Modelling divided visual attention with a winner-take-all network

Dominic I. Standage[1], Thomas P. Trappenberg[1]
and Raymond M. Klein[2]

[1]Faculty of Computer Science, [2]Department of Psychology
Dalhousie University, Halifax, Canada

Neural Networks (2005) 18(5-6):620-627

Abstract

Experimental evidence on the distribution of visual attention supports the idea of a spatial saliency map, whereby bottom-up and top-down influences on attention are integrated by a winner-take-all mechanism. We implement this map with a continuous attractor neural network,and test the ability of our model to explain experimental evidence on the distribution of spatial attention. The majority of evidence supports the view that attention is unitary, but recent experiments provide evidence for split attentional foci. We simulate two such experiments.Our results suggest that the ability to divide attention depends on sustained endogenous signals from short term memory to the saliency map,stressing the interplay between working memory mechanisms and attention.

1 Introduction

Attention is an old concept in psychology correlated with enhanced processing of objects or regions in space (Posner, Snyder, & Davidson, 1980). While attention is a multi-modal phenomenon (Cherry, 1953; Zelano et al., 2004), the majority of research has focused on selective visual attention (SVA). The limited capacity of the visual system necessitates a mechanism to select stimuli from the visual field, and Tsotsos pointed out that attention solves the complexity problem of sensory processing (Tsotsos, 1992).

A distinction can be drawn between pre-attentive and attentive visual processing (Neisser, 1967). Pre-attentive processing refers to bottom-up (BU) feature saliency of visual stimuli whereby items that differ from their surroundings 'pop out' to the viewer. Attentive processing refers to top-down (TD) influences on perception of stimuli determined by object and locational bias such as task instructions or foreknowledge of stimulus characteristics. Determining saliency, then, is both a BU and TD requirement, and computational models of SVA include maps that integrate BU salience across object features (Koch & Ullman, 1985), TD bias (Treisman, 1998), and the interplay of both (Wolfe, 1994; Deco, Pollatos, & Zihl, 2002).

Koch and Ullman (1985) (Koch & Ullman, 1985) provide a neural network model of SVA in which topographic feature maps are integrated by a winner-take-all (WTA) saliency map of BU stimuli. In their model, inhibiting the selected location causes a shift to the next most salient location. Wolfe (1994) (Wolfe, 1994) builds on Neisser's pre-attentive/attentive distinction (Neisser, 1967), integrating BU and TD saliency criteria in his Guided Search model. Treisman (1998) (Treisman, 1998) provides a model of spatial attention to solve the Binding Problem, in which a TD saliency map determines object features selected for further processing, and suggests parietal cortex as the biological correlate of her 'master' map. Deco et al. (2002) (Deco et al., 2002) use inhibition to mediate BU and TD influences in an instantiation of Duncan and Humphreys' biased competition model (Duncan & Humphreys, 2002), simulating saliency in posterior parietal cortex (PP) with a Continuous Attractor Neural Network (CANN). Spatial saliency in PP interacts with BU feature maps to converge on a winning location. See Shipp (2004) (Shipp, 2004) and Itti and Koch (2001) (Itti & Koch, 2001) for a review of these and other models.

There is long-standing debate about the distribution of SVA. Many cognitive models propose a unitary focus of attention, likened to a roving spotlight

over the visual field (Posner et al., 1980). Variants of the spotlight metaphor include gradient (Downing & Pinker, 1985; LaBerge & Brown, 1989) and zoom lens (Eriksen & James, 1986) models, suggesting that attention may be a graded phenomenon, attenuated around a central focus. A large body of evidence supports such unitary models (Posner et al., 1980; McCormick, Klein, & Johnston, 1998), but several more recent experiments have provided evidence for non-contiguous allocation of SVA (Hahn & Kramer, 1998; Awh & Pashler, 2000; Muller, Malinowski, Gruber, & Hillyard, 2003).

Here, we study how split attention can be achieved by a dynamic implementation of a WTA map. Despite their WTA nature, CANNs are able to account for split attention when network dynamics facilitate long transition states between regimes (Trappenberg & Standage, 2005) and when dominated by sustained inputs (D. Standage, Trappenberg, & Klein, 2005). We simulate the experiments of Muller et al. (2003) (Muller et al., 2003) with a 1-dimensional (1D) CANN model. We build on simulations presented in Standage et al. (2005a) (D. I. Standage, Trappenberg, & Klein, 2005) that use a narrow weight profile, facilitating steeply sloped regions of activity that occupy a small portion of the network. Because we don't know the size of the active region of PP and its relation to coordinates in the visual field, we run similar experiments with a wide weight profile, resulting in activity that spans the majority of the network. We demonstrate that the ability of the model to account for divided attention doesn't depend on fine tuning this network parameter.

We simulate two experiments by Awh and Pashler (2000) (Awh & Pashler, 2000) with a 2-dimensional (2D) CANN model, demonstrating how the model accounts for their finding divided attention in one experiment and unitary attention in the other. Preliminary simulations in 1D are reported in Standage et al. (2005a) (D. I. Standage et al., 2005). Our simulations are consistent with their experimental findings, but our model offers an alternative conclusion.

2 Methods

In 1D and 2D simulations, we use a fully connected recurrent rate model with N nodes, where $N = N_x N_y$. We model only PP from the model by Deco et al. (2002) (Deco et al., 2002). WTA is implemented by local cooperation and long distance competition in the laterally connected network. The average

state u_i of a node with index i is given by

$$\tau \frac{\mathrm{d}u_i(t)}{\mathrm{d}t} = -u_i(t) + \sum_j w_{ij} r_j(t)\ a + I_i^{\mathrm{ext}}(t), \tag{1}$$

where τ is a time constant, I_i^{ext} is external input to the network, $a = 2\pi/N_x$ is a scale factor, and r_i is a normalized square of u_i given by

$$g(u_i) = \frac{{u_i}^2}{1 + \frac{1}{2}\ a \sum_j {u_j}^2}. \tag{2}$$

We use this normalization through divisive normalization (shunting inhibition) to force more biologically realistic smooth (Gaussian) bubbles (Deneve, Pouget, & Latham, 1999).

The weight matrix **w** is determined by a shifted Gaussian function

$$w_{ij} = A_{\mathrm{w}}\ e^{-d^2/2\sigma_w^2}\ -\ C \tag{3}$$

between node i and node j where d is given by

$$\begin{aligned} d &= \sqrt{d_x^2 + d_y^2} && (4)\\ d_x &= min(|i_x - j_x|a, 2\pi - |i_x - j_x|a) && (5)\\ d_y &= min(|i_y - j_y|a, 2\pi - |i_y - j_y|a), && (6) \end{aligned}$$

and i_x and i_y are the x and y components of node i, $d_y = 0$ in the 1D case, C is an inhibition constant describing the activity dependent inhibition of an inhibitory pool of neurons, and A_{w} is a scale factor.

The external input I_i^{ext} is Gaussian shaped around input location j, determined by

$$I_i^{\mathrm{ext}} = e^{-d^2/2\sigma_{ext}^2} \tag{7}$$

where d is given by Equation 4.

In 1D, $N_x = 100$, $N_y = 1$ ($N = 100$). In 2D, $N_x = N_y = 30$ ($N = 900$). In all simulations, $C\ \epsilon\ \{0.1, 0.3\}$, $A_{\mathrm{w}} = 10$, $t = 1$, $\tau = 10$, and $\sigma_w = 1.2$ (1D) and 1.3 (2D), $\sigma_{ext}\ \epsilon\ \{0.3, 0.5\}$.

We classify our inputs along exogenous (exo) and endogenous (endo) dimensions. Exo inputs refer to neural responses to stimuli, here representing

visual cues. Endo inputs refer to voluntary control of attention, here representing task instructions in behavioural studies. Exo and endo inputs thus correspond to BU and TD signals respectively.

Simulations are run with transient and sustained inputs. We equate network activity with SVA. Because transient inputs elicit WTA behaviour in CANN models, we start by demonstrating one-bubble attractor states as models of a unitary attentional focus. Transient input stimuli are the norm in biological networks, as evidenced by high firing rates at stimulus onset followed by lower rates when stimuli are sustained in experimental settings. In the exo case, this initial burst of activity serves as input to higher cortical areas such as PP. Sustained firing after transient stimulation is a property of highly specialized neural assemblies (Funahashi, Bruce, & Goldman-Rakic, 1989), and as such is the exception among biological networks, not the norm. Where we simulate exo stimuli as sustained inputs, the stimuli being modelled are spatially static, rapidly changing symbols. We interpret these changes as providing continual 'refreshment' of neural representations due to novelty effects (Colby & Goldberg, 1999). We interpret sustained endo inputs as STM representations of task instructions in PFC.

We use a Gaussian shaped input profile to approximate typical tuning curves of neurons, so their firing profiles are well approximated by smooth curves. In the case of transient input, the specific shape of localized input is unimportant because the network dynamic dominates after cessation of input. With sustained input, a Gaussian input profile leads to a good approximation of a Gaussian output profile, achieving the biological realism of our input profile described above.

Finally, we compare our CANN model of SVA to one with no lateral interaction, modelling the latter by simply adding together its Gaussian inputs.

3 Simulations

Muller et al. (2003) (Muller et al., 2003) provide evidence for sustained division of visual attention by recording steady state visual evoked potentials (SSVEP) while subjects viewed a horizontal array of four stimulus elements following instructions to attend to two locations. On separate blocks of trials, subjects attended to adjacent and separated positions. The SSVEP is the electrophysiological response in visual cortex to a rapidly flickering stimulus, and has been shown to increase in amplitude when attention is paid to the

location of the stimulus (Muller et al., 2003). They found that SSVEPs were lower at the location between separated targets in a detection task. Additionally, they showed that split locations were attended just as well as adjacent locations in their experiment.

We model these experiments in Simulations 1 and 2, however, we widen the network weight profile from our earlier work, increasing σ_w from 0.4 to 1.2 and C from 0.1 to 0.3. This change results in an increase in the width of a stable post-stimulus bubble from $\sigma = 4 * a$ to $\sigma = 8 * a$, demonstrating that our findings are robust in this respect.

Awh and Pashler (2000) (Awh & Pashler, 2000) use a partial report procedure to test subjects' ability to divide spatial attention. Subjects viewed a 5x5 array of alpha-numeric characters containing 23 letters and 2 digits. Subjects fixated a central location before the presentation of two cues, either side of fixation, indicating the probable location of the digits. The character array was subsequently presented, and the subjects' task was to identify the digits. During eighty percent of trials, digits appeared at the cued (valid) locations. During the remaining twenty percent of trials, digits appeared either side of fixation in the orthogonal direction. Thus, on invalid trials, one of the digits appeared directly between the cued locations. Performance at the cued and intervening locations was compared. To the extent that SVA can be divided, subjects should perform better at the cued locations than in the middle. If division of attention were perfect, performance on the two unattended locations would be equal. Subjects' ability to divide SVA was found to depend on the presence of a subsequent noise mask, but the removal of array noise alone was sufficient to significantly reduce division of attention, regardless of subsequent masking. We model this work in Simulation 3.

3.1 Simulation 1

Adopting Müller's terminology, we refer to the locations of stimuli as 1, 2, 3 and 4, where 1 is the left-most location and 4 is the right-most location (Figure 1A). A 1+2 trial refers to trials in which subjects were instructed to direct their attention to locations 1 and 2, a 2+4 trial refers to instructions to attend to locations 2 and 4, and so forth for other combinations of the four locations.

In these transient-input trials, we give the network exo and endo inputs for 300 iterations of dt/τ, simulating the changing symbols in the visual field and task instructions respectively. This input activity is followed by

300 iterations without either source of input. These iterations are sufficient for the network to stabilize under both dynamic regimes (both during and after input). In adjacent trials (1+2 and 3+4) and split trials (1+3 and 2+4), network activity merges into a single winning bubble between target locations, predicting a unitary focus of attention. These results are shown in Figure 1.

In comparison to our earlier study, the wider activity profile predicts a more even distribution of attention once two bubbles merge into one. Specifically, in the 1+2 trial, activity at locations 1 and 2 is 67% and 90% of maximum activity respectively, compared to 21% and 37% in Standage et al. (2005a) (D. I. Standage et al., 2005). The wider activity profile also effects split trials. The bubble drifts into the area between attended locations, shown in Figure 1F and G. Both WTA effects conflict with Müller's findings.

3.2 Simulation 2

Network configuration and the shape and location of inputs is identical to Simulation 1. Exo and endo inputs are sustained simultaneously for 500 iterations, sufficient for the network to stabilize.

Under sustained inputs, our model replicates Müller's findings in split trials, as network activity is greater at locations 1 and 3 than in between (Figure 2B). Two distinct bubbles are also seen in adjacent trials (Figure 2A) suggesting that Müller's subjects may have divided their attention between adjacent locations. Because Müller et al. didn't test subjects' attention between adjacent stimuli, this effect doesn't conflict with their results. Having found similar results with a narrow weight profile in Standage et al. (2005a) (D. Standage et al., 2005), results here show that sustained inputs dominate the network regardless of the width of its weight profile.

To achieve a single bubble in adjacent trials, we increase the overlap between representations of input stimuli from $\sigma_{ext} = 0.3$ to $\sigma_{ext} = 0.5$. The model predicts divided attention in split trials (Figure 2D, but no longer in adjacent trials (Figure 2C).

Because network output so closely resembles the shape of sustained inputs, we investigate the contribution of the network dynamic to the output profile. In the extreme case, complete neglect of the network dynamic reduces the model to a simple addition of Gaussian (AOG) input curves. To test if an AOG provides a model of CANN behaviour under sustained inputs, we measure the reduction or 'dip' in activity between bubbles at different distances

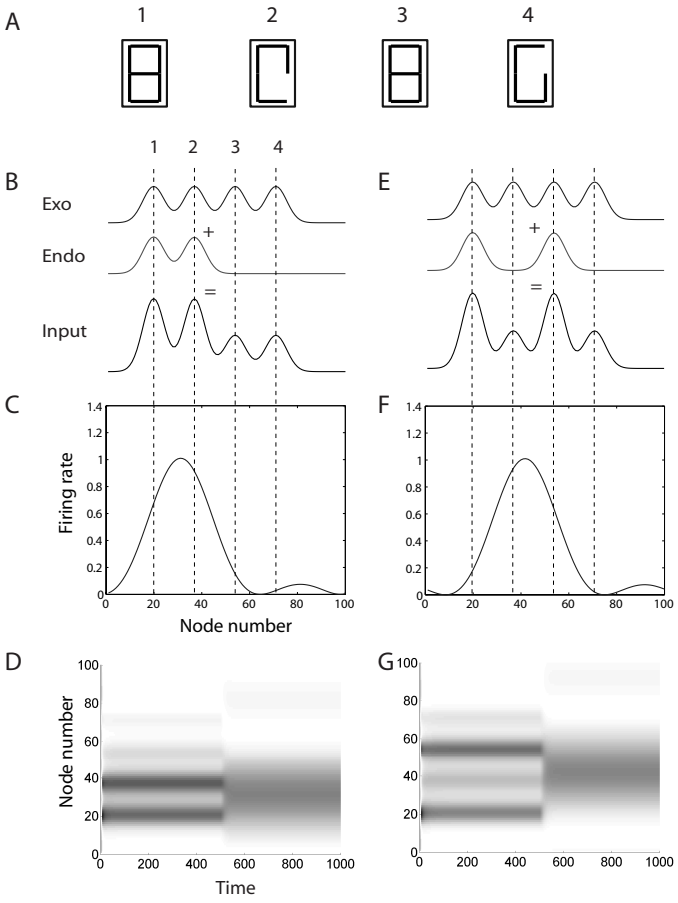

Figure 1: (A) Subjects in Müller's experiment attended two of four horizontal locations, here labelled 1, 2, 3 and 4. Rectangles were flashed at different rates, creating SSVEPs. Random sequences of five symbols were provided at all locations. Subjects' task was to report simultaneous occurrence of a target symbol at two attended locations. The figure depicts a 1+3 trial with target symbol '8'. (B) Nodes 20, 37, 54 and 71 correspond to locations 1, 2, 3 and 4 respectively. Exo inputs are applied to all locations. Endo inputs are applied to locations 1 and 2 only. Combined exo and endo input activity shown on bottom. Gaussian width factors $\sigma_w = 1.2, \sigma_{ext} = 0.3$, constant of inhibition $C = 0.3$. Dashed vertical lines run through target locations. (C) Stable bubble following transient input. The bubble is centred on node 31, reflecting the merge between locations 1 and 2. (D) Network activity over time. Input is stopped after 500 iterations, followed by transition to a one bubble (merged) state by approximately 550 iterations. (E) All parameters are identical to B except endo inputs are applied to locations 1 and 3. (F) Stable bubble following transient input, centred on node 42 (location 3). (G) Network states over time. A merged bubble is stable by approximately 550 ms.

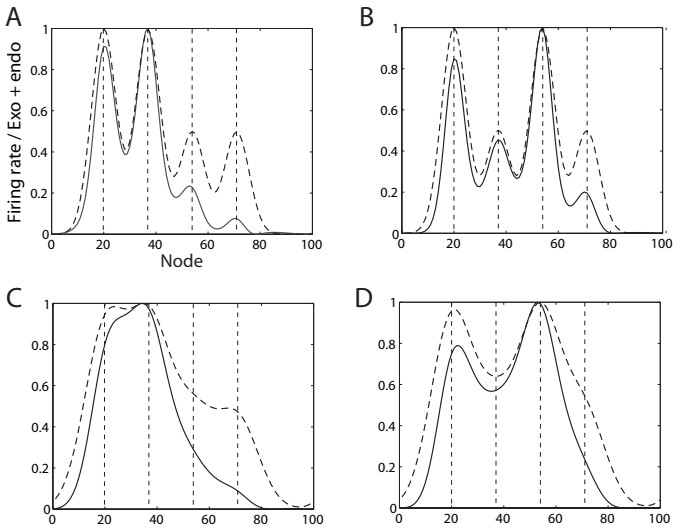

Figure 2: Network configuration as described in Figure 1B. Activity conforms to the sustained input profile, sharpened by lateral inhibition. For clarity, input (dashed) and output (solid) are normalized to 1. Dashed vertical lines show target locations. (A) 1+2 trial, $\sigma_{ext} = 0.3$ (B) 1+3 trial, $\sigma_{ext} = 0.3$. (C) 1+2 trial, $\sigma_{ext} = 0.5$. Wider input profile abolishes divided attention in adjacent trials. (D) 1+3 trial, $\sigma_{ext} = 0.5$. Under sustained wide inputs, the CANN model still predicts divided attention in split trials.

between inputs, comparing it to peak activity in the bubbles. Correspondingly, we measure the height of the midpoint between two Gaussian curves as a function of the distance between them, comparing it to their maxima. The solid line in Figure 3A represents the CANN model. The dotted line represents AOG. Both curves predict unitary attention when targets are spatially proximal ($\approx$ 10 nodes) and divided attention between more distant targets ($\approx$ 30 nodes). In between, the slight difference between curves reflects the effect of lateral inhibition, reminiscent of Mountcastle's two-point discrimination (Mountcastle & Darian-Smith, 1968). In contrast, the effect of local excitation is not evident in the figure, as the onset of a divided activity is not right-shifted for the CANN curve.

Figure 3A shows results of our analysis for only one value of σ_{ext}. Figure 3B shows how the AOG curve in A depends on the width of Gaussian inputs. The same effect is observed for the CANN. Our model predicts that attention cannot be divided at close distances, but without a means to map σ_{ext} to physical parameters, we do not predict specific distances over which attention may be divided.

3.3 Simulation 3

We simulate Awh and Pashler's experiments in 2D. In keeping with our 1D simulations above, we use a wide weight profile in our 2D model ($\sigma_w = 1.3, C = 0.1$). The effects of a sustained input profile are similar to the 1D case. In the 2D CANN model, the number and location of inputs are arranged to reflect Awh and Pashler's xperimental conditions (described in Figure 4). Inputs are centred on every fourth node (horizontally and vertically) in a square bounded by nodes (7,7) and (23,23) in the 900-node 2D network. Exo inputs at these locations represent the character array in Awh and Pashler's Experiments 1, 4 and 4a. In all trials, endo inputs centred on nodes (11,19) and (19,19) represent subjects' attention to the probable location of the digits. In valid trials, digits are represented by exo inputs centred on these nodes. In invalid trials, digits are represented by exo inputs centred on nodes (15,11) and (15,19). Subjects' fixation is represented by an exo input centred on node (15,15). In these simulations, we model Awh and Pashler's experiments only as far as the presentation of the character array. The effect of noise masks and subsequent identification of target digits presumably involve STM and object recognition processing not included in our model.

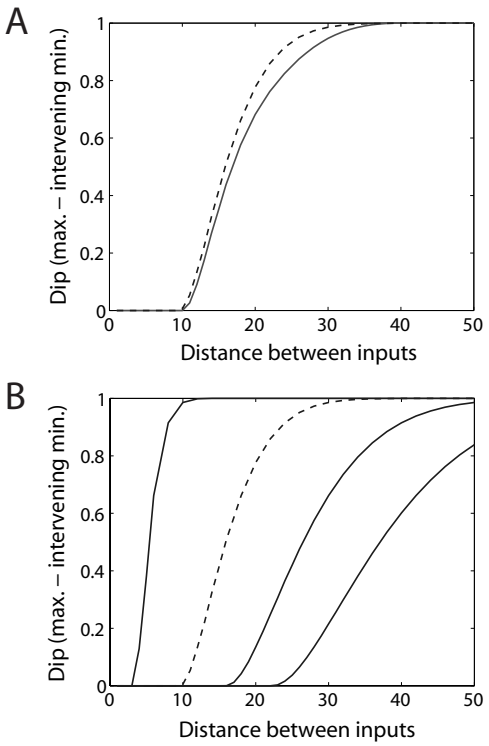

Figure 3: (A) Dip between CANN bubbles (max. - intervening min.) plotted against the distance between peaks (solid line). $\sigma_w = 0.8, \sigma_{ext} = 0.3, C = 0.3$. Dip between peaks of summed inputs, plotted against distance between them (dashed line). (B) AOG distance vs. dip (as in A). From left to right, $\sigma_i = 0.1, 0.3, 0.5, 0.7$. Dashed line shows curve in A.

To model Awh and Pashler's Experiment 1, we provide an exo input to the fixation point for 500 ms. This exo signal is then accompanied by endo inputs to target locations for 750 ms. Finally, exo inputs are centred on all character locations and endo inputs are continued for 118 ms, where $dt/\tau = 1$ ms. These iterations model the duration of input screens in Awh and Pashler's study.

Our results replicate those of Awh and Pashler's Experiment 1. The sustained endo inputs dominate the network, facilitating the activity profile shown in Figure 4C.

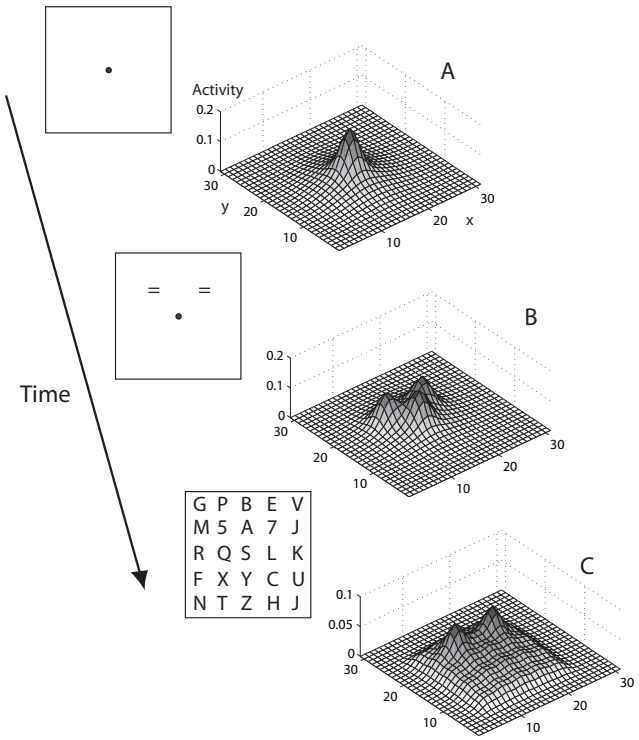

Figure 4: Simulation of Awh and Pashler's Experiment 1. Subjects fixate on the central dot for 500 ms, then fixate on the dot *and* attend to the 'equals' signs for 750 ms before presentation of a 5x5 character array for 118 ms. On 80% of trials, digits appear at the attended locations (valid trials). On invalid trials, digits appear at the locations shown in Figure 5B. Here, the full character array is presented on all trials. The 2D CANN's response to corresponding input signals is shown on the right.

Our simulation of Awh and Pashler's Experiments 4 and 4a uses the same network configuration and width and duration of inputs as Experiment 1, except the character array is provided for only 62 ms, as in Awh and Pashler's Experiment 4.

We simulate removal of non-target characters during valid trials by removing all exo inputs except at attended locations. Similarly, we simulate removal of non-target characters during invalid trials by limiting exo inputs to the middle and far locations. Results of our invalid trials (Figure 5B) coincide with those of Awh and Pashler. The middle location is more active than the attended locations. Results of our valid trials (Figure 5A) paint a different picture. The absence of letter noise reduces competition between stimuli in the character array, so that activity at attended locations is more cleanly divided than in our simulation of Experiment 1. Our interpretation of these results is that the probe stimulus may have been responsible for the reduction of divided attention in Awh and Pashler's Experiment 4 and its abolition in their Experiment 4a.

4 Discussion

In our simulations of Müller's experiments, we address a parametric issue raised in Standage et al. (2005a) (D. I. Standage et al., 2005). By increasing the width of the model's weight profile, and comparing network output with similar experiments in Standage et al. (2005a) (D. I. Standage et al., 2005), we show that under transient input, the model outputs a flatter bubble, predicting that the magnitude of subjects' attention in Müller's adjacent trials was more evenly distributed across attended locations than predicted in Standage et al. (2005a) (D. I. Standage et al., 2005). This input/output paradigm still supports a unitary gradient model, but with a much less extreme slope.

In Simulation 2, we show that under sustained inputs, network output capitulates to input, regardless of the width of our weight profile. Under this paradigm, we no longer need the CANN to account for divided attention. A simple AOG model suffices when inputs are Gaussian shaped. In this case, our ability to model divided SVA is mediated by the width and proximity of input activity. Because inputs to PP are highly pre-processed, coming from PFC and V4 in Deco's model (in addition to less pre-processed input from V1), we believe these inputs may be characterized by considerable overlap,

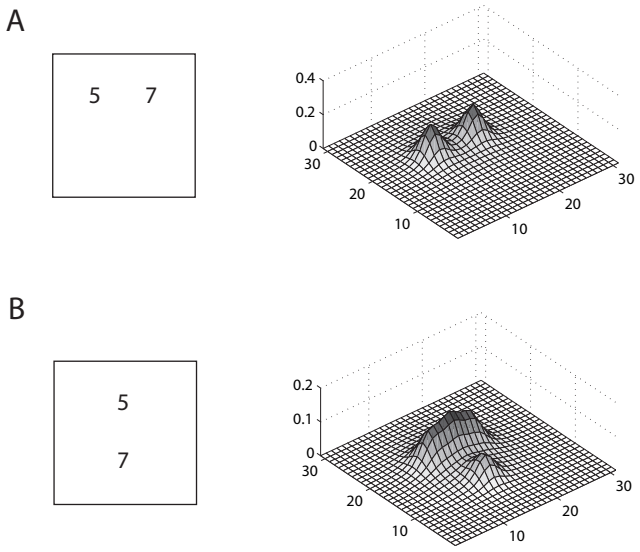

Figure 5: Simulation of Awh and Pashler's Experiments 4/4a. Awh and Pashler's partial character arrays are shown on the left (valid array top, invalid array bottom). The 2D CANN predicts divided attention on valid trials (top right). On invalid trials (bottom right), the model predicts unitary attention over the target and middle locations, with less attention focused on the 'far' vertical location. This output coincides with Awh and Pashler's results on invalid trials.

given the expansion of receptive fields in hierarchical processing. We conjecture that the overlap between integrated object representations may increase as a function of their number of common features. For example, neural representations of a red circle and a red square may overlap more than those of a red circle and a blue square, and attention may be more difficult to divide between them.

Simulation 2 predicts that attention should be easier to divide as foci become more distant (within a reasonable visual area). This prediction is parameter-dependent, as a stiffer gain function would still predict divided attention, but in all-or-none fashion. With a Sigmoid gain function, the model no longer resembles an AOG. These predictions could be tested by a probe stimulus between adjacent locations in Müller's experiment, and the addition of a 1+4 trial. Such experiments are important to further constrain computational models of attention. Additionally, although Müller et al. recorded target detection rates (TDR) in their experiment, the subjects' task was to detect simultaneous occurrence of the target symbol at the attended locations only. TDRs at adjacent locations were no better than at split locations (indeed, they were slightly worse), but without testing subjects' ability to detect simultaneous occurrence of target symbols at *any* two locations while subjects attended to two specific locations, TDRs provide no direct psychophysical evidence of divided attention. For example, we don't know that TDRs would be better at positions 1 and 3 during a 1+3 trial than at position 2. We believe that adding a task requiring subjects to detect simultaneous occurrence of a target symbol at any two locations during the same trial block would strengthen Müller's conclusions.

A parameter largely unexplored in these simulations is the strength of connectivity in the network. By greatly reducing input strength in comparison to the network's connection strength, connectivity dominates sustained input and the WTA nature of the model re-emerges. It is also likely that the strengths of exo and endo inputs are not equal. If we assume a model that tends to WTA in the general case, that is, that attention is unitary 'by default', and that strong endo inputs are able to override this tendency in unusual cases, then perhaps exo inputs to the model should be weak and wide, and endo's should be strong and narrow. In this regard, our model suggests that an understanding of the modulation of signals from PFC to PP is crucial to understanding SVA.

Awh and Pashler found the ability to divide SVA was greatly reduced following removal of noise surrounding target stimuli. We believe the model

can account for this result if their probe stimuli on invalid trials dominated voluntary attention. That is, in the absence of competition from letter noise in the character array, exo signals at the invalid locations dominated subjects' attention. In valid trials, exo and endo signals were directed to the same locations, so attention may have been divided when Awh and Pashler weren't testing for it, only to be unified by the testing procedure.

Simulations 2 and 3 show that the CANN model is able to account for divided SVA under sustained inputs. As such, we believe that divided SVA may be possible for as long as endo and/or exo signals are provided to PP, and that differences in behavioural findings may reflect differences in experimental conditions rather than subjects' ability to divide their attention. The nature of these conditions is largely unexplored. This conjecture echoes that of Schneider (Schneider, 1998) that different experimental paradigms may facilitate measurements of different attention-related phenomena.

Because we interpret endo inputs as subjects' representations of task instructions in WM, our model predicts that interference with STM should abolish split attention in both Müller's and Awh and Pashler's experimental conditions. This prediction could be tested in a dual task paradigm. By equating task instructions in behavioural studies with STM representations in WM, and by modelling these representations as sustained endo inputs to the CANN, we revisit the relationship between WM and attention. Attention has often been cited as the primary constraint on WM capacity (Cowan, 2001), but here we view WM representations as the driving force behind attention.

Our focus has been on stable attractor states in this paper. Transitions between dynamic regimes tend to be rapid, and given the large number of parameters that effect the model, stable states provide a better foundation for our simulations. As we show in Standage et al. (2005a) (D. I. Standage et al., 2005), parameter adjustments effect transitions between regimes. A possible explanation of the findings of Müller et al. and Awh and Pashler is that divided SVA corresponds to the transition between two-bubble and one-bubble states in a WTA model. Thus, subjects may only be able to divide attention during these meta-stable states; given sufficient time for the network to settle, subjects may be unable to divide their attention.

5 Conclusions

The model of SVA by Deco et al. (2002) (Deco et al., 2002) implements a saliency map in PP with a CANN network. This instantiation of biased competition (Duncan & Humphreys, 2002) integrates BU and TD influences in a biologically realistic computational architecture. Our simulations test this promising model's ability to explain behavioural and physiological evidence on the spatial distribution of SVA.

Our results demonstrate that CANNs provide a model of spatial attention in PP capable of explaining divergent experimental findings. With transient inputs, the model's WTA nature predicts a unitary attentional focus. With sustained inputs, the model accounts for divided SVA. As such, our predictions depend on the nature of exo and endo signals in attentive phenomena. Here, the use of sustained inputs replicates the findings of Muller et al. (2003) (Muller et al., 2003) and some of the findings of Awh and Pashler (2000) (Awh & Pashler, 2000). Where Awh and Pashler interpret unitary attention on invalid trials as demonstrating unitary attention on valid trials, we believe subjects' attention may have been divided on valid trials, only to be unified by their probe stimulus.

The interplay between WM and SVA is paramount to our model. If divided attention is facilitated by STM representations providing endo inputs to PP, then disruption of STM should abolish divided attention. We believe further research in this area would improve our understanding of the relationship between WM and attention.

Acknowledgment

This work was supported in part by the NSERC grant RGPIN 249885-03.

References

Awh, E., & Pashler, H. (2000). Evidence for split attentional foci. *Journal of Experimental Psychology, Human Perception and Performance*, *26*(2), 834-846.

Cherry, E. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the Acoustical Society of America*, *25*, 975-979.

Colby, C. L., & Goldberg, M. E. (1999). Space and attention in parietal cortex. *Annual Review of Neuroscience*, *22*, 319-349.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87–114.

Deco, G., Pollatos, O., & Zihl, J. (2002). The time course of selective visual attention: theory and experiments. *Vision Research*, *42*, 2925-2945.

Deneve, S., Pouget, A., & Latham, P. E. (1999). Divisive normalization, line attractor networks and ideal observers. *Advances in Neural Processing Systems*, *11*, 104–110.

Downing, C., & Pinker, S. (1985). The spatial structure of visual attention. In M. Posner & O. Marin (Eds.), *Attention and performance* (Vol. 11, p. 171-187). Erlbaum.

Duncan, J., & Humphreys, G. (2002). Visual search and stimulus similarity. *Psychological Review*, *96*, 433-458.

Eriksen, C., & James, J. S. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception and Psychophysics*, *40*, 225-240.

Funahashi, S., Bruce, C., & Goldman-Rakic, P. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of Neurophysiology*, *61*, 331–349.

Hahn, S., & Kramer, A. (1998). Further evidence for the division of attention between non-contiguous locations. *Visual Cognition*, *5*, 217-256.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews: Neuroscience*, *2*, 194-203.

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, *4*(4), 219–227.

LaBerge, D., & Brown, V. (1989). Theory of attentional operations in shape identifacation. *Psychological Review*, *96*, 101-124.

McCormick, P., Klein, R., & Johnston, S. (1998). Splitting vs. sharing focal attention: comment on castiello and umilta (1992). *Journal of Experimental Psychology*, *24*(1), 350–357.

Mountcastle, V., & Darian-Smith, I. (1968). Neural mechanisms of somethesia. In V. Mountcastle (Ed.), *Medical physiology* (12th ed., pp. 1372–1423). St. Louis: Mosby.

Muller, M., Malinowski, P., Gruber, T., & Hillyard, S. (2003). Sustained division of the attentional spotlight. *Nature*, *424*.

Neisser, U. (1967). *Cognitive psychology.* New York: Appleton, Century, Crofts.

Posner, M., Snyder, C., & Davidson, B. (1980). Attention and detection of signals. *Journal of Experimental Psychology, General*, *109*, 160-174.

Schneider, W. X. (1998). An introduction to 'mechanisms of visual attention: a cognitive neuroscience perspective'. *Visual Cognition*, *5*, 1-8.

Shipp, S. (2004). The brain circuitry of attention. *Trends in Cognitive Sciences*, *8*(5), 223–230.

Standage, D., Trappenberg, T., & Klein, R. (2005). A continuous attractor neural network model of divided visual attention. *Proceedings of the International Joint Conference on Neural Networks, July 31 - August 4, 2005, Montreal, Canada.*

Standage, D. I., Trappenberg, T. P., & Klein, R. M. (2005). A continuous attractor neural network model of divided visual attention. In *Proceedings of the international joint conference on neural networks 2005* (pp. 2897–2902).

Trappenberg, T. P., & Standage, D. I. (2005). Multi-packet regions in stabilized continuous attractor networks. *Neurocomputing*, *65–66*, 617-622.

Treisman, A. (1998). Feature binding, attention and object perception. *Proceedings of the Royal Society.*

Tsotsos, J. K. (1992). On the relative complexity of active vs. passive visual search. *International Journal of Computer Vision*, *7*, 127-141.

Wolfe, J. (1994). Guided search 2.0: a revised model of visual search. *Psychonomic Bulletin and Review*, *1*(2), 202-238.

Zelano, C., Bensafi, M., Porter, J., Mainland, J., Johnson, B., Bremner, E., ... Sobel, N. (2004). Attentional modulation in human primary olfactory cortex. *Nature*, *8*, 114-120.