A biased competition computational model of spatial & object-based attention mediating active visual search

Authors: **Linda J Lanyon*, Susan L Denham**

Centre for Theoretical & Computational Neuroscience, University of Plymouth, U.K.

*Corresponding author: Email Linda.Lanyon@plymouth.ac.uk

Abstract

Colour is more important than form in driving the scan path during active visual feature conjunction search. A model shows how feature-based attention in V4 provides information to LIP allowing it to represent behaviourally relevant locations and attract attention to target coloured locations. Attentional effects in V4 develop such that an initial spatial attention focus, scaled according to stimulus density, becomes object-based later in the response. Thus, this is the first model to replicate, with temporal precision, different attentional effects observed in single cell studies of V4 and IT, in addition to the systems level scan path behaviour.

Keywords: Visual attention; Visual search; Biased competition; Object-based attention; Saccades

1. Introduction

The biased competition hypothesis [5] is currently a very topical theory used to explain attention effects observed in single cell studies of the primate visual system (e.g. [1,2,15,16]). Biased competition suggests that neuronal responses are determined by competitive interactions that are subject to a number of biases, such as "bottom-up" stimulus information and "top-down" cognitive requirements, in particular a working memory template of the target object from prefrontal cortex. When more than one stimulus is present within a cell's receptive field, the response of the cell is determined by which of the stimuli are attended. Biased competition models have recently developed from small scale (e.g. [20]) to more complex systems level models (e.g. [4]). Here, we describe a model for *active* visual search, where retinal inputs change at each fixation as the focus of attention moves. The model is able to simulate both spatial and object-based attentional modulations seen in single cell studies of inferior temporal (IT) region [1], and extrastriate area V4 [2,12,16]. This appears to be the first model to accurately replicate the time courses of these different attentional modulations. At the systems level, the model is able to reproduce scan path behaviour found [18] in monkeys performing feature conjunction search, where fixations tended to land near stimuli of the target colour.

Target orientation did not appear to influence the scan path. The model achieves this by means of a "cross-stream" interaction between extrastriate areas in the ventral pathway leading to temporal cortex and the dorsal pathway leading to parietal cortex.

A detailed description of the model's simulation of single cell recordings in IT [1] and V4 [2,12,16] has been given by [10] and a subsequent examination of the scan path under different scene conditions was provided by [11]. Here, we investigate the effect of the aperture of the initial spatial attention focus being scaled when stimulus density varies across the scene. We also consider the sensitivity of the model to the weighting of V4 connections to parietal area LIP and the strength of novelty feedback.

2. The model

The model focuses on emergent attentional effects in intermediate stages of the visual system and Figure 1 shows the areas modelled. Featural "pre-processing" is provided by retinal and V1 modules, but these do not form part of the dynamics of the main active vision portion of the system. Attention arises as an emergent property of this portion of the system in modules representing V4, IT and the lateral intraparietal area (LIP). These modules use mean field population dynamics [6], where representation is at the level of cell populations, known as pools or assemblies. The model operates by moving its retina around the scene. The size of the dynamic cortical areas is dependant on the size of the retina, which can be adjusted for each simulation. The model's ventral stream (running from V1 to V4 and IT) operates as a feature/object processing hierarchy with receptive field sizes increasing towards IT and encoded features becoming more complex, such that objects are encoded in an invariant manner in IT. In V1 and V4 both orientation and colour features are encoded. LIP encodes the locations of behaviourally relevant features [3], perhaps acting as a "salience map" [7] and may form part of the cortical network responsible for selecting possible targets for saccades. In the modelled LIP, spatial locations compete and activity is used to decide the next location to be fixated. Featural information from V4 is integrated within the spatial competitive framework in LIP and this allows it to represent behaviourally relevant locations most strongly, and guide the scan path accordingly.

Object-based attention has not been observed in V4 until at least 150ms post-stimulus [2,16]. However, spatial attention seems able to module responses from the onset of the stimulus-related

response at ~60ms post-stimulus, and even enhances baseline firing in advance of the stimulus-related response [12]. In the model, both effects are simulated at the appropriate times by means of two different types of bias to V4. The spatial enhancement is due to a spatial bias to competition in LIP (where spatial attention is also found: e.g. [3,9]), which then biases V4. The source of the spatial bias is suggested to be the frontal eye field (FEF), because microstimulation of FEF leads to spatial modulation of V4 [14]. The spatial bias results in the creation of an initially spatial *attention window* (AW). The aperture of this window is scaled according to coarse resolution information reflecting local stimulus density [17], which is assumed to be conveyed rapidly by the magnocellular pathway to parietal cortex. All other information within the system is assumed to be derived from the parvocellular pathway. Thus, during a scan path, the size of the AW is dynamic, being scaled according to stimulus density found around any particular fixation point. The following gives the radius of the AW and is inspired by the psychophysical findings of Motter and Belky [17]:

$$AWrad = \min\left[Round\left[\sqrt{\frac{m*n}{d}} * 2 \right], \min(m,n) \right] \quad (1)$$

where:

d is the number of non-zeros in *f* below

m,n are the dimensions of the retinal image

$$f = \psi\left(\sum_{k=1}^{K} I_{ijrk}\right) \quad (2)$$

where:

I_{ijrk} is the output of V1 simple cells, selective to K orientations, over the area of the retinal image.

r is the spatial resolution of the orientation information; set to the lowest resolution detected

ψ is a function that removes the lowest 95% of activity and reduces to zero the activity at points where a neighbour has been found within a Euclidean distance equal to the length of the stimuli (simple bar stimuli are used).

Over time, attention gradually becomes object-based. Object-based attention develops in parallel across the cortical region and is not constrained by the spatial AW. However, object-based attention is facilitated by spatial attention such that responses are strongest within the AW, resulting in a combined attentional effect [13]. It is suggested that object-based attention takes longer than spatial attention to

develop because it is reliant on the resolution of competition between objects and features. Object-based attention results from an object-related bias from prefrontal cortex (responses of prefrontal "late" neurons [19] are modelled) to IT. Objects in IT provide a "top-down" bias to features in V4, such that features not related to the object are suppressed, i.e. a feature-based attention effect. Under conditions of visual search involving a colour (or luminance-defined) target, object-based attention in V4 results in target coloured (or target luminance) stimuli becoming effectively "highlighted" across the scene in V4 [16] and this enables V4 to bias LIP towards representing these locations most strongly. Object-based effects also operate within the orientation feature in V4 but the connection from colour features in V4 to LIP is slightly stronger than the connection from orientation features in V4. This allows the scan path to be drawn to target coloured locations across the scene.

Inhibition of return (IOR) in the scan path is implemented by means of a frontal bias to LIP related to the *novelty* of the location. IOR has been suggested to be influenced by recent event/reward associations linked to orbitofrontal cortex [8], which is a possible source of the bias.

3. Simulation Results

During a scan path, next fixation points tend to be chosen from within the current AW because this area of LIP is subject to a spatial enhancement of activity. For any given retinal image size, the AW is much smaller for dense scenes than that within sparse scenes. This results in small saccades being made in dense images and larger saccades in sparse scenes. The AW is dynamically scaled according to local information as the retina moves around the scene. Therefore, when stimulus density varies across the scene, the AW changes size during the scan path so that dense areas are examined more thoroughly by a series of smaller saccades, as shown in Figure 2. The AWs for two of the fixations in contrasting areas of this scan path are shown in Figure 3.

Spatial attention has an early effect on V4 responses and subsequent responses are modulated by the combined effect of the spatial AW and object-based attention [13]. Object-based attention in V4 is able to influence activity in LIP such that the scan path is attracted to areas containing the target colour, as shown in Figure 2. The amount of non-target coloured locations attracting attention is influenced by the weight of the novelty bias to LIP, the weight of connections from V4 to LIP and the strength of

object-related feedback from IT to V4. The latter was suggested by [10] to be affected by training, i.e. familiarity with objects and the task. The relative weight of connection from colour and orientation assemblies in V4 to LIP has an effect on the number of non-target coloured locations attracting attention. However, target coloured locations continue to dominate the scan path even when the colour connection is only very slightly stronger than the orientation connection. The strength of novelty bias to LIP appears to have a large effect on the scan path. If the novelty bias to LIP is very strong, it does not allow LIP to discriminate any stimuli within the vicinity of a previous fixation and this tends to reduce the choice of stimulus locations for the next fixation. In very sparse scenes, this can result in blank areas of the display becoming more active in LIP than stimulus locations. Therefore, when the novelty bias is set to a very high value, the scan path tends to select more blank areas (20% of fixations were found in blank areas in [18]).

4. Conclusion

This model has been used to simulate attentional effects at the single cell level, with accurate onset times for both spatial and object-based attention in IT and V4 [10]. At the behavioural level, it has also been able to simulate visual search scan paths found in monkeys performing feature conjunction search [18]. IT feedback to V4 allows object-based attention to develop therein and this influences spatial competition in LIP such that the scan path is attracted to behaviourally relevant locations. Here, we find that, in addition to other factors, the scan path is influenced by the strength of novelty feedback to LIP. The novelty bias may reflect a cognitive strategy to seek new locations and could reflect a previous event-reward association [8]. During search within a scene where stimulus density varies, large saccades tend to traverse sparse areas but much shorter saccades take place within dense regions, allowing detailed investigation of the area.

References

[1] L. Chelazzi, E.K. Miller, J. Duncan, R. Desimone, A neural basis for visual search in inferior temporal cortex, Nature 363 (1993) 345-347

[2] L. Chelazzi, E.K. Miller, J. Duncan, R. Desimone, Responses of neurons in macaque area V4 during memory-guided visual search, Cerebral Cortex 11 (2001) 761-772

[3] C.L. Colby, J.R. Duhamel, M.E. Goldberg, Visual, presaccadic, and cognitive activation of single neurons in monkey lateral intraparietal area, J. Neurophysiol 76(5) (1996) 2841-2852

[4] G. Deco, T.S. Lee, A unified model of spatial and object attention based on inter-cortical biased competition, Neurocomp. 44-46 (2002) 775-781

[5] R. Desimone, J. Duncan, Neural mechanisms of selective visual attention, Ann. Rev. Neurosci. 18, (1995) 193-222

[6] W. Gerstner, Population dynamics of spiking neurons: Fast transients, asynchronous states, and locking, Neural Comput. 12 (2000) 43-89

[7] J.P Gottlieb, M. Kusunoki, M.E. Goldberg, The representation of visual salience in monkey parietal cortex, Nature 391 (1998) 481-484

[8] T.L. Hodgson, D. Mort, M.M. Chamberlain, S.B. Hutton, K.S. O'Neill, C. Kennard, Orbitofrontal cortex mediates inhibition of return, Neuropsychologia 431 (2002) 1-11

[9] S. Kastner, M. Pinsk, P. De Weerd, R. Desimone, L. Ungerleider, Increased activity in human visual cortex during directed attention in the absence of visual stimulation, Neuron 22 (1999) 751-761

[10] L. J. Lanyon, S. L. Denham, A biased competition model of spatial and object-based attention mediating active visual search, Vision Research (2003) (submitted)

[11] L. J. Lanyon, S. L. Denham, A Model of Active Visual Search with Object-Based Attention Guiding Scan Paths, Neural Networks (2003) (submitted)

[12] S.J. Luck, L. Chelazzi, S.A. Hillyard, R. Desimone, Neural mechanisms of spatial attention in areas V1, V2 and V4 of macaque visual cortex, J. Neurophysiol. 77 (1997) 24-42

[13] C.J. McAdams, J.H.R. Maunsell, Attention to both space and feature modulates neuronal responses in macaque area V4, J. Neurophysiol. 83(3) (2000) 1751-1755

[14] T. Moore, K.M. Armstrong, Selective gating of visual signals by microstimulation of frontal cortex, Nature 421 (2003) 370-373

[15] J. Moran, R. Desimone R, Selective attention gates visual processing in the extrastriate cortex, Science 229 (1985) 782-784

[16] B.C. Motter, Neural correlates of attentive selection for color or luminance in extrastriate area V4, J. Neurosci 14(4) (1994) 2178-2189

[17] B.C. Motter, E.J. Belky, The zone of focal attention during active visual search, Vision Research 38(7) (1998) 1007-1022

[18] B.C. Motter, E.J. Belky, The guidance of eye movements during active visual search, Vision Research 38(12) (1998) 1805-1815

[19] R. Romo, C.D. Brody, A. Hernandez, L. Lemus, Neuronal correlates of parametric working memory in the prefrontal cortex, Nature 399(6735) (1999) 470-473

[20] M. Usher, E. Niebur, Modeling the temporal dynamics of IT neurons in visual search: A mechanism for top-down selective attention, J. Cog. Neurosci. 8(4) (1996) 311-327

Linda Lanyon has a 1st class honours degree in Computing & Informatics and comes from a background in computer systems development, most recently acting as IT Advisor to a major U.K. e-government project. She is currently completing the final year of her PhD, from which this work is drawn, at the Centre for Theoretical & Computational Neuroscience, University of Plymouth, U.K.

Dr Sue Denham obtained her PhD in 1995 from the University of Plymouth where she holds the post of Principal Lecturer in the Centre for Theoretical and Computational Neuroscience. Her research interests lie in developing theoretical and computational models of perception and learning, with particular interest in the representation of timbre and the segregation and grouping of sounds within natural acoustic environments.

Figure 1

Overview of the model

Figure 2

Scan path over a scene with varying stimulus density. Attention is attracted to the locations of target coloured stimuli (in this case the target colour is green, shown here as black; the non-target colour is red, shown as grey). From an initial fixation point in the centre of the image, large saccades are made within the sparse area of the scene but subsequent saccades within the dense area are shorter.

Figure 3

Shows the difference in aperture of the AW for two points during the scan path shown in Figure 2. The retinal image is shown as the outer of the plotted boxes. In the dense stimulus region, the AW (inner boxes plotted within each retinal image box) is much smaller than that in the sparse region.

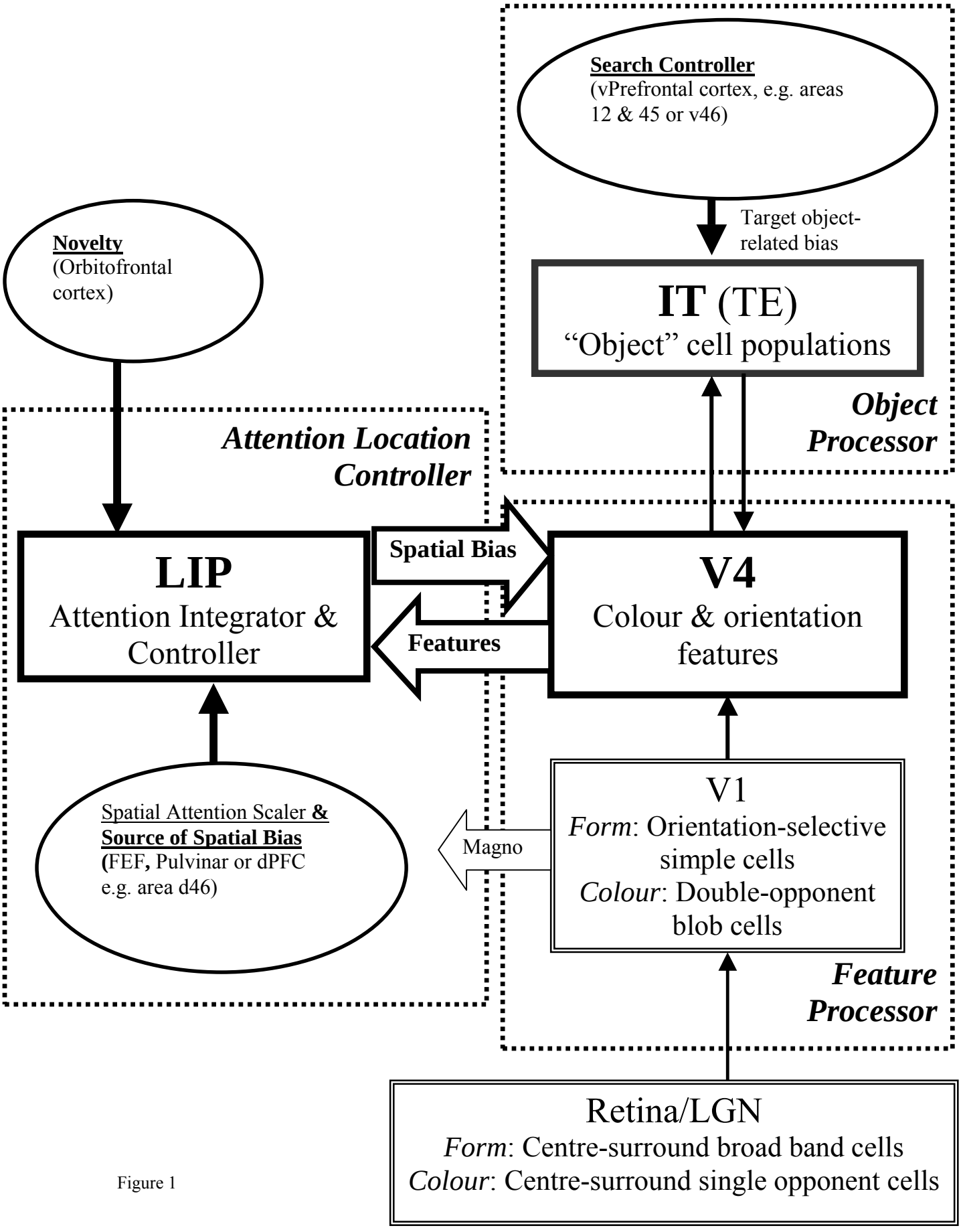
Search Controller
(vPrefrontal cortex, e.g. areas 12 & 45 or v46)
Target object-related bias
IT (TE)
"Object" cell populations
Object Processor
Novelty
(Orbitofrontal cortex)
Attention Location Controller
LIP
Attention Integrator & Controller
Spatial Bias
Features
V4
Colour & orientation features
Spatial Attention Scaler & Source of Spatial Bias
(FEF, Pulvinar or dPFC e.g. area d46)
Magno
V1
Form: Orientation-selective simple cells
Colour: Double-opponent blob cells
Feature Processor
Retina/LGN
Form: Centre-surround broad band cells
Colour: Centre-surround single opponent cells

Figure 1

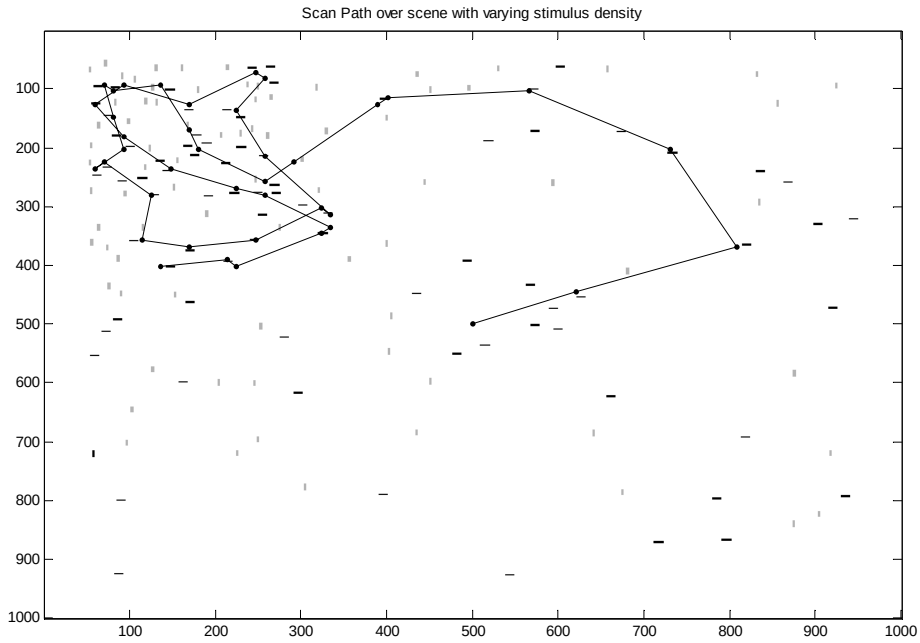

Figure 2

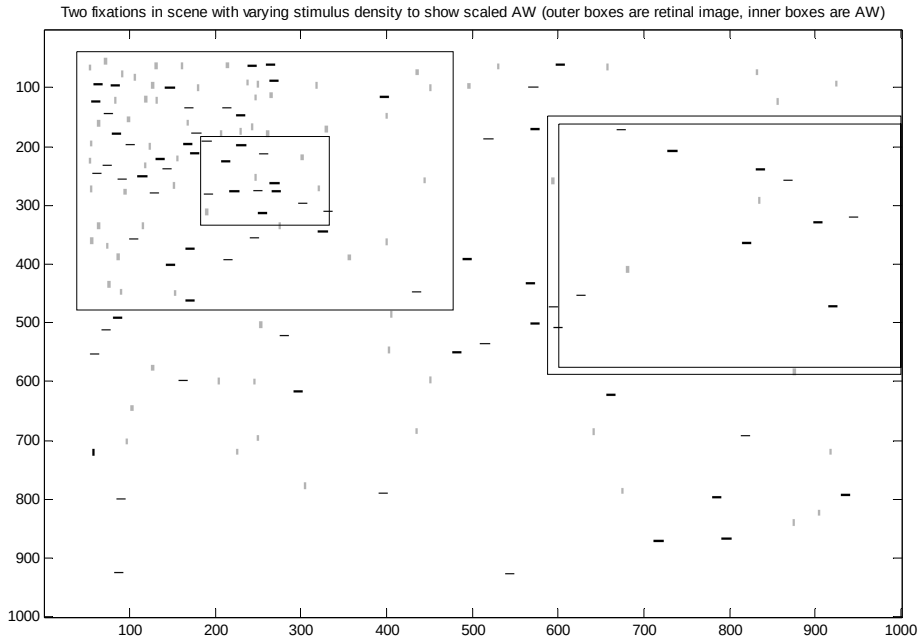

Figure 3