Available online at www.sciencedirect.com

ScienceDirect

Cognitive Systems Research 12 (2011) 102–112

Cognitive Systems RESEARCH

www.elsevier.com/locate/cogsys

The persistent visual store as the locus of fixation memory in visual search tasks

Action editor: David Peebles

David Kieras *

Computer Science and Engineering Division, Electrical Engineering and Computer Science Department, University of Michigan, 3641 Computer Science & Engineering Building, 2260 Hayward Street, Ann Arbor, MI 48109-2121, United States

Available online 10 August 2010

Abstract

Experiments on visual search have demonstrated the existence of a relatively large and reliable memory for which objects have been fixated; an indication of this memory is that revisits (fixations on previously fixated objects) typically comprise only about 5% of fixations. Any cognitive architecture that supports visual search must account for where such memory resides in the system and how it can be used to guide eye movements in visual search. This paper presents a simple solution for the EPIC architecture that is consistent with the overall requirements for modeling visually-intensive tasks and other visual memory phenomena.

Keywords: Visual search; Cognitive modeling; Eye movements

1. Introduction

Many everyday and work activities involve visual search, the process of visually scanning or inspecting the environment to locate an object of interest that will then be the target of further activity. Many human–computer interaction tasks require such visual search to be made in a visual environment that is much simpler than natural scenes. For example, a particular icon coded by color, shape, and other attributes must be located on a screen and then clicked on using a mouse. This domain combines relative simplicity of the visual characteristics of the searched-for objects with practical relevance: the task is a natural one in the sense that such activities are very common in current technology. Visual search is so heavily relied on in many computer-based systems that it probably is a major bottleneck in human-system performance. Thus understanding in detail how visual search works in such domains can lead to better system designs. In addition, if visual search can be understood in the context of a comprehensive computational cognitive architecture, then it will add to our knowledge of human perception, cognition, and action in the especially rigorous and coherent way characteristic of computational cognitive architectural modeling.

1.1. Visual search and active vision

In a laboratory visual search task, a display of objects is presented, and the participant must locate a particular object specified by perceptual properties and make a response based on whether such an object is present or exactly which properties it has (e.g. the specific shape). In most experiments, the display is static and contains some number of objects, only one of which is the target that must be responded to; the others are distractors. The properties of the display or the displayed objects are manipulated, and reaction time (RT) and/or eye movements are measured. The empirical literature on this task was dominated for a

* Tel.: +1 734 763 6739; fax: +1 734 763 1260.
E-mail address: kieras@umich.edu

1389-0417/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.cogsys.2010.06.007

long time by studies that measured only RT, and often for tachistoscopically presented displays that ruled out eye movements, but more recently the cost of eye movement data collection has decreased to the point that it has become much more common, and deservedly so. While any behavioral measurement only indirectly reflects the mental processes that produce it, RT is clearly much less diagnostic of what goes on during visual search than eye movements. Furthermore, tasks in which the eye is free to move about a static display in a naturalistic manner, typical of eye movement studies of visual search, will be more representative of the normal operation of the visual system and the role of attention in visual activity. This point was argued eloquently by Findlay and Gilchrist (2003) in presenting an *active vision* framework for understanding visual activity; it is markedly different from traditional approaches to visual attention which have ignored both the role of eye movements and extra-foveal information.

A key process in visual search is choosing the next object for inspection. A variety of studies (see Findlay and Gilchrist (2003), for a review) have shown that this choice is not at all random; rather the color, shape, size, orientation, or other visual properties of objects influences which object is chosen for the next fixation; the phenomenon is called *visual guidance* or *eye guidance*. In the active vision framework, these properties are available in extra-foveal or peripheral vision to some extent, meaning that visual attention, which in the context of normal visual activity is almost synonymous with where the eye is fixated, is a process of selecting for detailed examination one of a large number of objects currently perceived to be in the visual scene, and doing this selection on the basis of the visual properties available in extra-foveal vision.

1.2. Fixation memory

An important fact about visual guidance in visual search tasks is that an object that was previously fixated will be only rarely selected for a new fixation. This is an old result in eye movement studies (e.g. Barbur, Forsyth, & Wooding, 1993), but it did not receive much attention until the controversial Horowitz and Wolfe (1998) claim that "Visual search has no memory." They compared search RTs of a static display with a changing display, in which the objects changed positions during search, and found no difference in RT. If the visual search mechanism remembered where it had already inspected, it should be disrupted if the objects changed location; the RT being unaffected argues that the search was not disrupted, which means in turn that there was no memory for the previous fixations. Peterson, Kramer, Ranxiao, Irwin, and McCarley (2001) countered with a study demonstrating that "Visual search has memory". They recorded eye movements during search of a static display, and discovered, as earlier studies had noted, that revisits were rare, meaning that the previous fixations were remembered in some way.

1.2.1. Encoding failures trigger revisits

Peterson et al. went further with a detailed analysis showing that most revisits were made immediately after only one intervening fixation, which rules out memory failure as the cause of a revisit. Rather, Peterson et al. proposed that revisits were due to *encoding failures*: soon after fixating an object and moving onto the next, the person would realize that the previous object had not been fully encoded, and so would revisit it. Using a Monte-Carlo model, they demonstrated that this explanation accounted for the statistical structure of the revisits considerably better than either a no-memory or memory-failure model.

1.2.2. Search strategies dominate

Several subsequent studies (e.g. Geyer, von Mühlenen, & Müller, 2006; von Mühlenen, Müller, & Müller, 2003) using RT, eye tracking, and changing displays make it clear that Horowitz and Wolfe (1998) results were an artifact of how the changing displays would force a change in task strategy. If the display is changing, the only way to perform the task successfully is to use a strategy that compensates, such as to "wait and see" whether the target appears in a subset of the display. In other words, the changing-display paradigm forces a strategy that produces a no-memory effect. Regardless of the methodological issues and the merits of the results, an important implication is that making use of memory for previous fixations is not "hard-wired" in the visual system, e.g. at the oculomotor level, but rather is an optional feature of a visual search task strategy.

1.2.3. Objects are remembered, not locations

Additional studies (e.g. Beck, Peterson, & Vomela, 2006) have attempted to determine whether what is remembered on each fixation is the location, the identity, or the properties of the objects. However, it should be clear that in a changing-display paradigm, if objects are identified in terms of their properties (e.g. shape), then they are "teleporting" from one location to the next, which is not a natural input to the visual system. Hulleman (2009) performed the most elegant and naturalistic test of whether fixation location was remembered simply by having the objects move around on the display during search similar to Pylyshyn and Storm (1988) multiple object tracking paradigm. He observed almost no difference in search rates compared to a static display. This strongly suggests that fixation locations themselves were not remembered, since the objects were continuously changing location. The conclusion would seem to be that previously fixated *objects* are being remembered, where object identity persists over changes in location. In a naturally static display, such as Peterson et al. (2001) paradigm, the issue does not arise: objects retain their location and properties.

1.2.4. Fixation memory has large capacity

The consensus of the empirical literature at this point is that memory for previous fixations exists. Moreover, it has a fairly large effective capacity. The Peterson et al. study