

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

A Recurrent Model for Spatial Attention

Permalink

<https://escholarship.org/uc/item/6c68b72t>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 27(27)

ISSN

1069-7977

Authors

Casarotti, Marco

Umilta, Carlo

Zorzi, Marco

Publication Date

2005

Peer reviewed

A Recurrent Model for Spatial Attention

Marco Casarotti (marco.casarotti@unipd.it)

Department of General Psychology; 8 Via Venezia
Padova, PD 35131 ITALY

Marco Zorzi (marco.zorzi@unipd.it)

Department of General Psychology; 8 Via Venezia
Padova, PD 35131 ITALY

Carlo Umiltà (carlo.umilta@unipd.it)

Department of General Psychology; 8 Via Venezia
Padova, PD 35131 ITALY

Abstract

The premotor theory of attention maintains that visuospatial attention originates from the activation of the same cortical circuits involved in saccadic planning. Attention orienting is supposed to be achieved by means of recurrent projections from premotor areas to spatial maps without involving specific modules separated from the circuits which transform sensory information into motor plans. In the present study, the basic claim of the premotor theory was tested by implementing a recurrent neural network model in which spatial attention is explicitly concerned with action-oriented representations and it is simulated in terms of feedback effects due to saccadic planning. The model employs basis function units which simulate parietal neurons involved in the representation of the oculomotor space. Simulation results bring computational evidence to the premotor theory and allow to make novel and testable predictions.

Keywords: Spatial attention; Spatial representations; Neural networks.

Introduction

The premotor theory of attention (Rizzolatti et al., 1987; Rizzolatti, Riggio & Sheliga, 1994; Umiltà et al., 1991) maintains that visuospatial attention originates from the activation of the same cortical circuits involved in saccadic planning. The premotor theory affirms that the preparation of a saccadic movement produces a processing facilitation for stimuli located in the region of space towards which the motor program is prepared. This processing facilitation is supposed to be achieved by means of recurrent projections from premotor areas to spatial maps without involving specific modules separated from the circuits which transform sensory information into motor plans.

Neurophysiological data strongly support the premotor theory indicating that cortical neurons located within the intraparietal sulcus (IPs) generate action-oriented representations of space for motor planning and are also crucially involved in the top-down (endogenous) control of spatial attention (Colby & Goldberg, 1999). Specifically, the neurons located in the lateral intraparietal area (LIP) show an increased firing rate when the monkey anticipates the onset of a stimulus. Moreover, they do not represent all objects of the visual field, but only salient targets, suggesting that they form a salience map of the visual world (Goldberg et al., 2002). Neural activity in parietal spatial maps describes the

spatial and temporal dynamics of attention (Bisley & Goldberg, 2003).

Neuroimaging studies showed that top-down control of spatial attention in humans recruits the IPs and the frontal eye fields (FEF) (Corbetta & Shulman, 2002, for review). That is, the network of brain regions involved in endogenous visuospatial attention strongly overlaps with the network subserving sensorimotor transformations for overt saccadic eye movements (Corbetta et al., 1998; Nobre et al., 2000).

Previous computational models concerned with spatial attention employed a separate subsystem or set of nodes to generate attentional effects. For instance, an "Attentional mechanism" (Mozer, 1991) or "Attention units" (Cohen et al., 1994) activated the relevant representations in the relevant module by means of ad-hoc connections. As noted in a review by Heslenfeld et al. (1997), in most of the (non-implemented) cognitive models (e.g., LaBerge, 1990; Treisman, 1988; Van der Heijden, 1992) or (implemented) connectionist models (e.g., Cohen, et al., 1994; Mozer, 1991; Phaf, Van der Heijden, & Hudson, 1990), attention is viewed as an additional resource without specifying where it comes from and how exactly the to-be-activated units are found in the system.

In the present study we propose a recurrent neural network model in which spatial attention is explicitly concerned with action-oriented representations and it is simulated in terms of feedback effects depending on saccadic planning. The model is based on the basis function approach to simulate the computational properties of the parietal neurons. Moreover, it includes a circuit that allows the system to update remembered spatial locations in eye-centred coordinates after intervening saccades.

As an introduction to the simulations, we briefly review neurophysiological data regarding the neural encoding of space.

Spatial representations in the PPC

The activity of many LIP neurons approximates a multiplicative combination of visual and posture signals (Andersen, Essick & Seigel, 1985; Andersen, 1989). The receptive field of a typical LIP cell is retinotopic like those found in V1, but the amplitude of the response evoked by a visual target increases monotonically as the eye moves along a particular direction, which is specific to each neuron. Cell encoding with multi-

plicative interaction of independent variables is called gain-field coding. Pouget and collaborators (Pouget & Sejnowski, 1997; Pouget & Snyder, 2000) suggested that parietal gain-modulated neurons may serve as basis functions with which the brain computes sensorimotor transformations. A population of these neurons may provide a complete basis set for the combined space allowing to approximate any arbitrary function of the input variables by taking a linear sum of their outputs (Poggio & Girosi, 1990). It follows that encoding space with basis function units renders it possible to reduce non-linear coordinate transformations to simple linear mappings. The resulting basis function representation codes spatial locations in a format which contains implicitly any frame of reference that can be derived from the input variables. For instance, a basis function map combining visual information with eye position may be used to compute target location in head-centred coordinates with a linear combination of a specific set of weights.

Retinotopic representations, however, are not sufficient to store spatial locations, because the remembered spatial location will not be in register with the external space after an eye shift (Miller & Bockisch, 1997). When the task requires to foveate a remembered location and the eyes move during the memory period, the saccadic system must take into account the change in eye position. Primates can easily and accurately perform saccadic tasks in which the retinal error is dissociated from the motor error, such as in the case of ocular perturbation by electrical stimulation (Mays & Sparks, 1980) and the double-saccade task (Goldberg & Bruce, 1990, Mazzoni et al., 1996). Mays and Sparks showed that a perturbation in the eye position, induced by electrical stimulation of the superior colliculus during the preparation of a saccade toward a visual target, is compensated by the saccadic system even when the target is no longer visible. In the double-saccade task, two visual targets are presented sequentially and very briefly. The task requires to direct the gaze at targets in order of presentation. When the first saccade is performed, the retinal coordinates of the second target do not match the motor coordinates required to foveate it. Thus, the system must take into account the displacement of the eye due to the first saccadic movement.

Two alternatives corresponding to different encoding schemes have been proposed to explain how the brain performs this computation (Bremmer, Pouget & Hoffmann, 1998). The first hypothesis is based on extraretinal encoding by means of a coordinate transformation. For instance, the position of the eye can be combined with the retinal coordinates of the target in order to obtain a head-centred representation, which is invariant with respect to eye movements. The second hypothesis is based on a perisaccadic updating mechanism. It has been shown that spatial representations within area LIP are remapped in the coordinates of the new fixation point after an intervening saccade (Duhamel, Colby & Goldberg, 1992), but the source and the properties of the signal that drives the updating process are still not clearly understood.

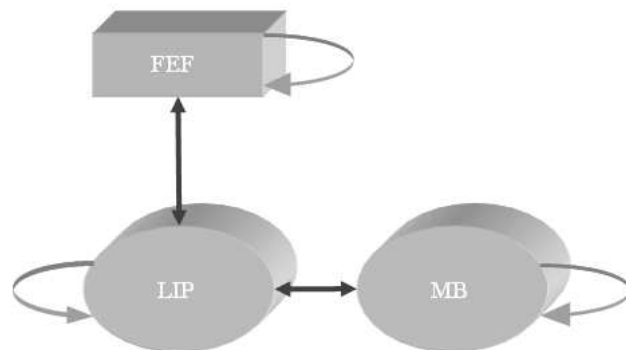


Figure 1: Recurrent neural network model for saccadic planning and spatial attention using basis functions. The model contains two basis function maps reciprocally connected: an input map (LIP) composed of basis function units which combine visual and posture information, and a memory buffer (MB) which stores target location in head-centred coordinates modulated by eye position. The output layer contains an oculomotor map (FEF). Each map has lateral connections with local excitation and long range inhibition.

Simulations

Description of the model

The architecture of the model (Figure 1) consists of two basis function maps which simulate the activation of gain modulated neurons described in the PPC and an output layer consisting of a computational motor map which is meant to simulate the activity within the FEF.

The first basis function map (LIP) is composed of 847 units which generate a representation of the oculomotor space by combining multiplicatively retinal input (r) with eye position (e) encoded in head-centred coordinates (Figure 2):

$$x^{LIP} = G(r)S(e)$$

The tuning curve of visual response is a Gaussian function of target location in retinal coordinates (r_x and r_y):

$$G(r)_i = e^{-\frac{(r_x - r_{xi})^2}{2\sigma_{xi}^2}} e^{-\frac{(r_y - r_{yi})^2}{2\sigma_{yi}^2}}$$

where r_{xi} and r_{yi} indicates the centres of the visual receptive fields (RFs), uniformly spread between -40° and 40° in increments of 8° , both in the horizontal and vertical dimensions. The width of the Gaussians, σ , changes linearly with the horizontal and vertical eccentricity of the units' RF according to the following equation: $2.491 + 0.439 * Eccentricity$ (Platt & Glimcher, 1998). As a result, the response field radius increases by nearly 0.5° for every degree increase in eccentricity. Gaussian functions optimally fit intraparietal neurons accounting for most of the variance in their activity (Gnadt & Breznen, 1996; Platt & Glimcher, 1998).

Eye position is encoded with a sigmoid function:

$$S(e)_i = \frac{1}{1 + \exp^{-\frac{e - e_i}{s}}}$$

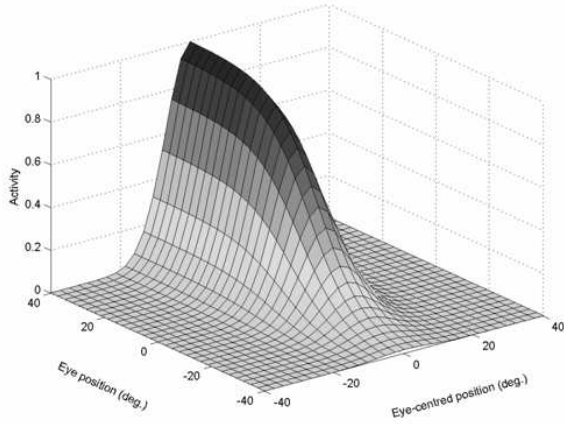


Figure 2: Activity of a basis function unit which computes the product of a Gaussian of target location (eye-centred position) multiplied by a sigmoid function of eye position.

where e is the current horizontal eye position and e_i is the inflection point of the function, spread between -24° and $+24^\circ$ in steps of 8° . The slope of the sigmoid, s , is set to 8.

LIP units are reciprocally connected with another basis function map (MB, memory buffer) which stores target location in head-centred coordinates modulated by eye position. Connections from LIP units to MB units allow the system to transform retinal coordinates into head-centred coordinates, while feed-back connections from MB units to LIP units allow system to achieve the opposite transformation, from head-centred to eye-centred coordinates. As a result, the extraretinal encoding in MB renders it possible to achieve perisaccadic updating in LIP after intervening saccades.

LIP units send connections also to the output layer that contains an oculocentric map composed of 11×11 units covering a $40^\circ \times 40^\circ$ space with 8° spacing. The activation of the output units is given by the linear sum of the weighted signals coming from the LIP units.

For sustaining activity in the absence of visual input, we introduced in each map of the recurrent model lateral connections that generate local excitation and long range inhibition. This kind of connectivity allows to maintain the hill of activity within a computational map over time after the stimulus disappears and has been used in a number of computational models of cortical activity (e.g., Somers, Nelson & Sur, 1995; Zhang, 1996; Salinas, 2003).

The dynamics of the model are captured by the following equations:

$$\frac{dx_i^{LIP}}{dt} = (G(r) + \sum_m w_{im} a_m^{MB} + \sum_j v_{ij} x_j^{LIP} - kx_i^{LIP})S(e)$$

$$\frac{dx_i^{MB}}{dt} = (\sum_m w_{im} a_m^{LIP} + \sum_j v_{ij} x_j^{MB} - qx_i^{MB})S(e)$$

$$\frac{dx_i^{OUT}}{dt} = \sum_m w_{im} a_m^{LIP} + \sum_j v_{ij} x_j^{OUT} - qx_i^{OUT}$$

where x_i is the activation of the unit i , w_{im} is the synaptic weight coming from the unit m in the connected map, v_{ij} is the lateral connection coming from the neighbouring neuron j , k and q are passive decays of activation.

Synaptic weights and learning procedure

Each set of weights in the model renders it possible to read out a specific frame of reference (*postsynaptic representation*) from the activity within a basis function map (*presynaptic representation*). The training procedure consisted of repeatedly presenting input patterns to the network, propagating the activation to the postsynaptic units and adjusting the weights by using the delta rule (Widrow & Hoff, 1960):

$$d\bar{w}_{ij} = \eta(a_i^* - a_i)a_j$$

where $d\bar{w}_{ij}$ is the change in the weight connecting the presynaptic unit j with the postsynaptic unit i , a_j is the activation of the presynaptic unit j , a_i^* is the expected postsynaptic activity and a_i is the predicted postsynaptic activity. The learning rate, η , was set to 0.01.

The training set consisted of 330 patterns generated by selecting random visual targets, given initial eye positions. Expected outputs were generated by computing Gaussian functions of target location in the frame of reference corresponding to the postsynaptic representation. Training was repeated until the mean square error between the centre of mass of the expected postsynaptic patterns and the predicted postsynaptic patterns was less than 1° .

The lateral connections were given by:

$$v_{ij} = e^{-\frac{(c_{xi}-c_{xj})^2}{2\sigma_{xi}^2}} e^{-\frac{(c_{yi}-c_{yj})^2}{2\sigma_{yi}^2}} - g e^{-d\frac{(c_{xi}-c_{xj})^2}{2\sigma_{xi}^2}} e^{-\frac{(c_{yi}-c_{yj})^2}{2\sigma_{yi}^2}}$$

where v_{ij} is the weight connecting the unit j to the unit i within a specific map, g and d are constants which control respectively the value and the width of the inhibitory region within the RF of the unit i .

Results

The premotor theory of spatial attention maintains that motor planning generates top-down signals that produce a processing facilitation for stimuli located in the region of space towards which they were prepared.

The basic claim of the premotor theory was tested in the model by implementing a spatial cueing paradigm (Posner, 1980), which requires to detect as fast as possible a visual target presented to the left or the right side of fixation. In endogenous cueing, participants voluntarily orient their spatial attention to the region of visual space indicated by a cognitive cue and the target can be presented on the same side (valid trials) or on the opposite side (invalid trials). In neutral trials attention is not spatially oriented. Typically, valid trials give rise to attentional benefits (faster reaction times with respect to neutral trials), while invalid trials give rise to attentional costs (slower reaction times with respect to neutral trials).

In accordance with the premotor theory, attention orienting was simulated by generating a saccadic plan in the output map of the model and feeding back the activation to LIP units through the same connections involved in sensorimotor transformations for saccadic movements. After a variable number

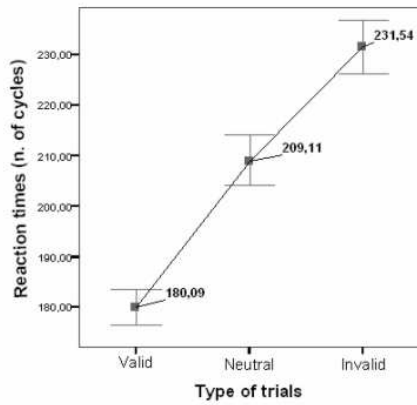


Figure 3: Reaction times for detecting visual targets in a spatial cueing paradigm. Orienting of attention was simulated by computing a saccadic plan in the output map and propagating the activation to the input units using the same connections involved in the sensorimotor mapping. Attention orienting produces reliable benefits (valid trials faster than neutral trials) and costs (invalid trials slower than neutral trials).

of time steps, a visual target was presented to the input units in the same location of the planned saccade (valid trials) or in another location (invalid condition). Neutral trials without attention orienting were used to establish a baseline for the detection task. The number of cycles required to reach a detection threshold in the LIP map was used as an index of the reaction times for detecting the target. In order to collect good statistics, we performed a number of trials by randomly changing the goal location of the planned saccade and the retinal coordinates of the target. An analysis of Variance (ANOVA) was performed on mean reaction times as a function of type of trials (Valid, Neutral, Invalid). Since the analysis of Variance was significant, $F(2, 198) = 39.01, MSE = 1705.44, p < 0.001$, we performed multiple comparison tests indicating that the valid condition produces faster responses (180 cycles) than the neutral condition (209 cycles), which is in turn faster than the invalid condition (232 cycles). These results clearly indicate that attention orienting in our recurrent model of saccadic planning produces reliable benefits and costs (Figure 3) consistent with behavioural data.

Perisaccadic updating and spatial attention

In order to investigate the role of the remapping process in spatial attention, a variant of the spatial cueing paradigm with an ocular perturbation before the onset of the target. The sequence of events was the following: a saccadic plan was first generated in the saccadic map and after a temporal delay an ocular perturbation was simulated by changing the eye position. After 1 or 10 time steps a random visual target was presented to the LIP units and the number of cycles required to reach a threshold value was measured. We performed 25 runs with 100 random valid trials, in which target location corresponded to the remapped motor error of the planned saccade, 100 random invalid trials in which target location corresponded neither with the remapped motor error nor with

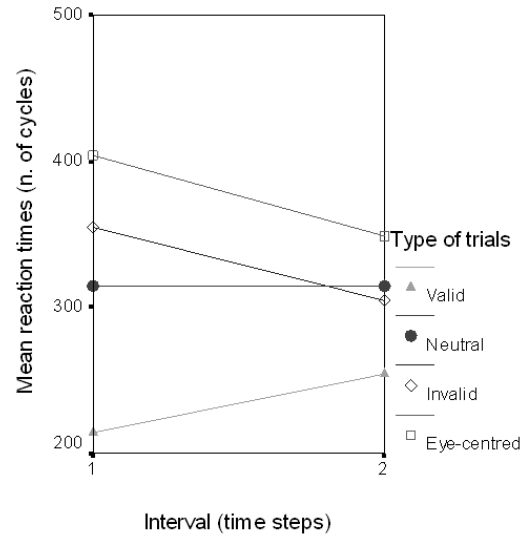


Figure 4: Reaction times for detecting visual targets in a variant of the spatial cueing paradigm with an ocular perturbation (OP). Attention orienting produces reliable benefits and costs, which decrease as the interval between the OP and the target presentation increases. Eye-centered trials are slower than the invalid trials (*remapping interference*).

the motor error, 100 random eye-centred trials in which the retinal coordinates of the target matched the original motor error, and 100 random neutral trials without attention orienting. An analysis of Variance (ANOVA) was performed on mean reaction times as a function of type of trials (Valid, Neutral, Invalid, Eye-centred) and interval between ocular perturbation and target presentation (1 vs. 10 time steps). The main effect of type of trials [$F(3, 72) = 479.25, MSE = 363.47, p < 0.001$] and interval [$F(1, 24) = 21.14, MSE = 622.13, p < 0.001$], as well as their interaction [$F(3, 72) = 49.47, MSE = 524.40, p < 0.001$] were significant (Figure 4). Reaction times were faster on valid trials than neutral trials [$F(1, 24) = 68986.59, MSE = 2.343, p < 0.001$], and slower on invalid trials than neutral trials [$F(1, 24) = 106.05, MSE = 717.41, p < 0.001$], indicating reliable benefits and costs. Both benefits and costs decreased at the longer interval as a consequence of the activation decay. Interestingly, we found that the eye-centred trials, in which the visual target appeared in the same location of the saccadic goal, were slower than invalid trials [$F(1, 24) = 1973.03, MSE = 87.68, p < 0.001$]. This result can be explained considering that after ocular perturbation the updating mechanism moves the attentional code away from the original location interfering with target detection. We will refer to this effect as "*remapping interference*".

Discussion

In the present research we examined whether a recurrent model of saccadic planning can account for attentional effects without requiring additional learning or specific computational mechanisms separated from the sensorimotor circuit. The model employs basis function units which simulate

parietal neurons involved in the representation of oculomotor space. Moreover, it incorporates the updating of remembered spatial locations after intervening saccades by means of coordinate transformations between basis function units.

We tested the model by implementing a spatial cueing paradigm which is a typical task used in behavioural research. The recurrent architecture allowed us to measure the number of cycles required to reach a detection threshold and to compare the simulation results with behavioural data. Simulations showed the typical patterns of results described in behavioural studies, with reliable benefits and costs in a spatial cueing paradigm under conditions of endogenous cueing. Moreover, we examined the role of perisaccadic remapping in attention orienting by implementing a variant of the spatial cueing paradigm in which an eye shift was interposed between attentional allocation and target presentation.

Simulations allowed to make novel and testable predictions: when attention is directed to a particular location with retinal coordinates R_x and R_y and then a saccadic movement is performed toward a different location, the attentional code is fast and efficiently remapped in the coordinates of the new fixation point. As a consequence, attention can be maintained over a specific region of the visual space while moving the eyes. Moreover, presenting a visual target with coordinates R_x and R_y immediately after the end of the saccade gives rise to additional costs with respect to any other retinal location. We called this effect *remapping interference*.

In summary, our simulations are consistent with the premotor theory of attention which maintains that attention and eye movements are tightly coupled, demonstrating that a plausible model of saccadic planning renders it possible to simulate attentional effects described in behavioural research.

References

- Andersen, R. A. (1989). Visual and eye movement functions of the posterior parietal cortex. *Annual Review of Neuroscience*, *12*, 377–403.
- Andersen, R. A., Essick, G. K., & Siegel, R. M. (1985). Encoding of spatial location by posterior parietal neurons. *Science*, *230*, 456–458.
- Bisley, J. W., & Goldberg, M. E. (2003). Neuronal activity in the lateral intraparietal area and spatial attention. *Science*, *299*, 81–86.
- Bremmer, F., Pouget, A., & Hoffmann, K.-P. (1998). Eye position encoding in the macaque posterior parietal cortex. *European Journal of Neuroscience*, *10*, 153–160.
- Cohen, J. D., Romero, R. D., Servan-Schreiber, D., & Farah, M. J. (1994). Mechanisms of spatial attention: the relation of macrostructure to microstructure in parietal neglect. *Journal of Cognitive Neuroscience*, *6*, 377–387.
- Colby, C. L., & Goldberg, M. E. (1999). Space and attention in parietal cortex. *Annual Reviews Neuroscience*, *22*, 319–349.
- Corbetta, M., Akbudak, E., Conturo, T. E., Snyder, A. Z., Ollinger, J. M., Drury, H. A., et al. (1998). A common network of functional areas for attention and eye movements. *Neuron*, *21*, 761–773.
- Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*, 201–215.
- Duhamel, J. R., Colby, C. L., & Goldberg, M. E. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. *Science*, *255*, 90–92.
- Gnadt, J. W., & Breznen, B. (1996). Statistical analysis of the information content in the activity of cortical neurons. *Vision Research*, *36*, 3525–3537.
- Goldberg, M. E., Bisley, J., Powell, K. D., Gottlieb, J., & Kusunoki, M. (2002). The role of the lateral intraparietal area of the monkey in the generation of saccades and visuospatial attention. *Ann. N. Y. Acad. Sci.*, *956*, 205–215.
- Goldberg, M. E., & Bruce, C. J. (1990). Primate frontal eye fields. III. Maintenance of a spatially accurate saccade signal. *Journal of Neurophysiology*, *64*, 489–508.
- Heslenfeld, D. J., Kenemans, J. L., Kok, A., & Molenaar, P. C. M. (1997). Feature processing and attention in the human visual system: An overview. *Biological Psychology*, *45*, 183–215.
- LaBerge, D. (1990). Thalamic and cortical mechanisms of attention suggested by recent positron emission tomographic experiments. *Journal of Cognitive Neuroscience*, *2*, 358–372.
- Mays, L. E., & Sparks, D. L. (1980). Saccades are spatially, not retinocentrically, coded. *Science*, *208*, 1163–1165.
- Mazzoni, P., Bracewell, R. M., Barash, S., & Andersen, R. A. (1996). Motor intention activity in the macaque's lateral intraparietal area. I. Dissociation of motor plan from sensory memory. *Journal of Neurophysiology*, *76*, 1439–1456.
- Miller, J. M., & Bockisch, C. (1997). Visual perception. Where are the things we see? *Nature*, *386*, 550–551.
- Mozer, M.C. (1991). *The perception of multiple objects: a connectionist approach*. MA: MIT Press/Bradford Books.
- Nobre, A. C., Gitelman, D. R., Dias, E. C., & Mesulam, M. M. (2000). Covert visual spatial orienting and saccades: overlapping neural systems. *Neuroimage*, *3*, 210–216.
- Phaf, R. H., Van der Heijden, A. H., & Hudson, P. T. (1990). SLAM: a connectionist model for attention in visual selection tasks. *Cognitive Psychology*, *22*, 273–341.
- Platt, M. L., & Glimcher, P. W. (1998). Response fields of intraparietal neurons quantified with multiple saccadic targets. *Experimental Brain Research*, *121*, 65–75.
- Poggio, T., & Girosi, F. (1990). Regularization algorithms for learning that are equivalent to multilayer networks. *Science*, *247*, 978–982.
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*, 3–25.
- Pouget, A., & Sejnowski, J. T. (1997). Spatial transformations in the parietal cortex using basis functions. *Journal of Cognitive Neuroscience*, *9*, 222–237.
- Pouget, A., & Snyder, L. H. (2000). Computational approaches to sensorimotor transformations. *Nature Neuroscience*, *3*, 1192–1198.

- Rizzolatti, G., Riggio, L., Dascola, I., & Umiltà, C. (1987). Reorienting attention across the horizontal and vertical meridians: evidence in favor of a premotor theory of attention. *Neuropsychologia*, *25*, 31–40.
- Rizzolatti, G., Riggio, L., & Sheliga, B. M. (1994). Space and selective attention. In C. Umiltà & M. Moscovitch (Eds.), *Attention and Performance XV*. Cambridge: MIT Press.
- Salinas, E. (2003). Self-sustained activity in networks of gain-modulated neurons. *Neurocomputing*, *52*, 913–918.
- Somers, D. C., Nelson, S. B., & Sur, M. (1995). An emergent model of orientation selectivity in cat visual cortical simple cells. *Journal of Neuroscience*, *15*, 5448–5465.
- Treisman, A. (1988). Features and objects: The Fourteenth Barlett Memorial Lecture. *Quarterly Journal of Experimental Psychology*, *40*, 201–237.
- Umiltà, C., Riggio, L., Dascola, I., & Rizzolatti, G. (1991). Differential effects of central and peripheral cues on the reorienting of spatial attention. *European Journal of Cognitive Psychology*, *3*, 247–267.
- Van der Heijden, A. H. C. (1992). *Selective attention in vision*. London: Routledge. London: Routledge.
- Widrow, B., & Hoff, M. E. (1960). Adaptive switching circuits. In *IRE Western Electric Show and Convention Record*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Zhang, K. (1996). Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. *Journal of Neuroscience*, *16*, 2112–2126.