological validity (Bertera and Rayner, 2000; Ojanpää et al., 2002; Vlaskamp et al., 2005). For example, Vlaskamp et al. (2005) used abstract shapes (e.g., squares) in their search task, and, Bertera and Rayner (2000) used an unstructured alphanumeric array. Ojanpää et al. (2002) used common-words, which reduced the task to a simple visual pattern match, rather than a match of information relevance but it is known that search behavior is contingent on label relevance (Brumby and Howes, 2008). The different materials may account for the different effects. Both tasks are far from an ecologically valid HCI task in which the stimuli are more heterogeneous and complicated.

The task of Brumby and Howes (2008) had high ecological validity but the task, involving words, required limited graphical information processing. To explore graphical information processing in a task environment involving search for pictures, Tseng and Howes (2008) used real photo thumbnails and the real display to simulate the pages of thumbnails returned by a search engine and found that the number of alternatives in a search set and the density of the display influence how people make small but significant changes to eye-movement strategy. For example, as item density changed their participants adjusted the duration that they attended to each item. There was a negative correlation between the number of items and the gaze duration. Longer gazes were only used when they were efficient, i.e. when expected information gain was high. Also, longer visits to items were combined with skipping. Tseng and Howes (2008) found that participants were observed to reduce the number of items that they visited, i.e. they skipped, when there was a larger number of alternatives in a search set. These findings support the view that people adjusted their visual search strategy to their expectations of information gain, and that these expectations were contingent on (a) the density of items, and (b) the prior likelihood that an item is the one that they will want to select.

2.2. Visual information processing constraints modulate strategy choice

There is a substantial body of literature showing how the constraints of the visual information processing system modulate strategy choice. Geisler (2011) reviews work that shows how visual search strategies are adapted to the biological constraints of the human visual system, such as the optics of the eye, the spatial and chromatic sampling by the photoreceptors, photon noise, and retinal spatial summation, all of which provide noisy neural representations of the stimulus.

Some of these constraints are embedded in an active vision model (Halverson and Hornof, 2011; Kieras and Hornof, 2014). By using the EPIC cognitive architecture (Kieras and Meyer, 1997) to model visual search behavior, they attempt to explain the relationship between strategies and the underlying architectural mechanisms. These models provide answers to the four questions of active vision: when do the eyes move? What can be perceived? Where do the eyes move to? What information is integrated between eye movements? In this model (Halverson and Hornof, 2011; Kieras and Hornof, 2014), they emphasized the importance of the constraints of the visual information processing system, particularly the spatial resolution of the retina which decreases as eccentricity increases (also see Geisler, 2011). Therefore, the saccade mechanism is required to bring the higher acuity visual receptors to where they can sample relevant