Modeling visual search experiments: The Selective Attention for Identification Model (SAIM)

Dietmar Heinke [a] Glyn W. Humphreys [a] Giovanni di Virgilio [a]

[a] *Behavioural and Brain Sciences Centre, School of Psychology, University of Birmingham, Birmingham B15 2TT, United Kingdom,*
d.g.heinke@bham.ac.uk

Abstract

We have presented a computational model called Selective Attention for Identification Model (SAIM), that can account for a broad range of psychological and neuropsychological phenomena on attention [1]. In this paper we report on work that extends SAIM to model data from visual search tasks. The results show that SAIM can capture important aspects of findings in visual search experiments, including variations of search slopes with the similarity between targets and distractors. SAIM is also capable of simulating experimantal findings for redundant targets, including the violation of the Miller inequality. SAIM explains these experimental findings as the consequence of competitive processes involved in object identification.

Key words: Connectionist modeling, Visual Search, Attention, Competitive mechanism, Redundancy Gain

1 Introduction

Recently, we have presented a connectionist model of human visual attention, termed SAIM (Selective Attention for Identification Model) [1]. SAIM's behaviour is dominated by competitive interactions between units and modules which can provide a qualitative account for a range of psychological phenomena on both normal and disorders attention. Simulations on normal attention demonstrate two-object costs on selection, effects of object familiarity on selection, global precedence, spatial cueing both within and between objects, and inhibition of return. When simulated lesions were conducted, SAIM demonstrated both unilateral neglect and spatial extinction, depending on the type and extent of the lesion. Different lesions also produced view-centred and

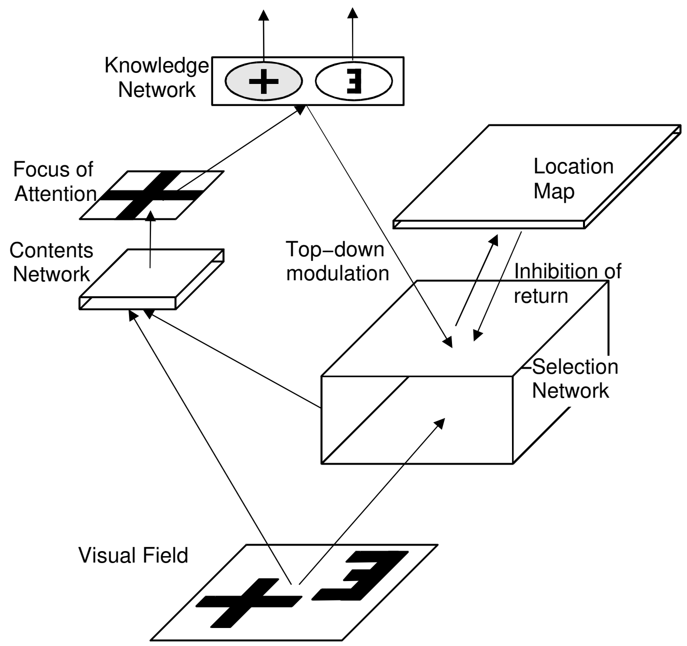

Fig. 1. Architecture of SAIM

object-centred neglect, enabling both forms of neglect to be simulated within a single patient. SAIM highlighted how emergent properties of the model unify object- and space-based theories of normal selection and 'attentional' and 'representational' accounts of visual neglect. In essence, SAIM suggested that attentional effects in human behaviour could result from competitive interactions in visual selection for object recognition and the related disorders are due to imbalanced competition following damage to certain areas of the brain. The results presented here aim at extending SAIM to modeling data from visual search experiments. In visual search experiments participants are asked to report the absence or presence of a specified target item amongst irrelevant items (distractors). The performance of the participants is measured in terms of time until response (reaction time). The number of distractors is varied across trials. The typical outcome of such experiments is a linear relation between reaction time and number of distractors. The slope of this linear relation is often interpreted as indicator for the underlying search mechanism. For instance, a small slope (0-10ms/item) is interpreted as parallel search and a steep slope (20-50ms/item) is assumed to indicate a serial search (see [7], for a recent review). The question here is whether SAIM's competitive mechanisms are able to simulate these basic findings in visual search experiments.

2 Overview of SAIM

Figure 1 gives an overview of SAIM's architecture and highlights the modular structure of SAIM with each module concurrently aiming at different objec-

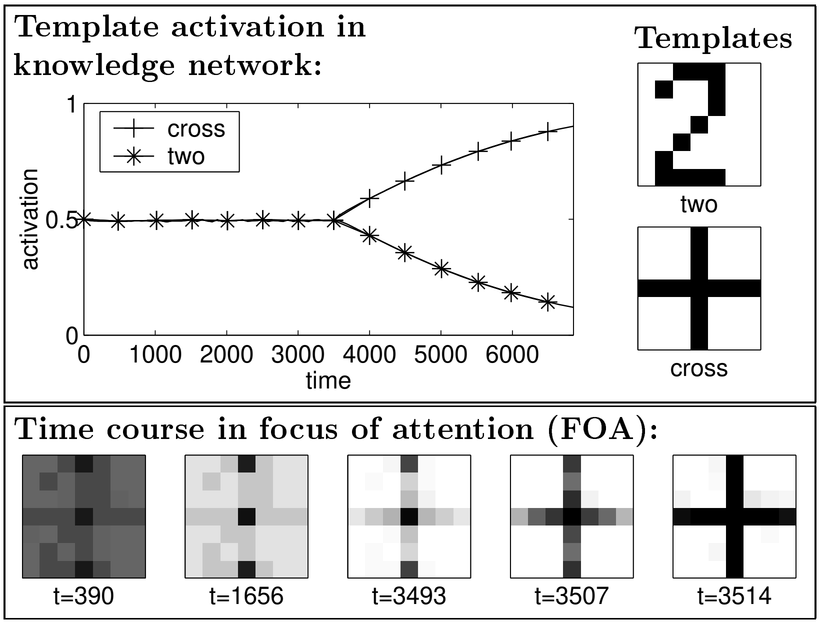

Fig. 2. Example of a simulation result. The input image is shown in Fig. 4

tives. The contents network maps a section of the input into a smaller Focus of Attention (FOA) taking into account a spatial selection process. In addition the mapping of the contents network is translation-invariant, enabling SAIM to perform translation-invariant object identification. The selection network control the contents network by selecting a certain spatial location and ensuring a correct mapping of the selected object. The knowledge network identifies the contents of the FOA using template matching and modulates the behaviour of the selection network with top-down activation so that known objects are preferred over unknown objects. The location map enables SAIM to scan serially by storing already selected location and preventing them from being return to through inhibition of return. The design of SAIM's network structure followed the idea of soft constraint satisfaction in neural networks. Examples for constraints in SAIM are, activating the best-matching template unit in the knowledge network and, for the selection network, ensuring a correct appearance of objects in the FOA. The resulting non-linear dynamics in SAIM consist of cooperative and competitive interactions between units and modules.

In order to model experimental data from visual search tasks, SAIM needed to be extended in two ways. First, the size of the input image had to be increased to allowed for more then two items in the input. The new input size allows for up to 40 input items. As a consequence of extending the input size, the parameters of the model had to be adapted so that the selection process can still suppress items. Second, to make the input image more realistic, noise was added. Apart from these modifications SAIM's basic structure and mechanism were kept the same, consequently, results reported earlier can be captured with the new version as well.

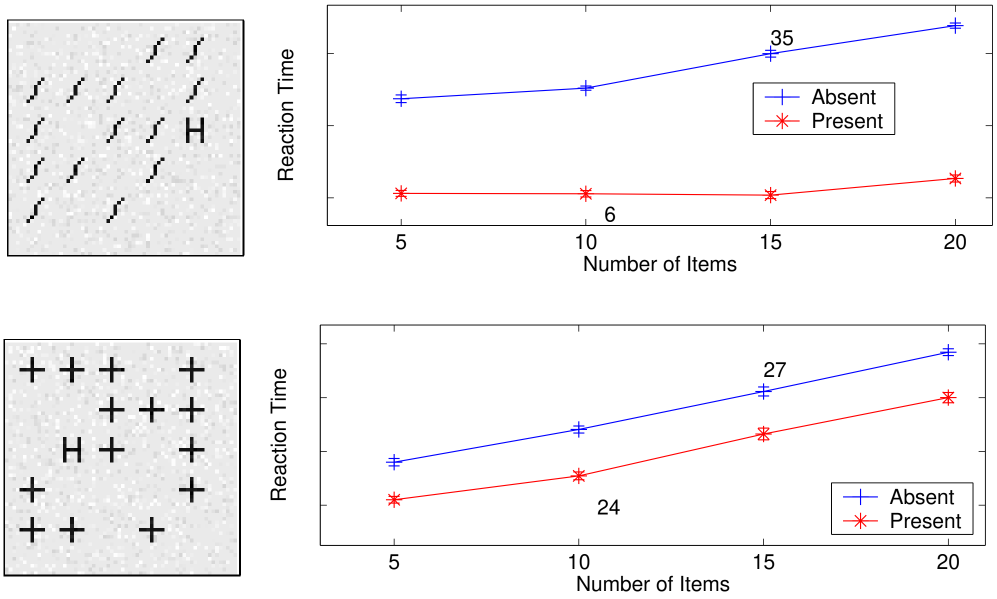

Fig. 3. Simulation of the parallel vs. serial dichotomie in visual search. The images on the left give an example of the displays used. The search functions on the right show different search slopes such as flat (6ms/item) for "parallel" search and steep (24 ms/item) for "serial" search.

3 Results and discussion

Fig. 2 gives an example of a simulation result for a cross amongst twos, highlighting a bottom-up selection advantage for the cross over the two. For simulating visual search the displays used contained one target (if present) amongst homogeneous distractors. Four different display sizes were used: 5, 10, 15 and 20 items. For each display size 20 simulation runs were run, varying noise and arrangement of the items. The templates in the knowledge network represented the items used in the displays. SAIM's reaction time is the point in time for which it becomes clear which template unit is the winner. This time is reported as "ms". Note that simulations should be understood as qualitative studies of human behaviour and that SAIM's behaviour does not necessarily mirror human performance for the particular type of displays used. This is particularly relevant for the crucial factor of SAIM's bottom-up prioritizing of items which is based on the shape of items, including their size.

In line with the most common experimental findings SAIM produced linearly increasing reaction times with increasing numbers of items in the input image (see Fig. 3). In SAIM this effect results from an increasing competition between increasing number of items. The explanation differs from the assumption of a serial search scan [6]. As in experiments the slope of the search function varies depending on the target/distractor pairing. In SAIM this variation results from the similarity of targets and distractors in term of their priority in the selection process. The similar the priorities are the steeper the slope.

The results, reported so far, used target/distractor pairs, where SAIM prior-

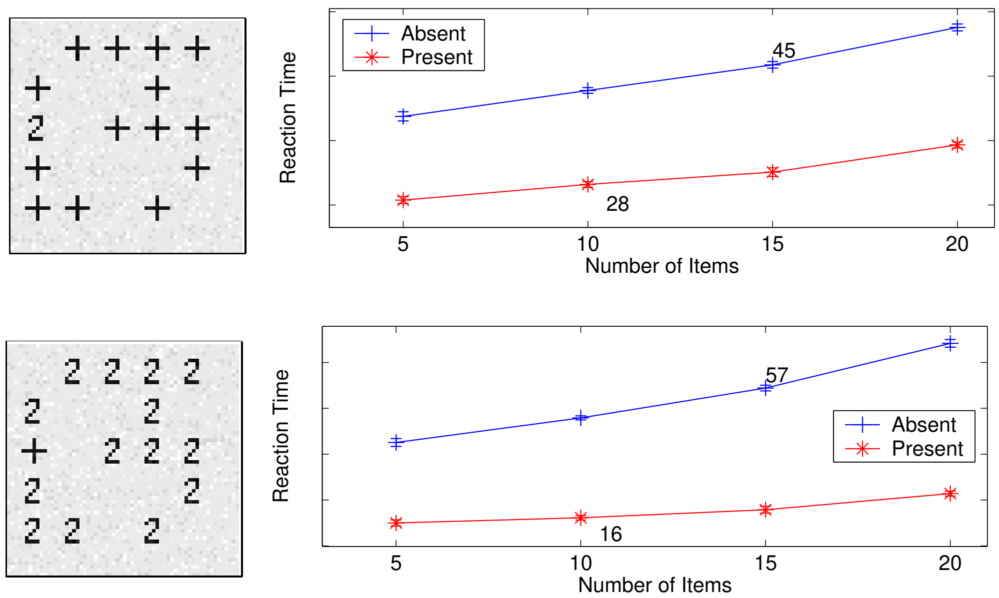

Fig. 4. Simulation of search assymetries. The search for two amongst crosses is slower (top) than cross amongst twos (bottom).

itizes the target in a bottom-up way. But SAIM can modulate the saliency through the recognition stage in a top-down way. This mechanism can explain the effect of search asymmetries where, for instance, Qs amongst O lead to quicker responses than O's amongst Qs [5]. In the faster condition, SAIM performs a simple bottom-up controlled selection, whereas in the slower condition the top-down modulation from the recognition stage needs to shift the bias towards an item which is normally not selected in a bottom-up way. This top-down modulation results in a response cost, enabling SAIM to simulate search asymmetries (see Fig. 4). Finally, SAIM is capable of simulating the reaction time gain resulting from redundant targets in search tasks and the violation of the Miller inequality (see Fig. 5). The Miller inequality compares the cumulative distribution function of reaction times from the single target condition with the cumulative distribution function of reaction times from the redundant target condition. A violation of the inequality is seen as a support for the assumption that a parallel selection process is speeded up by redundant targets, consistent with there being coactivation from the multiple targets [3]. In fact, this model is supported by SAIM, where coactivation occurs in the focus of attention through the presence of redundant targets, leading to a speeded response from the recognition stage.

4 Conclusion

Results of visual search task are often discussed in terms of serial and parallel processing [5] or terms of saliency guided serial search [6]. On the contrary

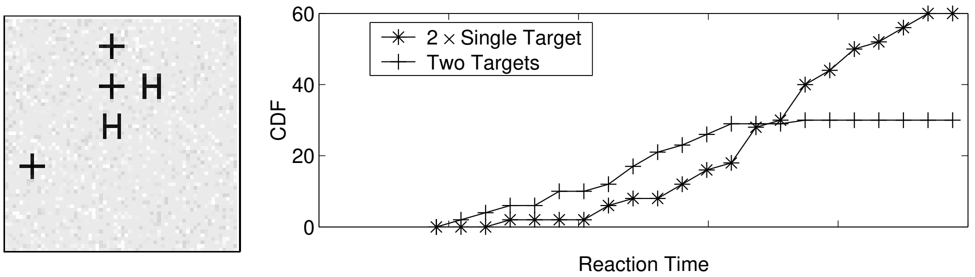

Fig. 5. Simulation of redundancy gain of 100ms. The simulation results also indicate a violation of the Miller inequality illustrated by the cumulative distribution function (cdf) on the right.

SAIM suggests that the results stem from a parallel competition process taking place during object recognition. This process is modulated by the degree of interference between items and an interaction between top-down and bottom-up factors, explaining a large variety of experimental findings in visual search tasks. It should be noted that the competition in SAIM is an integral part of the recognition process. This is a major conceptual difference from common models of visual search (e.g. [4]) which must assume that experimental findings in visual search result from an external "attentional" system containing a saliency map. SAIM can account for visual search data without these assumptions (see [2] for a complete discussion). Hence, SAIM seems to give a more parsimonious account of the data than other models.

References

[1] D. Heinke and G. W. Humphreys. Attention, spatial representation and visual neglect: Simulating emergent attentional processes in the Selective Attention for Identification Model (SAIM). *Psychological Review*, in press.

[2] D. Heinke and G. W. Humphreys. Computational Models of Visual Selective Attention: A Review. In G. Houghton, editor, *Connectionist Models in Psychology*. in press.

[3] J. Miller. Divided Attention: Evidence for Coactivation with Redundant Signals. *Cognitive Psychology*, 14:247–279, 1982.

[4] M. C. Mozer and M. Sitton. Computational modeling of spatial attention. In H. Pashler, editor, *Attention*, pages 341–393. Psychology Press, 1998.

[5] A. Treisman. Features and Objects: The Fourteenth Bartlett Memorial Lecture. *The Quartely J. of Exp. Psychology*, 40A(2):201–237, 1988.

[6] J. M. Wolfe. Guided Search 2.0 A revised model of visual search. *Psychonomic Bulletin & Review*, 1(2):202–238, 1994.

[7] J. M. Wolfe. Visual Search. In H. Pashler, editor, *Attention*, pages 13–74. Psychology Press, 1998.