

The Attentional Routing Circuit: A Neural Model of Attentional Modulation and Control of Functional Connectivity

by

Bruce Bobier

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Systems Design Engineering

Waterloo, Ontario, Canada, 2011

© Bruce Bobier 2011

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Several decades of physiology, imaging and psychophysics research on attention has generated an enormous amount of data describing myriad forms of attentional effects. A similar breadth of theoretical models have been proposed that attempt to explain these effects in varying amounts of detail. However, there remains a need for neurally detailed mechanistic models of attention that connect more directly with various kinds of experimental data – behavioural, psychophysical, neurophysiological, and neuroanatomical – and that provide experimentally testable predictions.

Research has been conducted that aims to identify neurally consistent principles that underlie selective attentional processing in cortex. The research specifically focuses on describing the functional mechanisms of attentional routing in a large-scale hierarchical model, and demonstrating the biological plausibility of the model by presenting a spiking neuron implementation that can account for a variety of attentional effects.

The thesis begins by discussing several significant physiological effects of attention, and prominent brain areas involved in selective attention, which provide strong constraints for developing a model of attentional processing in cortex. Several prominent models of attention are then discussed, from which a set of common limitations in existing models is assembled that need to be addressed by the proposed model. One central limitation is that, for many existing models, it remains to be demonstrated that their computations can be plausibly performed in spiking neurons. Further, few models address attentional effects for more than a single neuron or single cortical area. And finally, few are able to account for different forms of attentional modulation in a single detailed model. These and other limitations are addressed by the Attentional Routing Circuit (ARC) proposed in this thesis.

The presentation of the ARC begins with the proposal of a high-level mathematical model for selective routing in the visual hierarchy. The mathematical model is used to demonstrate that the suggested mechanisms allow for scale- and position-invariant representations of attended stimuli to be formed, and provides a functional context for interpreting detailed physiological effects.

To evaluate the model’s biological plausibility, the Neural Engineering Framework (NEF) is used to implement the ARC as a detailed spiking neuron model. Simulation results are then presented which demonstrate that selective routing can be performed efficiently in spiking neurons in a way that is consistent with the mathematical model. The neural circuitry for computing and applying attentional control signals in the ARC is then mapped on to neural populations in specific cortical laminae using known anatomical interlaminar and interareal connections to support the plausibility of its cortical implementation.

The model is then tested for its ability to account for several forms of attentional modulation that have been reported in neurophysiological experiments. Three experiments of attention in macaque are simulated using the ARC, and for each of these experiments, the model is shown to be quantitatively consistent with measured data. Specifically, a study by Womelsdorf et al. (2008) demonstrates that spatial shifts of attention result in a shifting and shrinking of receptive fields depending on the target's position. An experiment by Treue and Martinez-Trujillo (1999) reports that attentional shifts between receptive field stimuli produce a multiplicative scaling of responses, but do not affect the neural tuning sensitivity. Finally, a study by Lee and Maunsell (2010) demonstrates that attentional shifts result in a multiplicative scaling of neural contrast-response functions that is consistent with a response-gain effect. The model accounts for each of these experimentally observed attentional effects using a single mechanism for selectively processing attended stimuli.

In conclusion, it is suggested that the ARC is distinguished from previous models by providing a unifying interpretation of attentional effects at the level of single cells, neural populations, cortical areas, and over the bulk of the visual hierarchy. As well, there are several advantages of the ARC over previous models, including: (1) scalability to larger implementations without affecting the model's principles; (2) a significant increase in biological plausibility; (3) the ability to account for experimental results at multiple levels of analysis; (4) a detailed description of the model's anatomical substrate; (5) the ability to perform selective routing while preserving biological detail; and (6) generating a variety of experimentally testable predictions.

Acknowledgements

First and foremost, I would like to thank my advisor, Chris Eliasmith. Without his guidance and inspiration, this work would have never been possible. Chris has been central to the development of these ideas, and has been an invaluable mentor in developing my scientific curiosity. From neuroscience debates, to demonstrating how to play an F# chord on a paper drawing of a piano keyboard, and to discussing the finer points of negotiating through mogul fields, Chris has taught me that with careful consideration, many of these problems are less complex than they at first seem.

I would also like to thank Terry Stewart for his continual guidance throughout my Ph.D. I am fortunate to have had the opportunity to work with Terry, who has been pivotal in setting up, analyzing and interpreting the simulations in this thesis. Terry provides a great deal of energy to the lab and those around him, while ensuring that we are always conducting good science. My sincerest thanks to Terry, along with the other members of the CTN, particularly Trevor Bekolay, Xuan Choo, Travis DeWolf, Albert Mallia, Daniel Rasmussen, and Charlie Tang, who have been tremendously helpful, supportive and hilarious throughout this whole process.

My parents and sister have been there for me all along, and I cannot thank them enough for their constant support and guidance. My dearest friends, Andy, Sarah, Clinton, and Alice, I thank you for reminding me that there exists a world beyond my notebook, and for rescuing me from my work on more occasions than I can count.

Finally, my fiancée Alana, for her invariable loving support that has given me the perseverance to see this through.

Contents

1	Introduction to Visual Attention	1
1.1	The Problem of Visual Attention	1
1.2	Motivation and Objectives	2
1.3	Experimental Observations of Attentional Effects	5
1.3.1	Behavioural and Cognitive Effects	5
1.3.2	Physiological Effects	7
1.4	Thesis Organization	11
2	The Neurobiology of Attention	13
2.1	Cortical Areas	13
2.1.1	Ventral Stream	16
2.1.2	Dorsal Stream	18
2.1.3	Frontal Eye Fields	19
2.2	Subcortical Structures	20
2.2.1	Pulvinar	20
2.2.2	Superior Colliculus	24
2.3	The Columnar Hypothesis	25
2.4	Dendritic Computation	26
3	Models of Visual Attention	31
3.1	Selecting Attentional Targets	31
3.2	Approaches to Selective Information Routing	33

3.3	Gain Fields and Dynamic Routing	35
3.3.1	Gain Fields	35
3.3.2	Shifter Circuit	37
3.3.3	Selective Attention for Identification Model (SAIM)	40
3.4	Competition Based Routing	41
3.4.1	Biased Competition	41
3.4.2	Normalization Model of Attention	43
3.5	Synchrony- and Oscillation-based Models	47
3.6	Summary of Existing Models	49
4	Attentional Routing Circuit - High Level Overview	53
4.1	Overview of Model Structure	53
4.2	Control Calculations	56
4.3	Dendritic Nonlinearities	66
4.4	Coarse to Fine Processing	67
4.5	Default Routing	69
4.6	Object-centred Reference Frame	71
4.7	Comparison to Existing Models	74
5	Attentional Routing Circuit - Computational Specification	79
5.1	Neural Engineering Framework	79
5.2	Spiking Neuron Implementation of the Attentional Routing Circuit	83
5.2.1	Simulation Results from Spiking Neuron Implementation of the ARC	89
5.2.2	Influence of the Number of Neurons and Dendrites in each Column	97
5.3	Separating Inhibitory and Excitatory Connections	98
5.4	Neurobiological Mapping	102
5.4.1	Overview of Cortical Layers	103
5.4.2	Laminar Mapping	105

6	Attentional Routing Circuit - Simulations of Attentional Modulation Experiments	107
6.1	Simulations of Womelsdorf et al. (2008)	107
6.1.1	Changes in Receptive Field Amplitude	112
6.1.2	Changes in Receptive Field Position	114
6.1.3	Changes in Receptive Field Size	115
6.1.4	Results Summary	117
6.2	Simulations of Treue and Martinez-Trujillo (1999)	121
6.2.1	Results Summary	129
6.3	Simulations of Lee and Maunsell (2010)	132
6.3.1	Results Summary	141
6.4	Summary of Simulation Results	144
7	Future Work and Conclusions	147
7.1	Summary of Results and Predictions	147
7.1.1	Predictions	149
7.1.2	Assumptions	150
7.2	Summary of Comparison to Past Models	154
7.3	Future Work	157
	Bibliography	183

Chapter 1

Introduction to Visual Attention

This introductory chapter provides a high level overview of visuospatial attention and discusses some general psychophysical, physiological, anatomical, and computational findings that together have contributed to our understanding of the mechanisms by which attentional routing is performed in cortex. The research presented here incorporates results from studies of these areas, which serve as constraints for constructing a biologically plausible mechanistic model of how attended information is selectively processed in the brain.

1.1 The Problem of Visual Attention

William James famously wrote, “Every one knows what attention is. It is the *taking possession by the mind*, in clear and vivid form, of *one out of what seem several* simultaneously possible objects or trains of thought. [...] It implies *withdrawal from some things* in order to deal effectively with others...” [emphasis added] [95]. This oft-quoted statement efficaciously highlights that attention need not be considered as a monolithic process, but rather as being composed of at least three key subcomponents: the selection of a particular stimulus, the selective processing of that target, and the inhibition of other irrelevant stimuli. The research presented in this thesis addresses the second and third of these attentional components, that of selectively processing attended visual information and inhibiting irrelevant information in cortex so that it may gain access to further processing.

The process of selecting an attention target can be driven by exogenous or endogenous factors, or a combination thereof. Exogenous selection draws attention to salient targets in the environment based on the conspicuity or salience of their features, such as luminance onset, colour, orientation, and motion [269]. In this thesis, the term “feature” is used to refer to a particular stimulus dimension to which a neuron may respond, such as orientation, motion direction, or colour. The term “feature value” then refers to a particular measure of

a given stimulus dimension, such as a contrast value of 50% or an orientation of 90°. Such targets are thought to be salient due to their features and feature values being sufficiently different from those of other nearby stimuli, so as to stand out among their surroundings. Conversely, endogenous selection is a process in which volitional control has a greater influence, with selection being goal-driven, as in visual search tasks, and generally involves a slower and more sustained attentional state [277].

Once the attentional target has been determined, either through an exogenous or endogenous selection process, attention may then be directed to the target overtly with eye movements, or covertly, without eye movements. Covert attention can be thought of as looking at something out of the corner of one's eye, where the eyes are directed to a separate location than that to which one is paying attention.

The third component of attention is closely related to the selective processing component, and involves minimizing the amount of extraneous information that is processed, so that more neuronal resources can be devoted to processing the target. Each second, between one million and one hundred million bits of information fall on the retina and are passed through the optic nerve [102]. This considerable volume of information far exceeds the processing capacity of the visual system in subsequent striate and extra-striate cortical regions. The selective processing of relevant stimuli and concomitant suppression of irrelevant stimuli are inextricably intertwined, and together allow the brain to circumvent its resource limitations.

For both exogenous and endogenous selection, as well as for overt and covert shifts, attention may be focused on an object (object-based attention), on visual features such as colour, orientation, and motion (feature-based attention), or on a spatial location (spatial attention). The model of selective attentional processing proposed in this thesis, the Attentional Routing Circuit (ARC), addresses the problem of selectively processing attentional targets that have been determined by endogenous selection mechanisms and attended to either overtly or covertly.

1.2 Motivation and Objectives

At any one point in time, the amount of information falling on the retinae far exceeds that which can be processed by the visual system. Attention can be functionally characterized as filtering out the extraneous information so that the information relevant to the task at hand may be processed as it traverses the hierarchy of visual brain areas, without interference from irrelevant and distracting information. Given the dense and redundant connectivity between neurons in visual areas, the problem of selective attentional processing can be thought of as a problem of selecting one out of many paths or routes through which information can travel. Although the past two decades have seen numerous proposals of

how this may occur in cortex, what is lacking is a single model that performs the necessary functions of selective routing, and does so in a way that is anatomically and physiologically plausible.

The routing problem can be contextualized as being part of a larger question, namely how does the brain recognize objects in a natural environment when they can appear at different locations, under different lighting, and with different sizes? Consider the case of looking at a butterfly and attempting to identify its species based on its appearance – When viewed from a distance of 5 metres, the patterns formed on the retina are drastically different than those formed when the same butterfly is viewed from a distance of 1 metre. A similar problem arises when directing the eyes to the side of the butterfly as compared with looking directly at it. That we can recognize it as being a butterfly invariant to its spatial position and size, and possibly identify its species, is a feat that computer vision research has met with limited success, especially when compared to human performance.

The neural realization of attentional functions can be observed in myriad forms and at different levels of analysis, from increases in neurotransmitters such as acetylcholine [85] to improved reaction times in detecting changes at an attended location [183]. Not surprisingly, a breadth of approaches have been proposed by theoreticians attempting to define the conditions and mechanisms that give rise to these phenomena. Unfortunately, few of the models based on these theoretical approaches address more than one level of empirical support. Nevertheless, computational modelling, in concert with experimental work, offers a powerful method for furthering our understanding of attentional function. Models can, in principle, integrate different levels of abstraction, addressing the influence of extracellular chemicals on neuronal responses [145], the temporal characteristics of sequences of action potentials [39, 163], changes in firing rates of single cells or populations [188, 204], changes in tuning curve profiles for a given feature [188, 189, 245], modulations at the level of cortical areas [168, 218], and psychophysical effects [2, 30, 82].

In short, theoretical models can approach the problem from a functional and implementational perspective simultaneously, aiming to specify the computations that might be performed, and how they might be implemented, in order for the system to exhibit selective processing. Thus, theoretical models have the potential to bridge the gaps that exist between the many levels of empirical description currently available. Unfortunately, most models currently on offer for attentional phenomena do not do so.

The approach taken in this thesis is to study the anatomical and physiological properties of areas known to be involved in attentive processing, and to then use this information to formulate a model that is not only capable of performing attentional tasks, but does so within the constraints of biology. A central assumption of the work presented here is that scale- and position-invariant recognition ability is subserved by routing attended information to an object-centred reference frame in cortex. Here, an object-centred reference frame refers a representation of an object that is formed in higher level cortex, invariant

to its retinal size and position, and in a manner that is relatively independent of the visual information surrounding that object. It is presumed that in each level of visual cortex, local spatial relationships of attended stimuli are maintained, such that the relative spatial topography of the object formed on the retina can be extracted at each cortical level.

Overall, this thesis aims to identify general principles of how an attentional system can dynamically modulate the manner in which visual information is processed in cortex. These principles propose that the same routing mechanism exists in each area of the visual hierarchy, and throughout the extent of each area. The principles are also largely independent of changes in its parameters, such as the neuron model being used (e.g. LIF, aLIF, Hodgkin-Huxley), the number of neurons, connectivity between areas, or stimulus features being represented.

The model derived from these principles is intended to be sufficiently general that it is able to account for results gathered using numerous experimental protocols, without need for significant alterations. In designing the model, attention has been paid to minimizing the number of free parameters. In models having many free parameters, it is increasingly possible for the model to match experimental data by sufficiently adjusting parameters, although doing so limits the model's explanatory and predictive power for other data and experiments [228]. However, while preserving its generality, the model is also intended to be sufficiently specific in detail to quantitatively match experimental data. Importantly, the model itself is specified at the level of individual spiking neurons and neural connectivity. Several examples are presented later in the thesis.

There are several benefits of first modelling the problem at the computational level. First, specifying the model as a functional mathematical abstraction provides a clear description of the model's components and their interactions. This allows the mechanism to be contextualized in a large-scale implementation and to demonstrate that it is capable of performing the functions that are required for attentional routing before introducing the additional complexity that comes with a neural model. However, there are many possible ways in which the same function could also be performed, although some of them may be implausible due to the limited processing and representational capacity of neurons. For instance, selective routing could be performed by providing each neuron in each visual area with a unique signal that indicates the part of their receptive field from which they should selectively process visual signals. Such an arrangement however, suffers from a combinatorial explosion in the number of unique control signals required and the vast connectivity from control neurons to visually responsive neurons, which quickly becomes implausible.

By subsequently implementing the abstracted model in spiking neurons it is possible to assess the neural feasibility of the mechanisms. If it can be successfully demonstrated that the abstracted model produces the expected effects, and can be plausibly implemented within biological constraints, then simulations with the spiking model can be conducted to verify whether the proposed mechanisms reproduce the phenomena to be explained.

The inherent neuronal noise and variability of spiking neurons yield results that vary from trial to trial, and repeating the trials for different neurons operating on the same general principles, allows the proposed mechanisms to be tested in numerous simulated animals. Across these levels of abstraction, the model can then be qualitatively assessed based on its behaviour, and quantitatively assessed against specific experimental data. This second kind of assessment is of particular significance, since if the model is shown to be consistent with experimental data, then it is possible for predictions to be made. Such predictions may suggest ways in which details of the proposed mechanisms may be tested experimentally, or the expected distribution of effects that would be observed if recordings were made from additional neurons and animals. As well, such a model may also be used to predict other properties and characteristics of neurons that were not measured.

The following section introduces some of the forms of attentional modulation that have been observed, many of which may be consistent with the model. Presently, the majority of these observations seem consistent with the model, though not all have been explicitly simulated in this thesis (see Chapter 6 for details).

1.3 Experimental Observations of Attentional Effects

Much of our understanding of visual attention comes from combining the findings of psychophysical, neurophysiological and computational modelling research. Psychophysical studies assess the functional capabilities of the visual system typically by measuring changes in subjects' reaction times or in their ability to discriminate stimuli in an attention demanding task. Neurophysiological studies examine how a subject's attentional state produces changes in the spiking activity, neuronal tuning or activation of different brain areas. Finally, computational modelling studies seek to theorize the underlying mechanisms that may produce the behaviourally or physiologically recorded phenomena, formulate theories that integrate the characteristics of neural activity and behavioural responses, and use these findings to model these cortical and subcortical visual areas so as to make testable behavioural predictions [267]. Currently, most models of attention address the problem by attempting to reproduce high level psychophysical data, or by attempting to reproduce physiological data using abstract or mathematical models that lack sufficient biological detail.

1.3.1 Behavioural and Cognitive Effects

Around the middle of the 20th century, two theories of attention were dominant: early and late selection models. Early selection models (e.g. [24, 248]) theorize that the recognition

mechanism can only operate on one stimulus at a time, and thus unattended visual information is attenuated before its contents can be further processed. A single stimulus is selected for further processing so that under focused attention, the limited capacity channel can operate specifically on the selected information. Since the attentional selection is said to precede recognition, such models are termed early selection models.

Late selection models (e.g. Deutsch and Deutsch [47]) on the other hand, theorize that recognition has an unlimited capacity and all sensory information is processed preattentively in parallel. Attention then serves to allow a subset of the processed information to gain access to memory and consciousness. Both models suggest a bottleneck in processing resources, either as a perceptual limitation (early selection), or as a response limitation (late selection). Since these proposals were originally made, numerous theories of attention have been presented that are more biologically detailed. These are largely in agreement with the early selection theory, and are discussed in Chapter 3.

To test these ideas, two experimental paradigms have been commonly used in psychophysics studies of visuospatial attention, namely visual search and spatial cueing. In visual search tasks, subjects are instructed to search for a target stimulus placed among distractors. In a typical task, the set size (number of distractors) is varied across trials, and the subject's reaction time (RT) is measured as a function of set size. After plotting RT versus set size, the search is said to be efficient if the slope is shallow (i.e. slope $< 10\text{ms/item}$), while inefficient searches are characterized by steep slopes (i.e. slope $> 10\text{ ms/item}$) [269]. Early on, it was suggested that efficient searches were indicative of parallel, preattentive processing, while inefficient searches reflected a serial, attentive processing [54]. However, this theory is now considered less tenable, since steep RT slopes may instead reflect a parallel process with limited capacity; i.e. multiple display items may be searched in parallel, but due to capacity limitations, searching through larger sets requires more time [52].

In spatial cueing tasks [183], a stimulus cue is presented in the display prior to target onset, and may either validly or invalidly predict the target's location. The cue can take a variety of forms, such as a peripheral flicker, arrow or word, and performance is typically measured as reaction time, discrimination accuracy, localization accuracy, or identification accuracy [270]. Cues and targets are typically displayed for short intervals ($\leq 200\text{ ms}$) to prevent covert attention shifts or eye movements to the cued location [53]. Subjects typically respond more quickly on trials with valid cues, while subjects with attentional deficits or lesions can exhibit significantly different behaviour.

At the behavioural level more generally, attending to a stimulus has been shown to produce perceptual benefits, including increases in spatial acuity [276], improved contrast sensitivity [29], and improved reaction times to events occurring at an attended location [183]. While the model presented here reproduces some of these effects, it is predominately focused on attention at the neuronal and population level, so the discussion

primarily addresses attentional effects at the physiological level.

1.3.2 Physiological Effects

In cortical and subcortical areas, electrophysiology studies have demonstrated numerous ways in which attending to a spatial location or non-spatial feature can modulate neuronal responses. Measuring the modulation is typically performed by first locating the recorded neuron's receptive field and then determining the cell's preferred stimulus. The receptive field is the region of the retina within which visual stimuli can influence the firing of the cell. The preferred stimulus is the stimulus to which the cell responds most strongly. The neuron's response to variations of that stimulus is then recorded. The animal typically then covertly directs their attention to a stimulus inside the receptive field or to one outside of the receptive field. In higher cortical areas such as MT, V4 and IT, the large receptive field sizes allow multiple stimuli to be placed inside the receptive field, and attention may be directed to either of the stimuli inside the receptive field or outside.

Some studies have examined receptive field modulation by recording neural responses when attention is directed to one of two stimuli placed within the recorded cell's receptive field, or to a target outside the receptive field, and examined how responses are affected by the stimulus configuration and attentional target. Reynolds et al. [188] found that when a probe and reference stimuli were placed symmetrically inside a V2 or V4 neuron's receptive field and attention was covertly directed to a stimulus outside the receptive field, the cell's activity was driven by both stimuli, yielding a response somewhere between that elicited by each stimuli appearing alone. When attention was directed to one of the stimuli within the receptive field however, the cell's response was similar to that elicited by the stimulus being presented alone, with the unattended stimulus having a lesser impact on its firing rate, suggesting that the neuron is selectively processing the attended stimulus. Recording from V4 and IT, Moran and Desimone [147] used a match-to-sample task in which they had monkeys attend to coloured bars at one location while ignoring a stimulus at another. When both stimuli were within the recorded cell's receptive field, the neuron responded strongly when attention was directed to the preferred stimulus, but its response was significantly reduced when the preferred stimulus was positioned at the unattended position. With two stimuli in the receptive field, responses were primarily determined by the feature values of the attended stimulus, being similar to that elicited by the attended stimulus being presented alone. This was taken to suggest that attention involves a suppression of unattended information, rather than an enhancement of attended information as is suggested by some models [188, 189]. Other studies have reported consistent effects using similar approaches in ventral stream areas V2 and V4 [118], as well as dorsal stream areas MT and MST [186, 246, 247]

Attention also modulates the spatial properties of receptive fields. As shown by Con-

nor et al. [36] and Womelsdorf et al. [272], when attention is directed toward a stimulus, the effective portion of a neuron’s receptive field may shift toward the attended stimulus. Receptive field shifts enable the neuron to selectively process information near the attentional target, with the non-attended object contributing less to its activity, and thereby allow attended information to be routed to subsequent areas for further processing. These effects have been reported in both ventral and dorsal areas, where the Connor et al. [36] study examined neurons in V4, while Womelsdorf et al. [272] examined neurons in MT. The results from both studies closely match each other, suggesting that attention may work in similar ways in both ventral and dorsal stream areas.

Neurons in more peripheral areas such as LGN and V1 have smaller receptive fields, thereby making it difficult to place multiple stimuli inside their RF, and thus more difficult to quantify changes in the size and position of the neuron’s RF [118]. However, several studies have observed changes in firing rates of LGN [132] and V1 [149, 196] neurons when an animal covertly attends to a stimulus inside the cell’s RF versus away from it, despite the same stimulus falling in its RF. Other studies have reported only weak modulation in V1 [131] or no modulation [118, 147], although the experimental methods used in these studies vary considerably. Consequently, the observation of attentional modulation, or lack thereof, may be related to the difficulty of the task being performed, where fine detailed information encoded in lower areas may or may not be required for a particular task.

The observation of weaker attentional modulation in V1 than in higher areas [26] has led to the idea that attentional modulation in V1 may be driven by attentional feedback signals from higher extrastriate areas. This hypothesis is supported by studies of attentional effects in ventral stream areas which have found that the timing at which attentional modulation occurs, as well as the strength and prevalence of the modulation varies considerably between areas. Recording laminar event-related potentials (ERPs) in V1, V2, V4, IT, and superior temporal sulcus, Mehta and colleagues [138] found that attentional modulation of higher cortical areas precedes that of earlier areas. Additionally, they found more powerful and earlier effects occurring in V4 100-300 ms after stimulus onset, with smaller effects occurring in V2 during the same period, and later in V1. Similar results have been reported in a study by Buffalo and colleagues [26] that measured single cell activity in V1, V2 and V4, which found that attention modulates the responses of neurons in higher areas earlier than in lower areas. They also report that the proportion of recorded cells showing modulation is greater in higher areas, and that there is a significantly stronger modulation of activity in higher areas than in lower areas. In other electrophysiology studies, attentional modulation has been observed to occur in TE with a latency of ~ 200 ms [32], while attentional modulation in V1 has a longer latency of ~ 235 ms [196] or more [26].

Consistent with these reports of modulation in ventral stream areas, Herrington and Assad [86] found that attentional effects appear significantly earlier in dorsal areas LIP than in MT, with which it is connected. Their study had monkeys covertly shift attention

between two peripheral stimuli that were validly cued 85% of the time, and detect a short increase at either dot patch. It was found that the onset of attentional modulation was ~ 60 ms earlier in LIP than in MT, with approximately 85% of LIP neurons showing an increase in the attentional index when attending to the receptive field stimulus compared with roughly 76% in MT. Further, the strength of modulation was greater in LIP than MT by a factor of ~ 3 .

With a single stimulus presented inside a neuron’s receptive field, the main attentional effect of directing attention to the spatial position of that stimulus is to increase the neuron’s response to all stimulus values, both preferred and non-preferred [43]. Such effects have been reported for orientation tuning in V4 [131] and motion direction in MT and MST [245], where the responses proportionally increase for all orientations and directions respectively. This form of modulation is referred to as “response gain” because it affects the overall gain or firing rate of the neuron across different feature values (Figure 1.1(a)).

Other studies have investigated the influence of stimulus contrast on attentional modulation, finding that the amount of gain is greater when attention is directed to low contrast stimuli than to high contrast stimuli [125, 190]. This form of modulation is referred to as “contrast gain” because it provides a similar effect as proportionally increasing the contrast of the attended stimulus (Figure 1.1(b)).

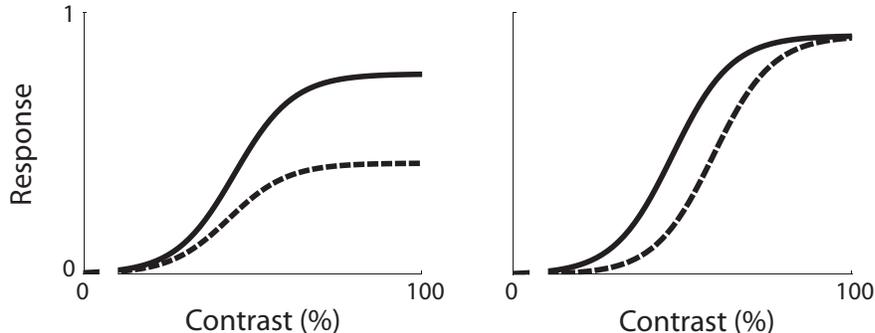


Figure 1.1: Effect of stimulus contrast on neuronal response. The x axis indicates contrast (%) on a logarithmic scale and the y axis indicates neural response normalized to the range [0, 1]. (a) Response gain, where the neuron’s activity is modulated by a similar proportion for all contrasts. (b) Contrast gain, where responses to low and intermediate contrast stimuli are scaled by a larger amount than for high contrast stimuli, resulting in a leftward shift of the contrast-response curve. (Adapted from [112]).

Under which conditions neurons exhibit a pure response gain or contrast gain remains debated in the literature for several reasons. First, detecting changes in low contrast stimuli requires a greater attentional vigilance than for high contrast stimuli and may artificially inflate the effect for low contrast stimuli, resulting in stronger modulation [112].

As has been shown in numerous studies, attentional effects are generally stronger with more complex tasks requiring fine discrimination [112, 118, 147, 222]. Second, long stimulus presentations may provide sufficient time for adaptation to low contrast stimuli, which may also obscure this effect. For example, Martinez-Trujillo and Treue [125] presented each stimulus for 1000 ms and conducted the analysis using the average activity from 200-1000 ms following stimulus onset.

Third, the study by Martinez-Trujillo and Treue [125] reports a contrast gain effect, although response gain was also found to provide a good fit to the data, as the correlation coefficients for both models were greater than 0.82 across neurons. An early study by Reynolds et al. [190] also reported contrast gain between attend-out and attend-in conditions with a single receptive field stimulus, although that study did not quantitatively test whether the data were better fit by a response gain or contrast gain model. As noted by Lee and Maunsell [112], only a small number of cells in the Reynolds et al. study could not be explained by response gain; rather, the data reported by Reynolds et al. present suggest that, were a quantitative analysis performed, the response gain model would also provide a suitable description of the modulation effect. In addition to the stimulus and task design, the discrepancies between these studies may arise from different methods being used for analysis. Reynolds et al. [190], as well as a study of attentional effects in superior colliculus by Li and Basso [115], both suggest that the effects were consistent with contrast gain. However, both studies examined neural responses using a receiver operator characteristic (ROC) analysis, which compresses differences at higher contrast values, where the effects of response gain are most evident [266]. Using a similar experimental method to that of Reynolds et al. [190], Williford and Maunsell [266] found that most of the attentional effects could be marginally better explained by response gain or activity gain than by contrast gain. In sum, it remains unclear the conditions in which a pure contrast gain may be evoked, and this issue is returned to in Chapter 6 where simulations with the presented model show some evidence for contrast gain, although response gain is demonstrated to significantly better explain the effect.

In addition to modulating the firing rates and tuning of cortical neurons, several studies have reported attentional modulation of the temporal structure of neuronal activity. Cook and Maunsell [37] examined the effects of spatial attention on motion integration for MT neurons, showing that when attention was covertly directed to a receptive field stimulus, there was a multiplicative scaling of the neurons' temporal integration window. In V4, Mitchell et al. [144] found a modest but significant reduction in the variance of firing rates as measured by the Fano factor (ratio of spike count variance to mean spike count), and a stronger and more reliable attentional modulation of putative interneurons. These interneurons, as identified by the narrow width of their action potentials, predominately project locally and not to other cortical areas. In another study of V4 neurons, Fries et al. [61] found that neurons activated by an attended stimulus showed increased gamma-

band (35-90 Hz) synchrony as compared with nearby neurons activated by distractors.

Overall, the primary effects considered by the model are responses being primarily driven by the attended stimulus, changes in receptive field profiles, proportional scaling of responses to attended stimuli, and response gain versus contrast gain, while the effect of earlier and stronger modulation in higher cortical areas is consistent with the model. Many of these forms of attentional modulation produce comparable effects in both ventral stream areas such as V4 and dorsal stream areas such as MT. Given the similarity of these effects across many studies, it is proposed that the brain has evolved to employ similar mechanisms in dorsal stream areas as well.

1.4 Thesis Organization

Chapter 2 reviews the literature pertaining to the neurobiology of attention and outlines significant findings that serve to constrain biologically plausible models. This chapter also serves to describe the relevant biological details of the brain areas involved in attentional processing, so as to provide a frame of reference for the discussion of existing models of attentional processing.

Chapter 3 begins with a discussion of possible mechanisms for attentional routing and reviews several related models. The reviewed models include gain field models [204, 272], shifter circuit [168], SAIM [82], biased competition [188], normalization model [189], and synchrony- and oscillation-based models [162, 163]. Across these models, there are several common limitations that remain to be addressed in a single model. These limitations include a spiking neuron implementation, defining the model for attentional processing across multiple cortical areas, and being able to account for multiple forms of attentional modulation using a single detailed mechanism.

Chapter 4 introduces the Attentional Routing Circuit at a high level to provide a general sense of the proposed mechanisms for selective routing by defining the computations that are performed in the model. Examples of selective routing using the model are presented, and several of the model's assumptions are discussed. Finally, the presented model is then compared with previous models introduced in Chapter 3, and their similarities and limitations are discussed.

Chapter 5 then introduces the Neural Engineering Framework (NEF) [55] which allows the mathematical model to be implemented in spiking neurons. Subsequently, the general model is defined in greater neuronal detail in order to elucidate the division of labour amongst groups of neurons in different cortical areas. Results of several simulations with the spiking neuron model are then presented and are shown to be consistent with that predicted by the mathematical model. With the functional roles of the neuronal ensembles

defined, further details of the model's anatomical substrate are given as a canonical laminar cortical circuit to which the model is mapped.

With the functional principles of the model defined, as well as their anatomical and physiological basis, Chapter 6 presents results of simulations using the ARC that allow the model to be more closely tied to experimental findings while preserving its overall functionality.

Chapter 7 then summarizes the simulation results and model predictions, and discusses several directions for future work.

Chapter 2

The Neurobiology of Attention

This chapter outlines the anatomy and physiology of several cortical and subcortical areas, and discusses their relationship and roles in the visual processing of attended information. This review focuses on the levels of the properties of single cells, the laminar and columnar structure within individual cortical areas, as well as inter-laminar and inter-areal communication. The goal of this chapter is to provide a sense of how single cells process information and how groups of neurons may interact as part of a cortical circuit so as to yield a frame of reference for constructing a detailed model of attentional routing.

2.1 Cortical Areas

At least 30 cortical areas, representing more than half of cerebral cortex [67], are known to be involved in visual processing [59]. Based on their connectivity and neuronal properties, they have been organized into a hierarchy with 10 levels. Within the visual system, there are two major processing streams that originate in V1 and include several extrastriate regions. The ventral or “what” pathway, is centrally involved in object recognition, with information projecting from V1 to V2, V4 and on to inferotemporal cortex, which is a complex of approximately six visual areas in the inferior portion of the temporal lobe [59]. The dorsal or “where” pathway, is involved with determining spatial relationships among objects and visual guidance toward them, and projects through V2, V3, MT, VIP, and on to LIP and area 7a. In general, neurons in the ventral stream respond to features that are relevant for object recognition, such as colour and shape, while neurons in the dorsal stream respond to more spatial features, such as direction and velocity of stimulus movement. Figure 2.1 illustrates the connectivity between visual cortical areas, with connections in the dorsal and ventral streams represented by dashed and solid lines, respectively. Although these processing streams are often thought of as being independent, this is not strictly the

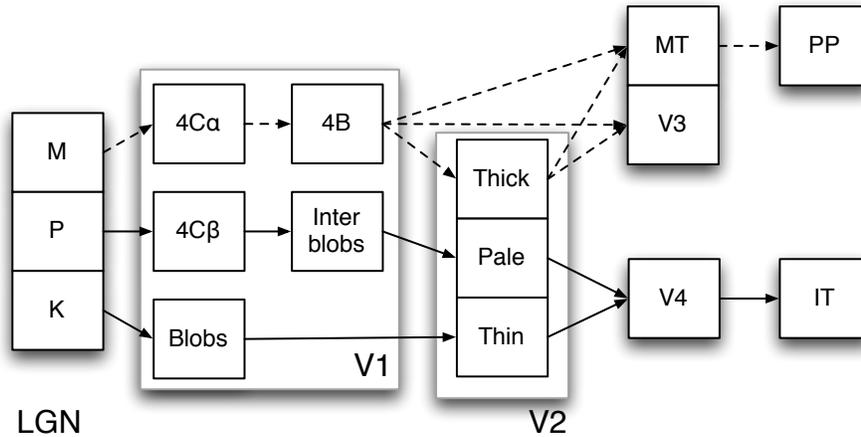


Figure 2.1: Connections between visual cortical areas. Dashed and solid lines indicate connections within the dorsal and ventral stream, respectively. LGN – lateral geniculate nucleus, magnocellular (M), parvocellular (P) and koniocellular (K) divisions. V1 – primary visual cortex, layers $4C\alpha$, $4C\beta$ and 4B, and blob and interblob regions. V2 – second visual area, containing thick, pale and thin stripes. V3 – third visual area. MT – Middle temporal area. PP – Posterior parietal cortex. IT – inferotemporal cortex. See text for details.

case, as there are numerous points in each stream at which there are reciprocal connections between dorsal and ventral stream areas, particularly in higher levels [12, 59].

In addition to feedforward connections from earlier processing areas, each area also has reciprocal feedback connections. A general rule of connectivity between two cortical areas is that if area A projects to area B, then area B reciprocates this connection [59, 109]. Although the exact function of the feedback connections is not fully known, it is tacitly assumed that feedforward connections serve to drive the target cells, while feedback connections serve to modulate the target cells' activity (e.g. for selective attention) [9, 79, 104].

In most topographically organized visual areas, the number of neurons representing an area of visual field decreases sharply and monotonically between foveal and peripheral representations. Quantitatively, this is expressed by the cortical magnification factor, which describes the area of cortex (mm) representing one degree of visual angle. This resembles the change in retinal ganglion cell density with eccentricity, and it has been proposed that the increased number of neurons representing the central field in V1 (approximately 55%-60% of the surface area corresponds to the central 10°) arises from the spatial density of retinal projections [264].

At progressively higher levels of the hierarchy, the surface area of each region, as well

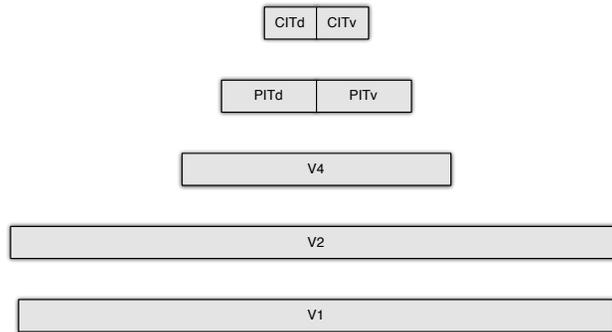


Figure 2.2: Proportional surface area of cortical regions in the ventral stream. At higher levels of the visual hierarchy, the surface area of each region, and the number of neurons they contain, decreases.

as the number of neurons they contain, typically decreases (Figure 2.2), while the size of the receptive fields increases (Figure 2.3). This latter property is thought to arise due to neurons in higher levels pooling visual information from multiple neurons within their receptive field. A hierarchical organization also allows lower level neurons to respond to relatively local visual features, while higher level neurons are activated to more global, complex visual features and respond to combinations of features encoded by their afferents. The combination of larger, overlapping receptive fields and increasing abstraction of visual features at higher levels results in the clear topographic organization found in V1 and V2 becomes being less obvious in higher levels, such as V4, PIT and CIT.

The hierarchical organization of the more than three dozen visual cortical areas proposed by Felleman and Van Essen [59] is based on the physical and connective properties of these areas. However, a different hierarchical organization can be formed by considering the temporal pattern of activation following stimulus onset. For example, Felleman and Van Essen [59] place frontal eye fields (FEF), an area involved in eye movements and attention, at the same level as central IT in the anatomical hierarchy, yet cells in FEF show activation with a similar latency to cells in V1, despite being situated seven levels higher in the anatomical hierarchy [27].

Visual area 1 (V1), or striate cortex, is among the most extensively studied visual areas, and primarily receives retinotopic projections from LGN, with central regions of the visual field being represented by a disproportionately large number of cells. V1 has a columnar organization, with interblob areas containing orientation columns made up of several simple cells having similar receptive fields and orientation tuning, as well as complex cells that receive intracolumnar projections from simple cells. Projections to interblob neurons can be traced to the LGN parvocellular layers, and V1-4C β [199, pp. 60]. Neurons in the

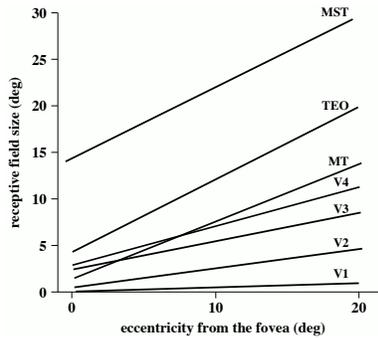


Figure 2.3: Changes in receptive field size with eccentricity. (Reproduced from [176, pp. 151]).

magnocellular layers of LGN project to V1-4C α and on to V1-4B to form the basis of the dorsal stream, sending afferents to the thick stripes of V2, as well as V3 and MT. By combining information from multiple simple cells, complex cells are able to provide more abstract representations of visual stimuli, and respond similarly to stimuli falling anywhere within their receptive field. Each orientation column is flanked by columns containing neurons tuned to slightly shifted axes of orientation, with a complete cycle of orientations being found approximately every 0.75mm [97, pp. 537]. Dispersed between these columns are blobs, which are regions of cells selective to colour rather than orientation, and receiving projections from LGN K cells. V1 also contains binocular neurons which receive input from both eyes, but typically respond more strongly to stimulation from one or the other. Similar to orientation-selective neurons, cells that preferentially respond to the same eye are grouped into ocular dominance columns, with the columns organized in stripes of alternating left and right eye dominance columns [89]. Grouping together sets of orientation columns spanning a complete cycle of orientations, blobs, and ocular dominance columns with a common receptive field, yields a hypercolumn [92].

2.1.1 Ventral Stream

In the ventral stream, information from V1 projects to alternating thin and pale regions in V2, which primarily encode colour and form information respectively. Information from the thin and pale stripes in V2 is projected through V4 to TEO (near the *temporal-occipital* border in macaque, homologous to PIT in humans), from where information about object colour, form and texture is projected to TE, the macaque homologue of CIT or AIT in humans (Figure 2.1).

In contrast with V1 neurons, most V2 cells are driven by binocular inputs, have larger

receptive fields and prefer stimuli having lower spatial frequencies [113]. Staining V2 with cytochrome oxidase shows intermixed thick, thin, and pale stripes that receive different inputs (Figure 2.1). These three types of stripes contribute to the representation of three distinct visual maps in V2, with the thin, thick and pale stripes containing neurons selective for colour, motion and form, respectively [199, pp 59]. Neurons in thin stripes of V2 receive input from blobs in V1 and project to V4, are typically not orientation selective, with over half being double-opponent colour cells. Neurons in thick stripes receive projections from the magnocellular layers of LGN via V1-4C α and V1-4B, and project to MT and V3. Thick stripe neurons are often characterized by their binocular sensitivity, as they respond most strongly when both eyes are simultaneously stimulated. They may also be orientation- and contrast-sensitive, colour-insensitive, and of the approximately 15% of V2 neurons selective for motion direction, the majority are localized here. Finally, pale neurons receive input from V1 interblobs and project to V4, and are selective to orientation but not to colour or motion, with over half being end-stopped (i.e. responsive to line ends and corners) [195, 199].

V4 receives direct projections from V1, V2 and V3, and sends strong projections to PIT. The majority of cells in V4 are selective to orientation, the length and width of oriented bars, or wavelength [44, 207], while some are responsive to contours such as angles and curves (particularly convex contours) [177], and binocular disparity [265]. Several studies have shown that in conditions where a V4 neuron’s receptive field contains two stimuli, its response is primarily driven by the attended stimuli [147, 188]; this effect is largely suppressive of unattended stimuli, rather than facilitatory, since the cell’s response is not significantly different when attending the preferred stimulus than in conditions where attention is directed to a stimulus outside the cell’s receptive field [256, pp. 546].

Inferotemporal (IT) cortex can be divided into at least six areas, arranged both hierarchically and anatomically and having ventral and dorsal subareas: posterior IT (PITv and PITd), central IT (CITv and CITd), and anterior IT (AITv and AITd). Note that the ventral and dorsal distinctions are based on their anatomical location rather than association with the ventral and dorsal processing streams. Some anatomical studies also make this ventral/dorsal distinction for lower areas including V2, V3, V4, V5, and MST. As a general principle, the dorsal division contains neurons having receptive fields in the lower visual field, while neurons in the ventral portion of a given area have receptive fields in the upper quadrants. However, this is less apparent in higher areas, where the large receptive fields can cross the horizontal and vertical meridians [77, 78].

PIT contains neurons that are selective to some basic features, as well as neurons that are selective for even more complex feature combinations than V4 neurons (e.g. complex shapes, and shape/texture and shape/colour combinations). Single cell recordings in macaque IT have shown that neurons recorded in a single vertical penetration respond to similar stimuli, and that with subsequent penetrations made nearby, cells respond pref-

entially to similar stimuli [62]. On average, each column spans ~ 0.4 mm, and although neurons within a column respond to similar stimuli, individual cells differ slightly in their preferred stimulus and tuning. The functional implications of such a configuration are that a given column may encode information about object features of multiple shapes, although the activity of different cells within that column allow the differences between those shapes to be encoded [62].

Neurons in CIT have larger receptive fields than those in PIT, although both areas also have a columnar organization, with similarly tuned neurons being grouped in the same, or nearby column [101]. Tsao et al. [251] found that almost 97% of the cells recorded in the anterior portion of macaque area TEO were selective for faces, while fMRI studies have reported greater activity in a region extending from V4 to rostral TE for pictures of intact objects than with scrambled objects [250].

2.1.2 Dorsal Stream

The dorsal stream is the complement to the ventral stream, and is a chain of visual cortical areas that contribute to representing object location, motor control (e.g. arm movements, saccades), and motion processing. The dorsal stream also begins in V1, where direction-selective neurons in layer 4B project to neurons in the thick stripes of V2, and on to V3 (Figure 2.1). Signals from these three areas then project to MT (V5) and other motion direction-selective areas in the superior temporal sulcus, and on to ventral intraparietal cortex (VIP), from where they are projected to other parietal areas including lateral intraparietal cortex (LIP) and area 7a. Several of these parietal areas also receive direct projections from peripheral representations in V1 and V2, which may serve to rapidly activate regions that mediate spatial attention [256]. Two primary functional characteristics of areas in this pathway are a larger representation of the peripheral visual field as compared to ventral stream areas, and specialization for detecting and processing moving stimuli [225, pp. 1229]. Each of these areas has topographic organization and is reciprocally connected with the pulvinar, with higher areas (e.g. MST, VIP and LIP) also having reciprocal connections with higher ventral areas such as V4, PIT, and CIT. As with areas in the ventral stream, the receptive field size of neurons increases at higher layers and at greater eccentricity (Figure 2.3).

Area V3 is located directly anterior to V2, and its representation of the central 40° of visual field is divided into two areas, with the dorsal (V3d) and ventral (V3v) portions representing the lower and upper areas of the visual field, respectively [121, 259]. Layer 4B of V1 provides the primary inputs to both portions of V3, which in turn project to dorsal areas MT, MST, and VIP [120]. V3 also receives inputs from V2 and is reciprocally connected with V4. V3d contains more directionally selective and fewer colour-sensitive

neurons than does V3v, although neurons in both areas are more sensitive to contrast, and prefer lower spatial and higher temporal frequencies than V2 neurons [68].

Visual area MT (middle temporal), or V5, receives strong afferents from V1-4B and from the thick stripes of V2, and is reciprocally connected with other topographically organized areas including V1, V2, V3, V4, VIP, LIP, inferior and lateral pulvinar nuclei, and superior colliculus. Neurons in MT largely project to MST which, as with MT, contains many directionally-selective neurons, but preferring more complex motion patterns such as clockwise or counterclockwise motion, expansion and contraction [76, 129, 199]. MT has a retinotopic organization, and most cells in MT are selective to the direction and velocity of stimuli, a characteristic shared with V1-4B and the thick stripes of V2 [66]. MT neurons have large receptive fields that are similar in size to those of V4 neurons at a given eccentricity, with the sizes increasing with eccentricity (see Figure 2.3). MT has a columnar organization with similar motion-directionally tuned neurons being grouped together. The majority of neurons in MT respond most strongly when both eyes are stimulated [127], and have directional selectivity for motion of random dots, bars and gratings, and depth-selectivity [128]. Functionally, MT appears to contribute to stereoscopic depth perception, the analysis of movement and motion integration [127].

Neurons in the lateral intraparietal area (LIP) receive projections from numerous cortical areas including V2, V3, V3a, V4, MT, MST, TEO, and TE, and have reciprocal connections with VIP and area 7a, prefrontal cortex, superior colliculus, and pulvinar [15, 18, 59]. Neurons in LIP have larger receptive fields than MT cells, and over half of the neurons respond to information from the central 6° of visual field [176]. LIP has been shown to be involved in encoding the location of an upcoming saccade [136], storing information in both retinal- and body-centric coordinates, and has neuronal activity that is modulated by the task relevance and salience of the stimuli [73, 230]. A study by Duhamel et al. [51] found that in LIP, locations of remembered objects are updated to new retinal coordinates before saccades in anticipation of upcoming eye movements. Similar effects have also been observed in V2, V3, and V3a [156].

2.1.3 Frontal Eye Fields

The frontal eye fields (FEF) play an important role in generating saccadic eye movements and attentional shifts [239] and show rapid activation following stimulus onset ($75 \pm 13\text{ms}$ [209]), well before most V2 cells. FEF neurons have large receptive fields, respond more strongly when their receptive field contains the target of an overt or covert attentional shift, and often discharge immediately prior to a shift. FEF makes strong projections to superior colliculus (SC), which in turn projects to pulvinar [179]. The rapid activation of FEF may occur via a pathway from retinal ganglion cells to superior colliculus, mediodorsal thalamus, and to FEF [100], or alternatively from retinal ganglion cells to the

magnocellular layers of LGN, V1, MT, and FEF. The magnocellular layers of LGN contain neurons selective for low spatial frequency/high temporal frequency visual information, making them suitable for a rapid first pass of the scene to determine potential attentional targets. Indeed, reversible deactivations of LGN-M with muscimol injection have been shown to cause significant impairments in an animal's ability to shift attention [191].

2.2 Subcortical Structures

2.2.1 Pulvinar

In addition to the discussed cortical areas, there are also several subcortical structures involved in visual attentive processing. The pulvinar is the largest nucleus of the thalamus and phylogenetically, it is the most recent thalamic nucleus, being relatively small in carnivores, progressively increasing in size in higher order primates, and becoming significantly larger in humans, where it is several times greater in size than the LGN [31, 162].

Pulvinar is typically subdivided into the inferior (PI), lateral (PL), medial (PM), and dorsomedial (Pdm) nuclei [227]. Dorsal pulvinar, including PM and Pdm, is connected with dorsal stream areas and in conjunction with parietal cortex, is thought to be involved in generating attentional signals. In contrast, ventral pulvinar (VP), including PI and PL, is connected with the ventral stream and is believed to play a more modulatory role in cortical processing (Shipp, personal communication).

Mapping studies of the macaque pulvinar have shown two retinotopically organized maps, one in inferior pulvinar (P1), and another in lateral pulvinar (P2) [180], both of which receive strong afferents from superior colliculus [1]. Whereas the P1 map receives projections covering the majority of the visual field, the P2 map contains only a representation of the central area [255]. The inferior and ventral portion of the lateral nuclei of the pulvinar also receive cortical projections from striate cortex and the adjoining extrastriate areas, as well as some retinal input, both directly and through superior colliculus (SC), with PI receiving more input from SC than ventral PL [227, 255]. Neurons in PI and ventral PL have well-defined receptive fields with sizes that increase with eccentricity, are retinotopically organized, respond to visual stimuli presented through either eye, and are either non-selective or broadly tuned to motion direction or orientation [179]. The characteristics and connectivity of neurons in these areas suggest that PI and ventral PL may be involved with guiding eye movements and attention [179, 192]. Within ventral pulvinar, at least two types of neurons may be distinguished. Parvalbumin-expressing neurons are relay cells that primarily drive neurons in layer-III and IV of cortex, whereas calbindin-expressing neurons are more regulatory in nature, and terminate in layer-I of cortex [71, 96], where they modulate the activity of their target neurons (Shipp, personal communication).

The medial nuclei (PM) and dorsal portion of PL have more widespread projections than PI and ventral PL, and have connections with several higher cortical areas including parietal areas 7a and LIP, frontal cortex, inferotemporal cortex, and V4, but not with V1 or V2 [214, 255]. PM is also interconnected with several other high level cortical areas including auditory and somatosensory cortex, and contains subregions that are predominately selective for information from visual or other sensory modalities. Neurons in PM and dorsal PL are directionally sensitive and have more widespread cortical projections than PI, but have a much coarser topographic organization. Based on its connectivity with higher level cortex, and the results of imaging studies, PM is also thought to be involved in directed attention and determining visual salience [110, 201].

PI and PL are more retinotopically organized than is Pdm [179], and both PI and PL are reciprocally connected with areas of the ventral stream, whereas Pdm is only known to be connected with more dorsal visual areas [227]. Many cells in Pdm are visually responsive and have saccade- and attention-modulated activity (i.e. higher activity when the receptive field contains the target), whereas neurons in PI and PL exhibit less attentional modulation [179, 180]. Although in these studies ventral pulvinar neurons showed weaker modulation than those in dorsal pulvinar, the study used a simple attentional task with only a single target stimulus without distractors. Studies of attentional modulation in V4 and MT have shown stronger effects with more difficult tasks and when two stimuli are in the receptive field than with one [21, 112, 118, 147, 222], and it is likely that neurons in ventral pulvinar will show similar modulation as dorsal pulvinar neurons if more complex tasks are used.

Recording from neurons in the dorsomedial region of pulvinar (Pdm, located in PL) of rhesus monkeys, Petersen, Robinson and Morris [180] found that Pdm neurons respond more strongly shortly before the animal saccades to a stimulus, or when it covertly attends to a peripheral stimulus. These characteristics are also found in parietal area 7a, which is thought to be involved in spatial attention, and is reciprocally connected with Pdm.

In the ARC and several other models of attention [2, 7, 162, 169], attentional signals for information routing are hypothesized to come from the pulvinar, and numerous studies have provided evidence to support this hypothesis. Posner and Petersen [184] presented an influential hypothesis, which posits that pulvinar is responsible for engaging attention, posterior parietal cortex for disengaging attention, and the superior colliculus for shifting attention. This hypothesis is consistent with several lesion and imaging studies showing that pulvinar lesions diminish an animal's ability to filter out irrelevant information and determine what is visually important or salient. Desimone et al. [45] found that by using muscimol to temporarily deactivate the portion of PL that projects to V4 and IT, macaque monkeys had reduced accuracy in colour and form discrimination tasks when distractors were present, but exhibited no such impairment in trials without distracting stimuli. Because the impairment was only observed when multiple stimuli were present, these findings

support the hypothesis of pulvinar being involved with visual selection, either by filtering out irrelevant information [45] or by enhancing relevant information [180].

Rafal and Posner [185] investigated the effects of hemorrhages in the thalamus (which included, but were not exclusive to the pulvinar) in human subjects in a spatial cueing task. Subjects had slower reaction times when detecting stimuli in the contralesional field whether or not attention had been cued to that location. However, with longer intervals between the cue and stimulus (stimulus onset asynchrony), validly cued targets in both hemifields produced similar improvements in reaction times, suggesting that they did not have a deficit in shifting attention. They hypothesized that the pulvinar lesion caused a deficit in their ability to engage attention to the target contralateral to the lesion, by impairing their ability to detect and respond to the target

Recent fMRI studies have shown increased pulvinar activity during luminance detection tasks [98] and rapid serial visual presentation tasks [275], while positron emission tomography (PET) studies have observed similar effects during object recognition tasks [110] and feature discrimination tasks [38] that require selective attention. Morris, Friston and Dolan [148] also used PET imaging to examine changes in pulvinar activity as a function of image salience. During the conditioning phase, subjects were presented with images of faces with happy, fearful or neutral expressions paired with either silence or loud white noise (making the stimuli have a more aversive and thus behaviourally significant and salient quality). During the PET scan, subjects again attentively viewed the face images but without noise. They found that for aversely conditioned faces, subjects showed increased pulvinar activity, which they hypothesize reflects the subject having increased attention to more salient stimuli, which supports pulvinar involvement in controlling processing in sensory cortex [148].

Bender and Youakim [16] recorded from cells in macaque LGN, PI, PL, Pdm, V2, V4 and parietal area 7a while the animal attended to a central fixation point. While they found no evidence of attentional modulation during fixation in LGN, approximately 26% of recorded pulvinar cells showed modulated activity, with no significant difference between subdivisions, as well as increasing amounts in areas between V2 (21%) and 7a (43%). In each of these areas, a roughly equal number of neurons were facilitated or suppressed, except in area Pdm, where all cells were suppressed during fixation. The observation of equal amounts of facilitation and suppression in PI and PL is particularly interesting, as it suggests a different interpretation of the results of Vandenberghe et al. [260]. In the study by Vandenberghe et al., no significant pulvinar activation was observed when subjects employed minimal selective attention to one of two stimuli, although they did observe activation of the left pulvinar when only a single stimulus was displayed. If pulvinar activity is facilitated or suppressed in approximately equal proportions, measurements of the regional glucose uptake in PET scans would lack the resolution to detect changes in neuronal activity as observed in single cell recordings.

Based on their findings from selective attention studies on pulvinar lesion patients, Ward and Danzinger [262] oppose the theory that the pulvinar is critical for filtering information and allocating attention. Instead, they propose that the pulvinar is involved in activating responses, particularly for situations of response competition between multiple stimuli, and in spatially coding visual features. In [263], the authors report of patient TN who has lesions in the anterior and most dorsal portions of the pulvinar, an area that corresponds to the PI and PL areas in macaque. TN has a selective impairment in localizing and binding features in the inferior contralesional quadrant, which they suggest to be a deficit of spatial coding. In that study, the subject performed a demanding central primary task while a target was briefly presented in the contralesional or ipsilesional field. When the target and distractor were in the contralesional field and separated by 1.7° , TN produced significantly more illusionary conjunction errors, although when the separation was increased to 5.1° , these errors significantly decreased.

In a different task, TN was asked to fixate a central digit, and report the central digit identity, as well as the target identity, colour and the quadrant of visual field in which it appeared. The subject again made significantly more mislocation errors for targets in the inferior contralesional quadrant, with no significant difference between the other quadrants. In a finding consistent with Desimone et al. [45], they also found that TN's ability to localize isolated, single objects was little affected in each quadrant. The authors hypothesize that damage to PI and PL causes information from the visual area affected by the lesion to be combined with erroneous or degraded spatial information.

The above studies provide support for pulvinar being involved with visual attention, with imaging and single cell recording studies showing attentional modulation of pulvinar activity, and studies of lesioned and deactivated pulvinar revealing difficulties in localizing and maintaining spatial feature relationships of contralateral stimuli. While Ward and Danzinger [262] argue that pulvinar's role is not to filter out irrelevant information and allocate attention, but rather is involved in preventing response activation and selection of objects, these hypotheses are not necessarily mutually exclusive, but rather may be addressing different aspects of the same function.

In the absence of a pulvinar signal (i.e. when it is deactivated or lesioned), the ARC model includes the proposal that the visual system operate in its default state, which causes the entire visual field to be processed at each level. Because of the increased spacing between afferents for columns in higher levels, in the default state, columns will integrate information from a greater portion of their receptive field. Placing the two stimuli more closely together may result in higher level columns integrating features from both nearby stimuli, resulting in illusionary conjunctions. This hypothesis may explain the observation made by Ward et al. [263], where fewer illusionary conjunction errors were made when the separation of stimuli was increased to 5.1° , and it is suspected that this error rate will decrease with increased separation.

2.2.2 Superior Colliculus

The superior colliculus (SC), or optic tectum, is situated at the roof of the midbrain. Histological staining of the SC shows a laminar organization that can be divided into superficial, intermediate, and deep layers [130]. In the ARC, the hypothesis of collicular function follows that of Posner and Petersen [184], who suggest that it contributes to encoding the location to which attention should be shifted, either by the deep layers as overt shifts involving eye movements, or by the superficial layers for covert shifts.

Cells in the superficial layers are topographically organized, respond more strongly to stimuli that are saccadic targets than to non-targets, more strongly to moving stimuli than non-moving stimuli, and are insensitive to the intensity, size and shape of stimuli [40]. A recent study by Ress et al. [187] had human subjects perform a difficult contrast decrement discrimination task for a slow moving stimulus while maintaining fixation. fMRI imaging of the superficial layers revealed distinct retinotopic maps for direct visual stimulation as well as covert attention that were in register. These findings are consistent with a study in rhesus monkeys by Müller et al. [154] in which the animal maintained central fixation, above which were two targets to the left or right of fixation. A peripheral cue, to which covert attention is thought to be facilitated through microstimulation of the superficial layers of SC, was given at eccentricities between 2° and 35° as a cluster of coherently moving dots whose direction (left or right) signalling the saccade target. They found that discrimination performance improved when focal stimulation was applied to neurons in the superficial layers having receptive fields covering the cued location. The authors conclude that although the SC is necessary for guiding covert attention, it is not sufficient, as similar results have been obtained through stimulation of the frontal eye fields [146], which suggests that covert attentional shifts are likely to arise through the interaction of multiple distributed cortical structures.

In addition to their involvement in signalling covert attentional shifts, the superficial layers have been shown to be involved in inhibition of return [48]. The superficial layers receive direct afferents from retinal ganglion cells and indirect projections from layer V of V1 and V2, and have efferents to PI and PL, and to dorsal and ventral LGN, while the deep layers project to PL, PM, and Pdm, as well as areas involved in saccade production, such as dorsal thalamus, oculomotor cortex and structures in the brainstem [130].

Within the deep layers, three primary neuron types have been identified: saccade-related burst neurons, that are thought to signal a saccade having the vector associated with their activity and position in the SC map; buildup neurons, which represent the preparation to make a saccade; and fixation neurons that are associated with maintaining fixation [273]. Neurons in the deep and intermediate layers exhibit multisensory response properties that result from converging inputs from cortical areas involved in processing information from different sensory modalities including visual, auditory and somatosensory

areas [48]. Similar to the receptive fields of neurons in the superficial layers, the response activation of cells in the deep and intermediate layers is produced by activity in their coarsely tuned movement fields. Movement fields have an activation gradient, such that a given neuron responds maximally to a saccade of specific direction and amplitude, termed an optimal saccadic vector [273]. Neurons with different optimal saccadic vectors are organized to form a neural map in SC, with vectors of similar direction and magnitude being grouped within a local subregion. The activity of neurons in this map of saccadic vectors is thought to collectively encode a vector of the change in eye position produced by the saccade [219]. Further, with over one quarter of the burst cells being active before a saccade [155], it is thought that the average or sum of the vectors represented by the population of active cells encodes the magnitude and trajectory of a saccade [220].

2.3 The Columnar Hypothesis

Mountcastle [150] proposed that neocortex has a relatively uniform structure composed of repeating elementary processing units called minicolumns. Unlike the laminar organization of the cortical sheet into six horizontally divided sections, these structures run orthogonal to the laminar arrangement and encompass all six layers. The basis for the columnar hypothesis comes from electrode recordings in cat somatosensory cortex, where it was observed that all cells in a given vertical penetration showed similar receptive field locations and similar response latencies [150]. After making multiple nearby penetrations, Mountcastle suggests that these columns span no more than $500\mu m$.

The minicolumn, containing 80-100 neurons across laminae II-VI, is the smallest level of vertical structure in cortex [28], and when bundled together in groups of 60-80 minicolumns, form a macrocolumn [152]. Primary visual cortex contains several specific types of columns such as ocular dominance columns, composed of cells that preferentially respond to retinal input from one eye, and orientation columns of neurons preferentially responding to stimuli of a given orientation, as well as hypercolumns, which are groups of several macrocolumns. The hypothesis proposes that these columns form distinct units, with adjacent columns having discrete boundaries, rather than a gradual transition of tuning across the cortical sheet.

While numerous past models have employed the columnar hypothesis [69, 81, 151, 163, 162, 168], the hypothesis of cortex being divided into columns with discrete boundaries is still debated more than 50 years after its original proposal [88]. In V1, a columnar organization can be observed for certain visual features (e.g. orientation, colour, ocular disparity), although a central criticism of the hypothesis pertains to the sharpness of columnar boundaries. In their pioneering studies on orientation tuning in cat area 17, Hubel and Wiesel [91] had difficulties finding evidence of such discrete boundaries, as the preferred ori-

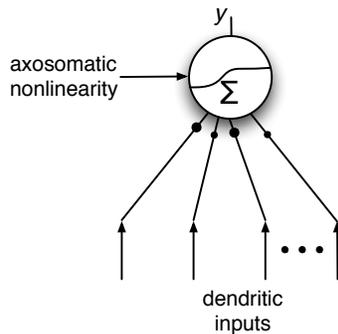


Figure 2.4: Linear dendrite neuron model. At the soma, a weighted summation of dendritic activity is passed through a global axosomatic nonlinearity to determine the cell’s firing rate.

entation of recorded cells often shifted directions rather unpredictably with small advances of the electrode, leading them to conclude, “The orientation columns, finally, if discrete, certainly fall within the definition of the term column. If orientation varies continuously with distance there are again semantic difficulties - either one must broaden the definition of the column or decide that the system is not strictly columnar” [91, pp. 289].

Regardless of there being abrupt and discrete column boundaries, or a graded transition of preferred tuning across a local cortical area of the cortical sheet, the basic idea of nearby neurons having similar receptive fields and predominately local connectivity, is a useful organizational simplification for modelling cortex. Thus, in the ARC, the broader definition of a column is used to refer to a predominately locally connected functional unit spanning the six cortical layers and having gradual horizontal boundaries that is periodically iterated spatially in each area.

2.4 Dendritic Computation

The dendrites are branching, tree-like structures that represent the primary source of electrochemical inputs to a neuron. These processes significantly increase the surface area for inputs to a neuron by a factor of $\sim 10-20$ [140], and provide a site for receiving input signals from neurons that are more distant from the cell body. Some cell types, such as pyramidal neurons, which constitute 70-90% of cortical cells [56, 144], have dendritic spines, which are a membranous extension on the dendrite that further increases the surface area.

The traditional view of dendrites is of them being passive cables surrounded by a mem-

brane having a conductance in parallel with a capacitance [102]. However, it is now known that pyramidal and other cell types also have an active component, resulting from voltage-dependant membrane conductances that arise from a relatively homogeneous distribution of sodium channels along the dendrites [102]. The functional role of these active conductances may be to mitigate at the soma, the attenuation of signals from distal synapses [140].

The classical view of neurons considers neurons as being single computational units or “point-neurons” (Figure 2.4) [134]. This view is common in models of brain function, and suggests that the combined effect at the soma of the activity of two or more synapses can be determined as a weighted sum of their individual responses, independent of their absolute or relative position in the dendritic tree. Global summation models of synaptic integration also include a single axosomatic nonlinearity, such as a logarithmic, exponential or sigmoidal function of the weighted synaptic input that determines the neuron’s firing rate.

Mathematically, this corresponds to neurons integrating inputs as a weighted sum of their excitatory and inhibitory inputs, which are then passed through a single output nonlinearity:

$$r = g\left(\sum_i \omega_i x_i\right), \quad (2.1)$$

where i indexes the presynaptic inputs, x_i is the firing rate of the i^{th} presynaptic neuron, ω_i is the synaptic efficacy or weight, and $g()$ is a nonlinear function that determines firing rate.

More recently, an alternative two-layer model has been proposed [181] based on physiological studies reporting that within-branch stimulation from two electrodes can produce sublinear, linear or superlinear responses, depending on the placement and timing of the stimulation [182]. In that study, dual-site focal extracellular stimulation was applied to various parts of the thin dendrites of neurons in rat barrel cortex while somatic voltage responses were recorded (Figure 2.5). Somatic recordings were made with electrodes targeting the same dendritic branch (within-branch) or separate branches (between-branches). In the within-branch case, the summed response at the soma was linear for weak stimulation, strongly superlinear for intermediate stimulation, and slightly sublinear for strong stimuli. Conversely, in the between-branches case, the effect at the soma was a linear summation for weak and intermediate stimulation, with a slight superlinearity with strong stimulation.

In another study, dendritic spikes evoked by focal synaptic stimulation were shown to typically remain confined to a single dendritic branch [208]. Specifically, stimulation within a single terminal basal dendritic branch evoked a sigmoidal subunit nonlinearity, but the activity summed linearly when different terminal branches were stimulated [208]. The

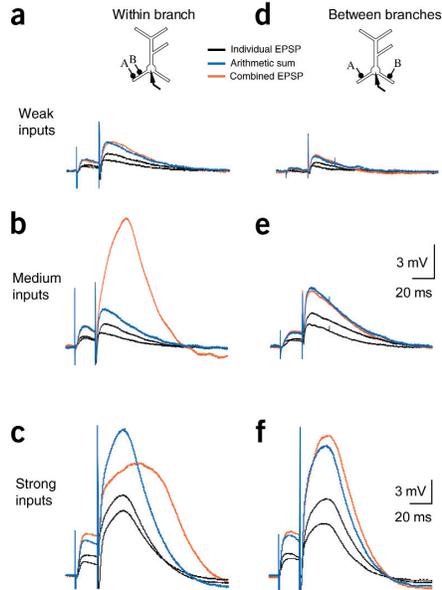


Figure 2.5: Somatic voltage traces when electrode stimulation was applied to the same dendritic branch (within branch) or to separate branches (between branch). Stimulation from two electrodes was applied individually (black traces) and simultaneously (red traces). Blue traces indicate the arithmetic sum of the two individual response. (a-c) Within branch stimulation showing linear summation for weak inputs (a), superlinear summation for medium inputs (b) and sublinear for strong inputs (c). (d-f) When separate branches were stimulated, the effect was a linear summation for weak and medium inputs and slightly sublinear for strong inputs. (Reproduced from [182]).

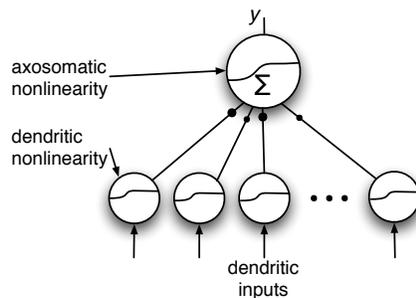


Figure 2.6: Pyramidal neuron model with nonlinear dendrites. Dendritic subunits compute a nonlinear function of their inputs and the soma receives a weighted sum of the dendritic activity, which is then passed through a global axosomatic nonlinearity. (Adapted from [181]).

two-layer model (Figure 2.6) suggests that thin basal dendrites of pyramidal cells, where the majority of excitatory synapses are located [137], act as processing subunits, where the activity of each branch is determined by its own sigmoidal nonlinearity [181, 182]. A second stage of processing then occurs at the soma, similar to the point-neuron model, where a weighted summation of the subunit activities provide the input to the axosomatic nonlinearity that determines the neuron’s firing activity.

$$r = g\left(\sum_i \omega_i s(x_i)\right). \quad (2.2)$$

In cases where $s(x_i) = x_i$, this is equivalent to the point-neuron model, or a communication channel (see Chapter 5).

As noted by Mel [140], if the subunit nonlinearity is a logical AND operation and the second nonlinearity is an OR operation, this produces a two-layer boolean logic network, whereas if the dendrites compute a product of their inputs and the second nonlinearity is a summation, this produces a sigma-pi network [141]. With a sigmoidal second nonlinearity, it effectively endows each pyramidal cell with the ability to operate in a similar capacity as a two-layer artificial neural network [181]. As has been shown in numerous artificial neural network models, any nonlinear function can be approximated to finite precision with a given number of sigmoidal hidden units, which here correspond physically to terminal dendritic branches [63]. In the ARC, it is shown how this nonlinearity can be exploited to result in efficient attentional routing.

Chapter 3

Models of Visual Attention

As noted in the introduction, visuospatial attention involves numerous processes including selecting a target, selectively processing information about this target, and limiting the processing of extraneous information. This chapter outlines several prominent models of attentional routing and discusses their strengths and limitations in terms of their biological plausibility, level of detail, and ability to explain experimental results. Although the past two decades have seen significant amounts of data being collected about the neurophysiology of attention, most models of attention do not currently address the mechanisms of attention while preserving significant amounts of biological detail.

3.1 Selecting Attentional Targets

The problem of selecting an attentional target, based on the saliency of bottom-up stimulus features or by top-down task demands, is included in several models of attention [168, 188]. Although the model proposed in this thesis does not explicitly include such a mechanism, this section briefly reviews several approaches in order to provide context for selecting attentional targets that may be combined with the proposed model.

The feature integration theory (FIT) of Triesman and Gelade [244] is largely based on psychophysical studies of visual search, and has had considerable influence on models of visual attention since its introduction. FIT proposes that a set of topographic feature maps represent information from the entire visual field, where a given feature map may represent the topographic distribution of vertically oriented bars, or the “redness” within the visual field. These maps are then integrated to form a “master map” and a spotlight of attention is then shifted over regions of high activation in the master map. Feature maps are updated preattentively, and their contents are not consciously accessible, with conscious access to information only being available from the attended regions in the master map. Models

that utilize a master or saliency map to guide attentional deployment (e.g. [94, 168, 253]) are considered early selection models and typically incorporate elements of FIT.

The central predictions of FIT are that (a) visual search in a single feature dimension is parallel and efficient, since the activity in the corresponding feature map can direct the spotlight to the appropriate location; and (b) visual search for conjunctions of features is serial and inefficient (i.e. search time increases linearly with the number of distractors) since the spotlight cannot be guided by a single feature map, and therefore elements in the display must be visited in sequence.

Since its introduction however, several problems with FIT have been noted. In its original form, a detailed account of how information from each of the feature maps is combined to form the master map is not provided, although later implementations (e.g. [94, 107]) attempt to specify this mechanism. Several studies have also provided evidence of efficient conjunction search, combining binocular disparity and colour [157], or motion and shape [135], which conflicts with the second prediction of FIT. Additionally, in feature and conjunction searches, distractors often have homogeneous orientation and/or colour, although when inhomogeneous distractors are used (e.g. distractor having different orientations in an orientated line search), there is little difference in subjects' performance when searching for elementary features and for feature conjunctions [52]. That study also found that search efficiency decreases when targets and distractors are increasingly similar, and when there is decreasing similarity between the distractors themselves.

Koch and Ullman's model of selective attention [107] seeks to provide a more detailed implementation of Triesman's FIT. As with FIT, attention is guided by a master or saliency map that is created by combining several feature maps, and suggest that the saliency map might be located in V1 or LGN. A winner-takes-all (WTA) network determines the most salient location and projects information from this region to a representation where local spatial relationships are maintained, and recognition proceeds on this representation. During visual search, locations winning the competition are attended and are then inhibited after being processed (inhibition of return), so that the focus of attention may shift to the next most salient location. However, the model is presented as a high level abstraction without much biological detail, and does not describe how the individual feature maps are represented in neurons and combined. Additionally, the model does not attempt to address a crucial element of attentional processing, namely the attentional routing mechanism, noting only that routing can be achieved by removing "some tonic inhibitory influence" or increasing the amount of excitation at the target [107].

The model of Itti, Koch and Niebur [94] extends Koch and Ullman's model in more detail by extracting visual features in parallel using 42 topographic feature maps representing orientation, intensity and colour at six spatial scales. At each spatial scale, feature maps are computed by linearly filtering the input image, convolving the filtered representation with a centre-surround structure and then normalizing. Feature maps from all spatial scales

are summed across scales by reducing each map to a common scale to form conspicuity maps of each feature dimension (orientation, intensity and colour), and these conspicuity maps are then linearly summed to produce a saliency map. Selection and inhibition of return of attended locations proceeds using the saliency map in the same manner as the Koch and Ullman model. Walther and Koch [261] further extend this model to infer the spatial extent of a “proto-object” at the attended location from the maps that are used to compute the saliency map.

The anatomical location of a saliency map has been ascribed to many cortical and subcortical areas by different groups including LGN [107], pulvinar [192], posterior parietal cortex [74, 168], lateral intraparietal area [108], V1 [116], frontal eye fields [83, 239], or distributed throughout the oculomotor system [45]. Although the ARC does not explicitly specify the process of selecting attentional targets, it is a straightforward extension for the model to incorporate any of these selection models. In this case, the output of the saliency model would be sent to pulvinar, from where the size and position of the target would be projected to PIT.

With this in mind, the remainder of this chapter addresses models of the subsequent stage of attentional processing, that of selectively processing information about the attentional target.

3.2 Approaches to Selective Information Routing

At the neuronal level, to selectively route visual information about an attended object between two adjacent cortical areas, there are several general approaches that may be taken (Figure 3.1). The need for selective processing is particularly evident in higher ventral stream areas where receptive fields can span more than 70° of the visual field [200]. At a given moment, such neurons may have multiple objects within their receptive field, yet the majority of these objects are irrelevant for the current task and should be ignored. How then, is this filtering of extraneous information performed at the cellular level?

Ignoring for the moment that the visual system is a hierarchy of multiple interconnected areas, we can first consider how visual information may be routed between two adjacent areas (e.g. V4 and IT). Based on the similarities of neurons in different ventral stream areas, as well as their inter- and intra-areal connectivity, it is presumed that the mechanisms for selective routing in one area are also employed in other areas as well. If we further presume that evolution has influenced the formation of these mechanisms to minimize the system’s energy consumption in terms of the number of neurons involved in controlling the routing, as well as the length of their axonal and dendritic processes, some of these approaches are more suitable than others. And, while limiting the resource requirements is an important factor for models of attentional routing, so too is ensuring that the model

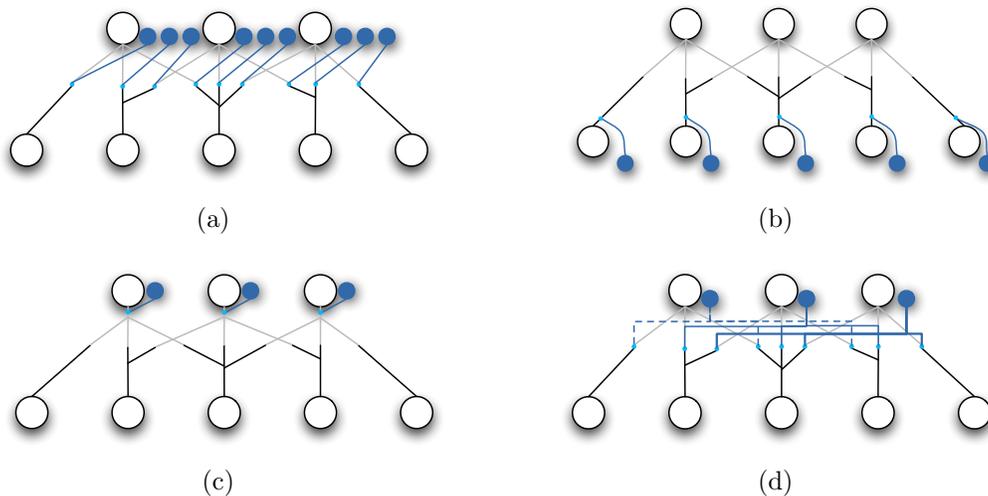


Figure 3.1: Four possible implementations of routing at the single cell level. (a) Control signals individually modulate each connection between input and output neurons. (b) Control signals modulate the output of a neuron, affecting all of its connections equally. (c) Modulating inputs to a neuron. (d) Modulating groups of connections to a variety of neurons with one controls signal. See text for details.

maintains its functional ability to perform this operation, and to do so with sufficient flexibility to operate in the variety of conditions that may be encountered.

Maintaining local spatial relationships at each level of the visual hierarchy is not agreed upon by different models of attention, with some models preserving them [168, 169, 82], while others do not [162, 253]. As suggested by Koch [105], models that maintain local spatial relationships would predict that neurons encoding the attentional target will have receptive fields that span a portion of the target, and that there is a retinotopic organization among these neurons. Models that do not maintain spatial relationships however would predict that those neurons may have receptive fields that span the entire attentional target, without a topographic organization among them.

A naïve approach to cortical routing is to dynamically modulate the connection strengths of all connected neurons between levels, such that only a subset of inputs to a given cell are driving its activity, with other inputs being inhibited (Figure 3.1(a)). However, to perform routing using this approach requires that feedforward connections from neurons in each input column are given a unique control signal, which places considerable demand on the neuronal resources in a way that quickly becomes intractable with larger networks. To effect these changes of a connection strength, existing models have proposed that either the synaptic strength is modified [168, 188], or that it may be modulated by imposing a

multiplicative gain signal on the connection [189, 204].

A second approach is to have the attentional control signals modulate the outputs of neurons in the lower level. This could be accomplished by having the control neurons form inhibitory synapses on the axon hillock or in the perisomatic region of the axon (Figure 3.1(b)). This arrangement would attenuate the output activity of the cell to all postsynaptic targets in the higher level. Although this approach provides a means of inhibiting the propagation of information from distracting and irrelevant stimuli, it is particularly problematic at higher cortical levels where receptive fields are large. Since a given neuron may project to thousands of cells in multiple columns in the higher level, such a configuration will only be able to alter the routing of visual information in a coarse manner, as the postsynaptic neurons will receive similarly modulated inputs, thereby precluding an undistorted and selective routing of the attentional target.

A third method to perform selective routing is similar to the inverse of the second approach, wherein control neurons modulate all feedforward inputs to a cell equally, and could be implemented by control neurons synapsing on the dendritic truck near the soma of the higher level neurons (Figure 3.1(c)). This has the effect of modulating all inputs received by the higher level neurons in a similarly coarse manner, without the selectivity required to minimize distortion of the stimulus' representation.

A fourth approach is to have each control signal modulate the efficacy of groups of connections. The configuration shown in Figure 3.1(d) is similar to that taken by the shifter circuit [168, 169]. In this example, each control neuron modulates the inputs from the same part of each output neuron's receptive field, which allows the attentional target to be translated between levels, although scaling the stimulus introduces additional distortion to the representation. To perform scaling without this distortion requires either additional control neurons, which increases the neuronal demand and may exceed biological constraints, or a different organization of the control signal projections. The ARC follows this last approach by using the same control signal for all connections of a column, but transforming it for each input column, and is discussed in greater detail in the following chapter.

3.3 Gain Fields and Dynamic Routing

3.3.1 Gain Fields

Numerous physiological studies of attention have reported a multiplicative modulation of neuronal responses with changes in attention [34, 35, 36, 131, 206, 245]. Gain field models suggest that such changes in activity can be fit by the product of two functions, one that

is based on the position of the stimuli in a neuron’s receptive field, and another that is based on the position of the attentional target.

The gain field model of Salinas and Abbott [204] consists of an ensemble of V4 neurons providing feedforward input to a single IT neuron. The V4 neurons are tuned to four orientations and three spatial frequencies, and their activity is computed as the product of the luminance of stimuli in their receptive field and an attentional gain field that depends on the position of the attentional target:

$$v_i = F_i(\mathbf{a}_i, I)G(\mathbf{y} - \mathbf{b}_i), \quad (3.1)$$

where v_i is the firing rate of neuron i , \mathbf{a}_i is its receptive field centre, I is the input image and $F_i(\cdot)$ is the output of the Gabor-like visual filter. The attentional gain field $G(\cdot)$ is a Gaussian function of the attentional target \mathbf{y} , and the cell’s preferred attentional locus, \mathbf{b}_i .

The response of the single IT neuron is then determined by taking a weighted sum of its V4 inputs, v_i , subtracting a threshold θ , and rectifying the result:

$$V = \left[\sum_i W_i v_i - \theta \right]_+, \quad (3.2)$$

where $V = 0$ if the summed inputs are less than 0.

In Womelsdorf et al. [272], a simple gain field based model was proposed to account for their observation of receptive fields in MT shifting and shrinking with changes in attentional targets. Their model consists of two levels, with feedforward visual inputs having a Gaussian connectivity profile with receptive field centre x_R , amplitude A_R and width σ_R . The feedforward inputs interact multiplicatively with a Gaussian shaped attention function having a centre x_{att} , gain A_{att} and width σ_{att} . Responses of MT neurons are computed as:

$$R = 1 + A_{att} \exp \left[-\frac{(x_R - x_{att})^2}{2(\sigma_{att})^2} \right]. \quad (3.3)$$

This multiplicative interaction of a Gaussian shaped attentional function is closely related to that used in the Salinas and Abbott model [204] and in the ARC. However, these past models have significant limitations that remain to be addressed. Both the Salinas and Abbott [204] and Womelsdorf et al. [272] models focus on explaining the mechanisms by which the multiplicative interactions of a Gaussian shaped attentional gain field and visual inputs can yield changes in the effective portion of a neuron’s receptive field. Both models also only consider the effects of attention between two adjacent cortical areas, without addressing the relationship of attentional modulation in different levels of the visual hierarchy. Further, neither model addresses how the receptive field profile modulations are related for neurons in other columns of that cortical area or how the attentional gain fields

are generated in the context of a large scale model. The Womelsdorf et al. [272] model is a purely mathematical implementation without spiking neurons, thereby precluding the assessment of it being plausibly implemented in the brain.

3.3.2 Shifter Circuit

Several models of visual attention have been proposed that utilize control neurons to dynamically route information under attentional influence and to model the dynamic changes in receptive fields, although perhaps the most neurally oriented is the shifter circuit [7, 167, 168, 169]. The key principles of the model are: 1) the flow of information is dynamically controlled across multiple hierarchical levels; 2) spatial relationships within the focus of attention are preserved; 3) attentional control is performed by dynamically modulating synaptic weights; and 4) scale and position invariance is handled by routing information from the focus of attention (FOA) to a 30×30 object-centred reference frame.

The basic architecture of the model is shown in Figure 3.2. Sample points in the input layer represent the retinal image that is sampled at three spatial resolutions by three separate stacks, each containing the same number of nodes. Nodes in the shifter circuit can be thought of as simplified columns. The bottom layer of each stack samples the retinal image, which is transformed at each stage by an attentional control signal to form an object-centred representation at the top layer, which corresponds to inferotemporal cortex. Salient locations are determined using a simplified version of the Koch and Ullman [107] mechanism by blurring the image into blobs and selecting the largest and brightest blob as the attentional target. The model can operate in blob search mode, which models visual search by having feature values of the target influence the saliency calculations, or in object recognition mode, where a signal indicating the spatial location of the target is projected to the saliency map. Control signals at each level modulate the connection strengths between levels as information is remapped from one level to the next. The visual information represented by node μ in level l , X_μ^l , is determined by combining its afferents X_μ^{l-1} with a control signal C_λ^l :

$$X_\mu^l = \sum_v W_{\mu v}^l X_\mu^{l-1}, \quad (3.4)$$

where $W_{\mu v}^l = \sum_\lambda C_\lambda^l \Gamma_{\lambda \mu v}^l$, λ indexes the control blocks, $\Gamma_{\lambda \mu v}^l$ is a control block specifying the amount by which λ modulates $W_{\mu v}^l$, and v indexes the afferents to μ .

Figure 3.3 shows a connection space diagram corresponding to the network in Figure 3.2, where an \times indicates a synaptic connection between a node in the input level (horizontal dots) and the node in the above level to which it projects (vertical dots). Level

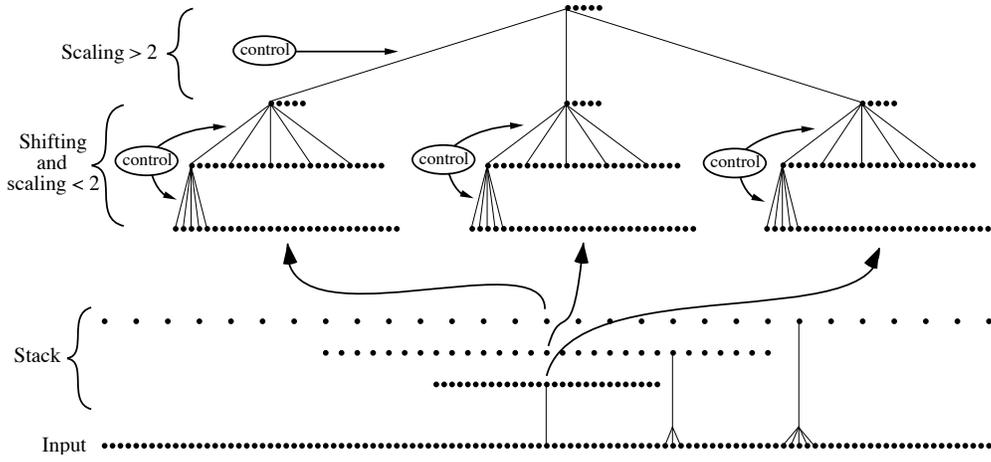


Figure 3.2: Shifter Circuit architecture. The model dynamically routes attended stimuli to an object-centred reference frame to provide shift- and scale-invariance. Three stacks receive retinal inputs with each stack having different spatial distances between sample points. Within a stack, the attentional target is shifted or scaled down by a factor of < 2 ; if a greater downsampling is required, the next larger stack is used. (Reproduced from [167]).

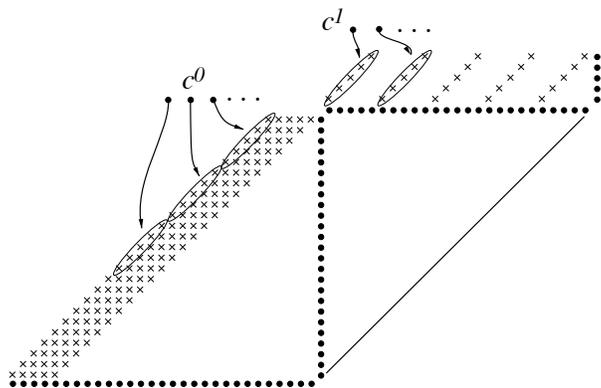


Figure 3.3: Shifter Circuit connection space diagram. Input nodes are arranged horizontally and output nodes vertically. Connections between input and output nodes are shown with an x, and control blocks that modulate connection strengths for selective routing encircle several connections. (Reproduced from [167]).

2 is represented twice, as it receives output from Level 1 and projects to Level 3. Here, an ellipse around several connections indicates a control block C_{λ}^l , which has a Gaussian profile centered in the control block, with all other diagonals having the same organization of control blocks. Each control block receives a scalar signal that is multiplied by the control block’s Gaussian weights to determine the effective connection weights for that block. To downsample the target’s representation between levels, multiple control blocks along a diagonal in connection space that correspond to the position of the target’s representation in each level are assigned positive control signals. This approach allows the values in one level to be approximately interpolated in the next. However, combining the representation of numerous inputs also causes the representation to be blurred at the next level, particularly when using larger control blocks to reduce the neuronal requirements, when downsampling by larger amounts, and in higher levels where connections are increasingly separated. To mitigate this effect, and to approximate the retinal sampling density, the model incorporates a multiscale lattice, where each stack contains approximately the same number of sample points, but lower resolution stacks cover increasing amounts of the periphery. This allows a single lattice to resize information from the target’s representation by a factor of two, with larger resampling being performed by the next larger lattice.

For a single level in each lattice, the number of control blocks is approximately $N_B \times fanIn$, where $N_B = \lfloor layerSize/w \rfloor$, $w = \lfloor moduleSize/2 \rfloor$, and $fanIn$ is the number of inputs to each node. For each stack shown in Figure 3.2, there are five control blocks at stage 2 and 25 for stage 1, where a stage refers to two adjacent levels and the control blocks that operate on the connections between them. Further, a module refers to a group of spatially adjacent nodes, their input nodes, and the connections between them, with each stage being composed of multiple partially overlapping modules. For the network shown in Figure 3.2 and 3.3, a module consists of nine input nodes and 5 output nodes.

The shifter circuit provides a compelling explanation of how dynamic routing can be employed for selective attentional processing, and approaches the problem from a more neurally oriented perspective than many computational models. Although the shifter circuit provides a good starting point for understanding the properties of attentional routing, it also has several limitations. As with most computational models, the shifter circuit lacks details of the computations and representations of cortical and subcortical neurons, which limits its explanatory and predictive capabilities. Specifically, it is a purely mathematical implementation without spiking neurons, which thereby precludes evaluating the plausibility of its neural implementation and the detailed neuronal requirements for the calculation, representation and communication of visual and control signals.

Second, when the model is scaled to approximate the size of visual cortex, the neuronal resources required by the model for computing control signals exceeds that available in the pulvinar. Specifically, Selemon et al. [211] estimate that pulvinar contains 2.52×10^6 neurons, and Petersen et al. [179] report that many cells in pulvinar fail to exhibit

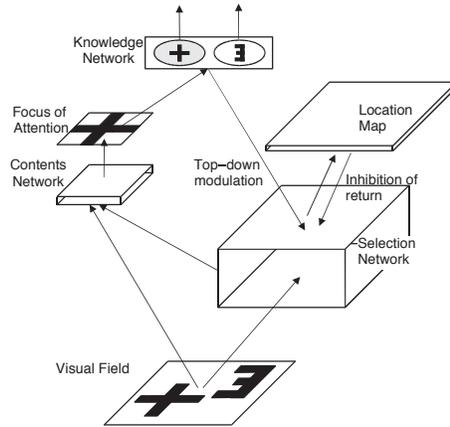


Figure 3.4: SAIM architecture. Visual information is presented to the system’s visual field from where it is projected to the contents network, where gating signals from the selection network interact with the visual signals. The selection network is biased for spatial attention by inputs from the location map or for object-based attention by the knowledge network. The focus of attention represents the target in an object-centred reference frame and projects to the knowledge network where template matching object recognition is performed. (Reproduced from [81]).

any significant attentional effects. Olshausen et al. [169] propose that for a full brain-scale shifter circuit for spatial attention, approximately 800,000 unique control signals would be required, which suggests that approximately three pulvinar neurons are available for computing and projecting each control signal. If it is assumed that each pulvinar neuron in the shifter circuit has a signal to noise ratio on the order of 10:1 (see Eliasmith and Anderson [55]), the model would in fact require $\sim 8 \times 10^6$ pulvinar neurons, which significantly exceeds the number of available neurons. In short, it is unlikely that only 3 neurons could sufficiently accurately encode the needed control signals in the shifter circuit. Third, the shifter circuit performs dynamic routing by modifying the synaptic connection weights, although there is little evidence that synaptic plasticity can occur at the time scale required for attentional processing.

3.3.3 Selective Attention for Identification Model (SAIM)

The SAIM model [81] follows the idea of the shifter circuit, but focuses on modelling the cognitive aspects of attention and effects of lesions, with less emphasis on biological correspondence. The architecture of SAIM is shown in Figure 3.4. Visual information enters the model in the visual field, from where it is projected to a selection network and

contents network. The selection network serves to selectively gate and route information from the visual field to the focus of attention, and contains a $k \times l \times i \times j$ matrix of connection strengths, where k and l represent the spatial size of the visual field and i and j are the spatial dimensions of the focus of attention. This configuration is similar to the first proposed routing mechanism shown in Figure 3.1(a). The location map is similar to a saliency map, in that it encodes potential or previously attended locations, but differs by not directly engaging in any competition between targets, and instead biases the selection of targets within the selection network. The location map is also involved in inhibition of return, wherein the salience of previously visited locations are suppressed. The focus of attention in the model is consistent with the notion of an object-centred reference frame, and provides gated visual signals to the knowledge network, where object recognition is performed using template matching.

Information from the visual field is selectively routed to an object-centred reference frame in the focus of attention by the selection network projecting gating signals to each possible connection between the contents network and focus of attention, where the control signals are multiplied by the signal carried along each connection, and the product is summed by each unit in the focus of attention. A central constraint of the model is that spatial relationships are preserved, and this is accomplished by imposing two additional constraints that each retinal input projects only to a single unit in the focus of attention, and that each unit in the focus of attention receives input from only one retinal unit.

More recently, Heinke et al. [82] extended the SAIM model to use spiking LIF neurons, aiming to close the gap between the behavioural and neurobiological levels of analysis. The model was shown to qualitatively reflect the behavioural effects of attention in normal and lesion patients, but has several limitations. The model addresses position invariance but not scale invariance, and assumes that at any one time, neurons in a given column receive inputs from exactly one input column, with all other inputs being fully suppressed. The visual hierarchy is also reduced to consist of two interconnected levels, with a large matrix in the selection network controlling the precise routing of each possible pathway between the two levels. The model also has scalability limitations, as incorporating additional visual processing levels in the model requires additional high dimensional matrices to be represented and transformed.

3.4 Competition Based Routing

3.4.1 Biased Competition

The biased competition model is another proposal for how attended visual stimuli may be selectively processed in cortex [43, 188]. The model proposes that the appearance of two

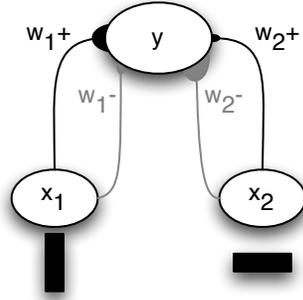


Figure 3.5: Biased competition model. The neuron on top receives excitatory (black lines) and inhibitory projections (grey lines) from input populations, shown at the bottom. Each input population represents separate stimuli, shown here as vertical and horizontal bars below them. (Adapted from [188]).

or more stimuli in a neuron’s receptive field activates neuronal populations that automatically compete with each other, and that directing attention to one of the stimuli biases this competition in favour of the attended stimulus [43]. The assumption of an intrinsic competition in cortex has been incorporated in several models of attention [199, 253], although a detailed account of the hypothesis was largely advanced in the computational model of Reynolds et al. [188].

A schematic of their feedforward model is shown in Figure 3.5. Two separate stimuli, shown as vertical and horizontal bars, activate populations of input neurons (ovals at the bottom of the figure) that provide feedforward input to the single output neuron, shown as an oval at the top of Figure 3.5. Input neurons i with firing rates x_i are connected to the output neuron with excitatory and inhibitory connections having weights w_i+ and w_i- respectively. The excitatory and inhibitory inputs to the output cell are determined as the sum of the activities of the input cells, multiplied by the excitatory and inhibitory weights, respectively:

$$\begin{aligned} E &= x_1 w_1^+ + x_2 w_2^+ \\ I &= x_1 w_1^- + x_2 w_2^- . \end{aligned} \tag{3.5}$$

The activity of the output cell, in terms of firing rate, is computed then as:

$$y = \frac{BE}{E+I+A}, \tag{3.6}$$

where B is the maximum response of the cell and A is a passive decay term. The model assumes that attention causes the strength of signals from attended stimuli to increase in a

multiplicative manner, and is implemented by increasing both the inhibitory and excitatory synaptic weights from the attended stimulus by a factor of 5. This method of selective routing is similar to the first approach presented in Figure 3.1(a).

The article in which the model was first presented also reported results from several single cell recording experiments in V2 and V4 with different attentional and stimulus conditions, with which the model was shown to be qualitatively consistent [188]. Specifically, Reynolds et al. had three monkeys attend to a target outside the recorded cell's receptive field and measured the activity when one or two stimuli were placed inside its receptive field. With a preferred and non-preferred stimulus in the receptive field, shifting attention to one of the targets caused the cell to respond at a rate that was similar to that elicited by the presentation of the attended stimulus alone. That is, with both a preferred and non-preferred stimulus in the receptive field, attending to the non-preferred stimulus yielded a lower firing rate than that elicited by the preferred stimulus being alone, yet higher than the firing rate elicited from the presentation of the non-preferred stimulus alone.

While the model was shown to capture some qualitative properties of attentional modulation in cortex, it has several limitations. As with many other models, the biased competition model is a purely mathematical abstraction of the computations, rather than the biological mechanisms, that may lead to selective routing. Second, the means of applying the attention signal by changing synaptic weights is debatable on the time scale needed for attentional routing. In their implementation, an attention influence was provided by increasing the synaptic weights by a factor of five, although the basis for this value, along with its generation in the context of a large-scale model, are not specified. This is effectively a multiplicative modulation of attended information, for which there is much experimental support. However, there are a few key differences in their approach and a gain modulation. As was discussed in Chapter 1, attention appears to attenuate unattended stimuli rather than to amplify attended stimuli [147]. The model also does not specify the basis for the value of the weight increase, nor how it is related to attentional modulation in other neurons in that area. Finally, similar to the gain field model of Salinas and Abbott [204], the model is only specified for a single output neuron, without detailing how attentional control applies in other ventral stream areas.

3.4.2 Normalization Model of Attention

The normalization model of attention [189] is an extension of the biased competition model that combines aspects of several previously proposed models [80, 188]. The model seeks to define, in a single computational model of attention, the mechanisms that can account for numerous forms of attentional modulation including contrast gain, response gain, sharpening of tuning curves, and reductions in firing rate with an additional non-preferred stimulus in the receptive field. The model comprises three general components that operate on a one

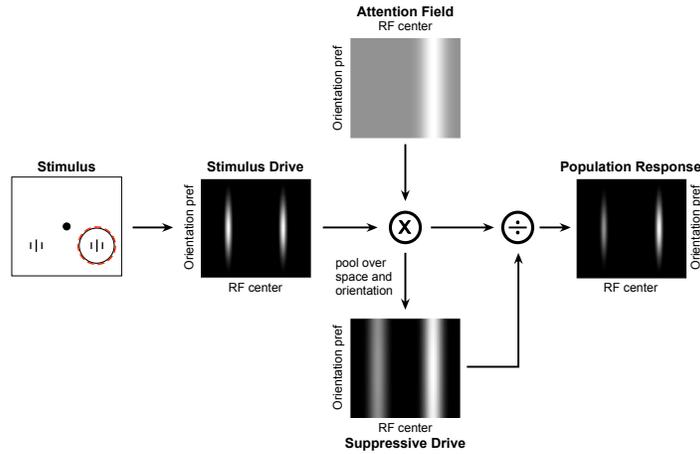


Figure 3.6: Normalization Model. Visual stimuli generate the stimulus drive and the attention field specifies the features (y -axis of attention field) or locations (x -axis) of the attentional target. The product of the stimulus drive and attention field is then convolved with a Gaussian to yield the suppressive drive, which divisively normalizes the activity to produce the population response. (Reproduced from [189]).

dimensional visual field and for one feature dimension: a stimulation field, a suppressive field and an attention field.

The stimulation field characterizes the response of a neuron to visual stimuli in its receptive field, based on the stimulus' feature values (e.g. orientation) and spatial position, without any attentional modulation. The suppressive field characterizes the spatial positions and feature values of a stimulus that can cause a suppressive effect on the stimulus evoked activity, or stimulus drive. The attention field describes the multiplicative gain that is applied to the stimulus drive of all neurons, based on the feature or spatial properties of the attentional target, as specified by a top-down mechanism. In cases of feature-based attention, the attention field has larger values for the target feature (i.e. $\gamma > 1$, where γ is the maximum value) across all spatial positions, whereas for spatial attention, the map has larger values for the target position across all feature values.

The population response represents the activity of neurons in the output level and is a function of the attention field, specifying the position and/or features of the target, and the stimulus drive, as determined by the visual features of the stimuli. The suppressive field is effectively a Gaussian blurred version of the stimulus drive and serves to divisively normalize the product of the attention field and stimulus drive.

Figure 3.6 illustrates the arrangement of these three components for one spatial dimension of a single cortical area (e.g. V2, V4 or IT) when attention is directed to a target on

the right side of the visual field (spatial attention). The leftmost box shows two vertically oriented stimuli in the visual field, with one stimulus in each hemifield. The receptive field of one cell is shown as a dashed circle, and the attentional target is outlined in red. The x -axis of each box indexes the spatial position of a neuron's receptive field centre, while the y -axis spans the range of preferred orientations at that location. The attention field has higher values, shown in white, for stimuli of all orientations, at the target location. The firing rates of the population, R , as a function of the stimulus drive and suppressive drive, are given by:

$$R(x, \theta) = \left| \frac{E(x, \theta)}{S(x, \theta) + \sigma} \right|_T, \quad (3.7)$$

where $E(x, \theta)$ is the stimulus drive of a neuron with receptive field centre x and orientation tuning θ , and $S(x, \theta)$ is the suppressive drive for that cell. σ is a constant that determines the neuron's contrast gain, and $|\cdot|_T$ represents rectification of the response with threshold T .

As with several other models [168, 188, 204], selection of attentional target is not handled by the model, but is provided to the system as $A(x, \theta)$ for all receptive field positions and preferred orientations. In cases of visuospatial attention, $A(x, \theta)$ has values of $\gamma = 1$ for all receptive field positions and orientations, except at the spatial position of the target, where it has a value of $\gamma > 1$. The peak of the attention field, γ can have various values, depending on the strength of attentional modulation required by the task. In cases where $\gamma = 1$, the attention field has equal strength everywhere, resulting in the stimulus drive of all neurons being multiplied by 1. Further, that $\gamma > 1$ for attentional targets, rather than $\gamma < 1$ for non-targets (i.e., enhancement of the target instead of suppression of distractors), is similar to the biased competition approach of increasing the weights of the attended stimulus to a value > 1 .

The effect of attention is to multiply the stimulus drive by an attentional gain term specified in the attention field, which in turn influences the suppressive drive. Incorporating the attention field, the resulting activity of the population of neurons is then expressed as:

$$R(x, \theta) = \left| \frac{A(x, \theta)E(x, \theta)}{S(x, \theta) + \sigma} \right|_T, \quad (3.8)$$

$$S(x, \theta) = s(x, \theta) * [A(x, \theta)E(x, \theta)], \quad (3.9)$$

$s(x, \theta)$ is the Gaussian kernel, $*$ indicates convolution, and $\int s(x, \theta) = 1$.

The model was shown to account for several forms of attentional modulation in cortex, including contrast and response gain. Contrast gain, the model proposes, arises in cases where the stimulus target is small and the attention field is large. The response of a neuron is then determined as:

$$R(c; x, \theta) = \left| \frac{\gamma E(x, \theta; c)}{s(x, \theta) * (\gamma E(x, \theta; c)) + \sigma} \right|_T, \quad (3.10)$$

where c is the stimulus contrast and γ is the peak of the attention field. Because the region of the attention field having values $\gamma > 1$ is large, it effectively multiplies the entire stimulus by γ , thereby shifting the contrast-response curve leftward for intermediate contrasts before saturating at higher contrasts. Assuming that all neuronal responses are proportional to contrast c and that the large attention field has a value of γ everywhere, Equation 3.10 can be rewritten as:

$$\begin{aligned} r(c) &= \alpha(\gamma c) / (\gamma c + \sigma) \\ &= \alpha c / (c + \sigma / \gamma), \end{aligned} \quad (3.11)$$

where α is the maximum response.

Response gain on the other hand, is proposed to occur in scenarios where the stimulus is large and the attention field is smaller than the stimulus. The effect of the small attention field is to multiply only the centre of the stimulus by γ , which also produces a similar effect to that created by decreasing the stimulus size, as it reduces the spatial spread of the stimulus drive. Thus, by using stimuli having contrasts with a Gaussian spatial profile (i.e. highest contrast at their centre), only the centre of the stimulus is affected by the attentional gain, while the edges of the stimulus having lower contrasts are unaffected.

To model response gain, the activity of a neuron with the stimulus and attention field at the centre of its receptive field is computed as:

$$r(c) = \alpha(\gamma c) / (\gamma c + \beta c + \sigma) \quad (3.12)$$

where $B \in [0, 1]$ scales the suppressive drive from the area surrounding the stimulus field. For low contrasts, the response to contrast c is approximately equal to $r(c) \approx \alpha \gamma c / \sigma$ where increasing γ scales the response, while for high contrasts ($c \gg \sigma$), $r(c) \approx \alpha \gamma'$, where $\gamma' = \gamma / (\gamma + \beta)$, causing saturation at higher contrasts. The model was used to simulate an experiment by Treue and Martinez-Trujillo [245], in which two stimuli were placed inside the receptive field, one moving in the anti-preferred direction, and the other moving in one of 12 directions in the range $[-180^\circ, 180^\circ]$. They found that directing attention to the variable motion stimulus produced an increase in the gain of the tuning curves, but there was no change in its width. This experiment is simulated in Chapter 6.

However, this model has several limitations. First, the model is a purely mathematical implementation without spiking neurons, and has limited predictive power in neurobiology as it stands. Although the model is not intended to provide a detailed mechanistic explanation of attentional modulation, the mathematical implementation makes it unclear how biologically plausible the computations are.

The normalization model is consistent with the experimental findings of increased directional gain when attending the variable direction stimulus, but is inconsistent with regard to the sharpening of tuning curves with spatial shifts of attention. The experiment by Treue and Martinez-Trujillo [245] that was modelled by the normalization model specifically tested whether the width of the tuning curve changes when attention is directed to the non-preferred or variable stimulus, finding that there was no significant sharpening. The normalization model however, predicts the opposite, that directing spatial attention to the variable stimulus will sharpen the tuning curve (i.e. the range of motion directions to which the neuron responds). Reynolds and Heeger suggest that this experimental result is not inconsistent with the model, as a later study by Martinez-Trujillo and Treue [126] demonstrated tuning curve sharpening. However, the later study by Martinez-Trujillo and Treue used a different experimental configuration that involved only one receptive field stimulus, and involved feature-based attention rather than spatial attention. From their simulations, they report that using a broad attention field, equal in size to the stimulus, results in the sharpening of tuning curves, and that the amount of sharpening depends on the width of the attention field.

Using their model ¹ to repeat this simulation, reducing the width of the attention field from 5 (as specified in their simulations in [189]) to 4 does reduce the amount of tuning curve sharpening. However, the smaller attention field causes only the centre of the stimulus to be gain modulated (similar to the response gain simulations), resulting in the attend-out case producing a greater peak response than attending the variable direction stimulus, which conflicts with the experimental data. Conversely, increasing the size of the attention field from 5 to 6 yields a change similar to contrast gain, further increasing the range of directions to which the neuron responds when attending the variable motion stimulus. This issue is returned to in Chapter 6, where the proposed model is used to simulate the Treue and Martinez-Trujillo [245] experiment.

3.5 Synchrony- and Oscillation-based Models

For many years, there have been reports of correlations between oscillatory activity in the gamma band (30-100 Hz, prototypically 40Hz) and mental processing, particularly attention and decision making [65, 75, 161]. Crick and Koch [39] proposed that this increased power in and around 40 Hz may play a role in binding together the components of attended stimuli by superimposing an oscillatory 40 Hz signal in the activity of neurons representing the attentional target. Their temporal tagging hypothesis was further developed in a computational model by Niebur et al. [163], which consists of a V1 and V4 module and

¹Available online at: <http://www.cns.nyu.edu/heegerlab/>

a saliency map based on the Koch and Ullman mechanism [107]. The saliency map, hypothesized to exist in superior colliculus or dorsomedial pulvinar, projects to cells in V1 having receptive fields that contain the attentional target, but not to V4. The signal from the saliency map affects only the temporal structure of spike trains from those V1 cells processing attended information, but not their mean firing rate.

For neurons in V1, the total firing rate λ is determined as the sum of spontaneous activity, λ_{spont} and stimulus dependent activity λ_0 . For neurons not receiving visual input, or stimuli having features of the cell’s non-preferred direction, $\lambda_0 = 0$, while for neurons receiving preferred stimuli, λ_0 is proportional to the amount of spatial overlap of their receptive field with the stimulus:

$$\lambda_o = \lambda_{max} \times \text{overlap}(\text{stimulus}, \text{receptiveField}), \quad (3.13)$$

where $\lambda_{max} = 200$ Hz is the neuron’s maximum firing rate and $\text{overlap} \in [0, 1]$, where 1 indicates a complete overlap. In their model, a neuron’s featural tuning is simplified to have a preferred stimulus to which it responds, with all other feature values being ignored. Attentional modulation is produced by an inhomogeneous Poisson process having a time varying mean rate of:

$$\lambda(t) = \lambda_0(1 + A \sin(\omega t + \phi)) + \lambda_{spont}, \quad (3.14)$$

where ϕ is a time-independent phase shift, $\omega/2\pi$ is the modulating frequency in the γ range (25-60 Hz) imposed by the saliency map, and $A \in [0, 0.75]$ is the strength of modulation based on the degree of spatial overlap between the receptive field and the attentional target. Individual cells are modelled using a linearized Hodgkin-Huxley model, with refractory periods drawn from a uniform distribution in the range [2,5] ms, with $\lambda_{spont} = 2$ spikes per second.

The V4 module has a columnar organization, with each column containing inhibitory and excitatory neurons tuned to various features, but having similar receptive fields. Specifically, each column contains pyramidal cells that process visual stimuli and are paired with interneurons that inhibit those pyramidal neurons that are tuned to opposing stimulus features. The model proposes that there is a competition between V4 neurons in each column representing different stimuli. The competition is biased toward neurons receiving input signals that are tagged with a γ -band oscillation imposed by the saliency map by using interneurons tuned to respond preferentially for spikes arriving every ~ 25 ms (40 Hz). These interneurons that are paired with pyramidal cells preferring a given stimulus, inhibit other pyramidal cells in that column tuned to opposing stimuli, thereby allowing its paired cell to “win” the competition.

A revised version of the model was subsequently presented by Niebur and Koch [162], which sought to identify an alternative means of imposing the temporal tag on attended

stimuli, as oscillatory firing in single cells was difficult to observe experimentally. For example, Luck et al [118] found that only $\sim 5\%$ of neurons showed evidence of oscillations in their autocorrelograms. In the revised model, no periodically recurring events were required, nor were V4 interneurons that are selective for spikes arriving every ~ 25 ms. Rather, neurons in the lower level representing the attended stimulus are tagged by having them fire synchronously in the γ range, an effect that unlike the former model would not be seen in the activity of a single cell, but only in the activity of multiple neurons. V4 interneurons then detected coincidences in their inputs, and similarly inhibit pyramidal cells tuned to opposing stimuli.

Both models predict that there will be no attentional effects on the firing rate of neurons in areas below V4, although such effects have since been reported in LGN, V1, and V2 [26, 118, 132, 138, 149, 196]. Further, Shadlen and Newsome [213] have argued that the existence of neurons in V4 or other visual areas that can detect coincident spikes at the time scale required by the model are unlikely. Several other models that incorporate rhythmically synchronized neuronal activity have been proposed to account for selective attentional processing in cortex (e.g., [10, 11, 19, 20, 25, 143, 242]), but will not be discussed in detail. Although several of these models aim to address phenomena at a similar level of detail as the ARC, few of them attempt to describe the relationship of the attentional effects between cortical areas (at most 2 areas are considered in any model) and none define the detailed laminar circuitry of their cortical implementation.

3.6 Summary of Existing Models

From the above discussions of these other models, there are several aspects of the proposals that are consistent with or incorporated in the ARC. I will return to a comparison of this past work with the presented model after describing it in the next chapter. The ARC is most closely related to the shifter circuit model [168, 169], which provides an appropriate starting point for generating a new model. Several concepts introduced by the shifter circuit are incorporated into the ARC. First, the ARC follows the notion that visuospatial attention can be considered as a problem of routing attended information through a hierarchy of visual areas. The ARC also seeks to route information from the attentional target to an object-centred reference frame using control neurons that serve to guide the routing of information between cortical areas while preserving local spatial relationships within the representation of the target.

Although the shifter circuit's proposed mechanism for performing dynamic routing by changing synaptic weights is questionable, a similar effect can be obtained through the use of gain fields. Gain modulation has been observed in numerous cortical and subcortical areas as a change in the response amplitude of a neuron, without affecting its selectivity [205].

Salinas and Abbott [205] propose that gain modulation, or gain fields, are a widespread computational mechanism that subserves the transformation of information from one coordinate frame to another. Within the ventral visual stream, attention provides a strong gain control signal, while in the dorsal stream, gaze direction serves a similar purpose.

In parietal cortex, Andersen and colleagues found that visual responses depend both on the retinal location of a stimulus and the direction of gaze [5, 6]. They demonstrated that with gaze held constant while spots of light were presented at various locations within the neuron’s receptive field, the resulting responses could be well fit by a Gaussian over the range of stimulus positions. When the animal directed its gaze to a different spatial position, similar results were obtained, but with a different amplitude. These results demonstrate that gaze direction has a multiplicative influence on neural responses that is well described by a Gaussian shaped gain field.

One functional advantage of gain fields in parietal cortex is that they allow for information to be transformed from eye-centric to head-centric and body-centric coordinate systems. In the higher levels of the dorsal stream, it has been reported that gain fields in LIP depend on gaze direction, placing information in body-centric coordinates, while in area 7a, to which LIP projects, gain fields depend on body position [217].

Functionally similar gain field effects have also been reported in ventral stream areas, where the gain fields depend on the position of the attentional target. Similar to the gaze-dependent experiments of Andersen et al. [5, 6], studies by Connor et al. [35, 36] examined the effects of attentional position on the responses of V4 neurons. As the animal maintained fixation, bars of light were flashed inside the receptive field while the animal covertly attended a target nearby, but outside the receptive field. These neurons also exhibited gain modulated response that can be characterized by an attentional gain field. Overall, then, it seems that gain fields provide a useful and well-established means of addressing how attentional multiplication might be accomplished in a shifter circuit.

The normalization model [189] provides a particularly interesting approach to understanding the neurophysiological basis of attentional processing by proposing that the many different forms and effects of attention may arise through a common mechanism. Several previous proposals have addressed only some forms of attentional modulation such as contrast gain, response gain, or receptive field modulation, although comparatively few have attempted to simulate several forms using a single mechanism. Consequently, it is a goal of this thesis to simulate several forms of attentional modulation with a single mechanism.

Across these models, there are several common limitations that need to be addressed. First, most of the above models are implemented in purely mathematical terms, without spiking neurons [168, 188, 189, 272]. Although they are able to capture the qualitative properties of attentional modulation and routing at various levels of abstraction, it remains unknown whether the necessary functions and representations can be plausibly implemented in the brain, within the constraints of timing, neural processing and storage.

In particular, constructing models with spiking neurons allows the model’s scalability to be tested; increasing the size of the visual field, the number of features being processed, and the number of connections between cells and thus synaptic weights, requires additional neurons to represent and process this information at a given precision, and it is necessary to verify that the neuronal requirements of larger models do not exceed those of the areas being modelled.

A second limiting factor with several of these models is that they describe a possible mechanism for attentional modulation subserving routing for only a single cell [188, 204, 272]. Although they serve a valuable role in advancing our understanding of attentional modulation at this level, there is a need for models to extend these proposals to specify how the control signals are generated and applied for other neurons in that area.

Of the models that do consider attentional modulation for an entire cortical area, they do not address how attention modulates activity in other areas of the ventral stream, aside from proposing that it operates in a similar manner [188, 189]. There is mounting evidence that attentional effects vary between areas of the ventral stream hierarchy, in terms of the timing of modulation, strength of modulation, and proportion of cells showing modulation [26, 138], phenomena which are not addressed by any of the models. With the recent advancements in electrophysiology equipment allowing an increasingly large number of cells to be recorded simultaneously from multiple areas, there is a need for theoretical models to provide predictions that may be tested, and to be able to model this data as it becomes available.

From the above discussions of existing models, a set of common limitations can be assembled that the proposed model seeks to address:

1. Implemented in spiking neurons, within biophysical constraints;
2. Defined for an entire cortical area, rather than for a single cell;
3. Able to account for the relationship of attentional modulation between multiple cortical areas, and the interaction of attentional routing between these areas;
4. Well-matched to physiological data; and
5. Able to account for different forms of attentional modulation using a common mechanism.

With these criteria established, the following chapter introduces the Attentional Routing Circuit (ARC), which is first described in functional terms for ventral stream areas V1-PIT. The following chapter elaborates this description by implementing the model in spiking neurons using the Neural Engineering Framework (NEF) [55], and then presents a detailed mapping of the model to specific neuron types and cortical laminae.

Chapter 4

Attentional Routing Circuit - High Level Overview

Understanding the detailed neuronal mechanisms of selective attentional processing in cortex remains an outstanding problem in neuroscience. Given the complexity of the cortical circuitry and myriad results from behavioural and physiological studies discussed in earlier chapters, attempting to understand the ways in which the brain selectively routes information can be pursued in several ways. The approach taken in this thesis is to first describe, at a high level, the functional mechanisms with which selective routing may be performed, while operating within the biological constraints previously outlined. Specifically, this chapter introduces the Attentional Routing Circuit (ARC) and describes the general mathematical basis of attentional routing, so as to provide a general sense of the proposed mechanisms. The model is further elaborated into a spiking neuron implementation in the subsequent chapter, but first, the general principles are established.

4.1 Overview of Model Structure

Figure 4.1 shows the general architecture of the ARC. The model is constructed as a hierarchy of levels that are taken to correspond to ventral stream areas V1, V2, V4, and posterior IT (PIT) for the purposes of exposition, but also map to the visual hierarchy in the dorsal stream (V1, V2, V5, PPC). Each level has a columnar and topographic organization, where each cortical column contains populations of neurons. Columns are shown with a uniform spacing between them, although more accurately, the spacing between columns and the receptive field sizes of the neurons they contain, increases at further distances from the fovea according to the cortical magnification factor. For explanatory purposes and due

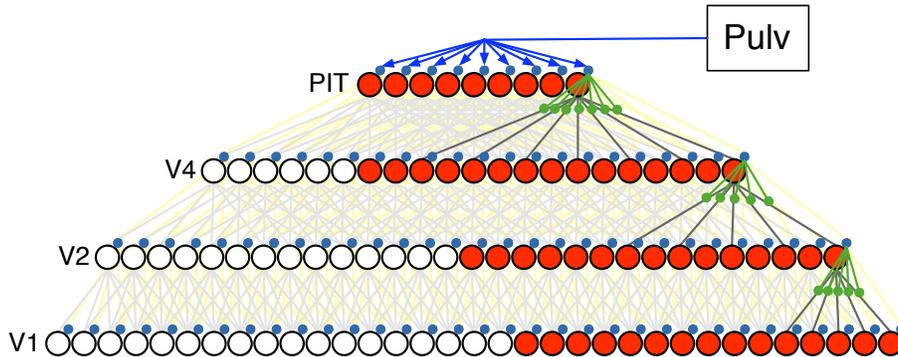


Figure 4.1: General architecture of the ARC. Each level has a columnar and retinotopic organization, where columns are composed of visually responsive neurons and control neurons (blue circles). Red circles indicate columns representing the attentional target. Each column receives feedforward visual signals (gray lines) and a local attentional control signal from control neurons (green lines), and these signals interact nonlinearly in the terminal dendrites of pyramidal cells (green circles). Global control signals from pulvina are fed back to control neurons in lower levels (yellow lines). Connectivity is highlighted for the rightmost columns only, although other columns in each level have similar connectivity.

to limitations of computational resources, the model is restricted to a one dimensional visual field, although extension to two dimensions is straightforward.

As with several past proposals [69, 81, 151, 168], the ARC adopts the view that the calculations used to determine local control signals are the same for columns in all levels. That is, the computations used to transform the global control signals to local control signals, and the means by which the local control signals are applied, are the same throughout the network.

In each column in the ARC, two types of neurons in each column can be distinguished based on their functional roles: visually responsive neurons that processes the feedforward visual signals, and control neurons that compute local control signals for dynamically routing the information that is processed by visually responsive neurons in that column. As is discussed below, these control neurons can be further divided based on the computations they perform and the cortical layers in which they are found.

Two types of control signals may also be distinguished. Global control signals indicate the size and position of the attentional target, and originate in pulvina from where they are projected to PIT and fed back through the hierarchy. The second type are local control signals that are used within the column and are derived from the global control signals. For each column, the local control signals simply specify the position within the receptive field from which visually responsive neurons should selectively process visual information.

The number of columns in each level of the ARC is based on the anatomical sizes of the cortical areas being modelled, and so the number of columns decreases exponentially at successively higher levels [271]. Given the considerable variance of the average surface area of visual areas in human cortex [49], this is a simple assumption, though one that does not affect the principles of the model. All columns in a given level receive the same number of inputs from the previous level, although in higher levels, neurons in each column receive input from a greater number of non-contiguous columns in the previous level, as demonstrated by numerous anatomical studies [57, 58, 60, 236]. Specifics of the connectivity between columns and neurons within columns (e.g., the number of columns providing input to a column or neuron) are not critical to the model’s routing mechanisms, since all of the control calculations are performed using locally available information and the same low-dimensional global control signals originating in pulvinar. In fact, the principles of the ARC still hold in networks having additional or fewer levels, with a linear or arbitrary reduction in the number of columns in each level, and with varied receptive field positions and sizes for individual neurons in a column.

In numerous areas of visual cortex, visually responsive neurons within a column have similar spatial receptive fields, but may respond preferentially to different visual features at that location. For simplicity, neurons in each level of the model are sensitive only to stimulus intensity. This simplification is made because although there is a fairly good characterization of the tuning and selectivity of neurons in V1 and MT, our understanding of the stimulus parameters needed to describe the complex receptive fields and tuning in the three dozen other visual areas is considerably less. Thus, there is no attempt to specify the features represented by these neurons. Instead, and each visually responsive neuron may be thought of as encoding a feature vector, in this case a scalar value corresponding to stimulus intensity. If more complex stimuli are encoded, these feature vectors would be defined in a higher dimensional space, with each dimension corresponding to a particular feature. Such an arrangement is found in models of object recognition and machine vision. None of the ARC’s principles prevent its extension to processing such higher dimensional spaces.

Figure 4.1 shows the primary elements for routing in the ARC. Red filled circles indicate columns in each level that encode visual signals from the attentional target and unfilled circles are columns operating in their default routing state and not encoding information from the target. Each column contains cortical control neurons (small blue circles) that compute local control signals for the visually responsive neurons in their column. The local control signals are computed based on the global control signals from pulvinar (blue lines), which projects to control neurons in PIT. Feedback projections from control neurons in PIT relay the global control signals to control neurons in the previous level (yellow lines). The local control signals are projected to the inputs of the visually responsive neurons (green lines) where they modulate the gain of the input signals from each input column.

Feedforward inputs to a column are shown with gray lines.

It is presumed that in the absence of, or prior to the arrival of a global attentional control signal from pulvinar (e.g. when the target’s location is not cued or known before stimulus onset), the system operates in a default routing state, wherein the entire visual field is routed to form a coarse representation in PIT. Since the retinal sampling density and cortical magnification factor cause more neurons to represent the foveal and parafoveal regions, information from these regions will be represented in PIT with greater fidelity than will peripheral regions. When the target is known, columns that are not encoding information from the attentional target still operate in their default routing state, and this assumption is further discussed in Section 4.5.

Following the assumption of control calculations being the same in all ventral stream areas, the following section presents the methods for computing the local control signals for a single column in the top most level, with columns in other levels employing the same mechanism.

4.2 Control Calculations

For columns in each level, cortical control neurons serve three general functional roles. First, control neurons determine the size of the attentional target’s representation in the level below, in terms of the number of columns that it will span. In determining this value, the model seeks to minimize the loss of visual information by having as many columns as possible represent signals from the target.

Second, control neurons determine the position of the target’s representation in the previous level. In computing the location of the target’s representation in the previous level, the model attempts to have the target’s representation centred at the lowest level as possible. That is, if the attentional target is located in the far periphery in V1, at each level, its representation is drawn closer to the centre of that level. This axiom follows a core concept of the model, namely minimizing the loss of information from the attentional target. Neurons in numerous visual cortical areas have Gaussian shaped receptive fields, with sensitivities that fall off at the edge of their receptive field [35, 36, 171]. By shifting the target’s representation toward the centre at each level, this effect can be minimized so that neurons, particularly those in higher levels, are selectively processing information from the centre of their receptive field.

Once the control neurons have determined the size and position of the attentional target’s representation in the previous level, the third function performed by control neurons is to determine the portion of the attentional target that each column should selectively process. This function is of particular importance in the ARC, as it ensures that only relevant

Table 4.1: Parameters used in the ARC

Variable Name	Usage
A_{len}	Length of attentional target
A_{pos}	Position of attentional target
θ_l	Position of attentional target in level l
a	Receptive field size
m	Maximum shift
s	Shift
sf	Sampling Factor
i	Position index of output column
j	Position index of input column
μ	Attentional target within receptive field
x_j	Input column at position j
\hat{x}_j	Visual signal carried by input column
$size$	Number of columns in level
σ_{att}	Width of routing function
σ_ω	Width of Gaussian connection strengths

information (i.e., the target) gets selectively processed, and that extraneous information does not.

With this overview in mind, the steps for computing local control signals when the attention target is known, are now defined mathematically. For reference, Table 4.1 provides a summary of the parameters used in the ARC.

The procedure for performing selective routing in the ARC begins when pulvinar projects to cortical control neurons in the top level of the hierarchy (PIT), a signal that coarsely encodes the size and location of the attentional target (A_{len} and A_{pos}). The first component of this signal A_{len} , specifies the size of the target in terms of the number of columns in V1 that it covers. The second component, A_{pos} specifies the position of the centre of the target, also in terms of V1 columns.

As mentioned above, the first computation performed by control neurons seeks to determine the size of the target’s representation in V4. Assuming an object-centred reference frame in PIT, where all columns represent the target, these control neurons simply determine, based on the global control signal A_{len} , whether the target’s representation in V4 will be larger than the number of columns in PIT. If the representation is larger, as will typically be the case, then the representation will need to be downsampled as it is projected from V4 to PIT. In this case, control neurons determine the downsampling factor by computing the number of V4 columns between which signals are selectively processed.

Assuming an object-centred reference frame in PIT where all columns represent the target, the first type of cortical control neuron receives A_{len} and uses it to compute the number of columns between values of μ for adjacent PIT columns:

$$sf = \frac{\min(\max(A_{len}, size_{PIT}), size_{V4}) - 1}{size_{PIT} - 1} \quad (4.1)$$

where $size_{PIT}$ and $size_{V4}$ are the number of columns in PIT and V4 respectively. As both of these values are fixed for a given network architecture, they are not explicitly signalled or represented, but are stored in the weights of the control neurons.

The second step in computing the local control signals is to determine the position of the target's representation in the previous level. This is performed using the second component of the global control signal from pulvinar, which indicates the position of the centre of the target, A_{pos} . As mentioned above, in performing this calculation, the model seeks to have the target's representation be centred in each level if possible. The position of the target's representation in V4 (θ_{V4}) is given by:

$$\theta_{V4} = \begin{cases} 0 & \text{if } A_{pos} \leq m_{PIT} \\ A_{pos} - (\text{sign}(A_{pos}) \times m_{PIT}) & \text{otherwise} \end{cases} \quad (4.2)$$

$$m_{PIT} = (a_{PIT} - 1)/2 + m_{V4} \quad (4.3)$$

where $\text{sign}(A_{pos}) = 1$ if the target is to the right of centre ($A_{pos} \geq 0$) and -1 otherwise, and m_{PIT} represents the maximum shift that can occur for columns in PIT. Stated simply, if the centre of the attentional target in V1 (A_{pos}) can be routed such that its representation is centred in V4, then route it to this position. If however, A_{pos} is at a greater eccentricity than can be routed to the centre of the level (due to the size of the receptive fields of neurons in that level), then route it as close to the centre of the level as possible. For a particular network architecture, the maximum possible shift m is fixed for control neurons in each level, and is computed by summing $\frac{1}{2}$ of its receptive field size ($(a - 1)/2$) and the maximum shift of the level below.

Having computed the position of the target's representation in V4 (θ_{V4}) and assuming an object-centred reference frame in PIT (θ_{PIT}), the number of columns by which the centre of the target's representation is shifted between V4 and PIT is given by:

$$s_l = \theta_{V4} - \theta_{PIT}. \quad (4.4)$$

Both θ and the global control signals A_{len} and A_{pos} are fed back to control neurons in V4 from PIT. These signals are provided to all V4 columns within the receptive fields of the PIT columns. However, depending on the values of the control signals, only some of the V4 columns will switch from their default routing state to one in which they are

selectively routing information from the target. To determine whether a given V4 column will be involved in encoding information from the target, at the point where the feedback signals arrive in V4, the following is computed:

$$state(x_j) = \begin{cases} selective & \text{if } A_{len} > size_{V4}, \\ selective & \text{if } \theta_{l-1} - \frac{A_{len}-1}{2} > x_j > \theta_{l-1} + \frac{A_{len}-1}{2}, \\ default & \text{otherwise,} \end{cases} \quad (4.5)$$

where x_j is the j^{th} input column in the previous level, *selective* indicates that the column will switch from its default routing state (*default*) to one in which information from the target is being selectively processed. In V4 columns not selectively routing attended information, the feedback signals are gated before they can influence the activity of the control neurons. Further, V4 columns remaining in their default routing state do not relay the feedback signals to the V2 columns within their receptive field, and thus, those V2 columns will also remain in their default routing state.

Before proceeding further, Figure 4.2 illustrates the role of sf and θ in the local control signal calculations. For V2 and V4, an equal number of columns encode the attentional target ($sf = 1$) since the size of the target (A_{len}) is less than the number of columns in those levels. As PIT has the fewest columns, the number of columns in V4 representing the target will typically exceed the number of columns in PIT. In Figure 4.2, the representation of the target spans seven columns in V4, while PIT contains just three columns ($size_{PIT} = 3$). In order to route the target's representation from V4 to PIT, the PIT columns sample every third column in V4 ($sf = \frac{7-1}{3-1} = 3$).

Figure 4.2 also illustrates that at each level, the target's representation is shifted as close to the centre of that level as possible. In V2, the centre of the target's representation $\theta_{V2} = 4$, due to the small receptive fields of neurons in lower levels. At V4 however, the larger receptive fields allow the target to be routed to the centre of the level ($\theta_{V4} = 0$).

Pyramidal cells with nonlinear dendrites have been reported in neocortex and hippocampus [117, 122, 160, 182], and based on the morphological and electrochemical similarities between those cells and pyramidal cells in visual cortex, it is proposed that such neurons may also be found in visual cortex (see Section 2.4 for further details). The functional advantage of the dendritic nonlinearities is that they endow the cell with the ability to compute two nonlinear functions of their inputs, similar to a two-layer artificial neural network [63, 181]. Presently, it is shown how this additional nonlinearity can be used to perform efficient selective routing.

The control neurons involved in computing Equations 4.1 and 4.4 project sf_{PIT} and s_{PIT} to the nonlinear terminal dendrites of the visually responsive pyramidal cells in column i in PIT. In the dendrites, these signals are used to determine the location within their receptive field to which attention should be directed:

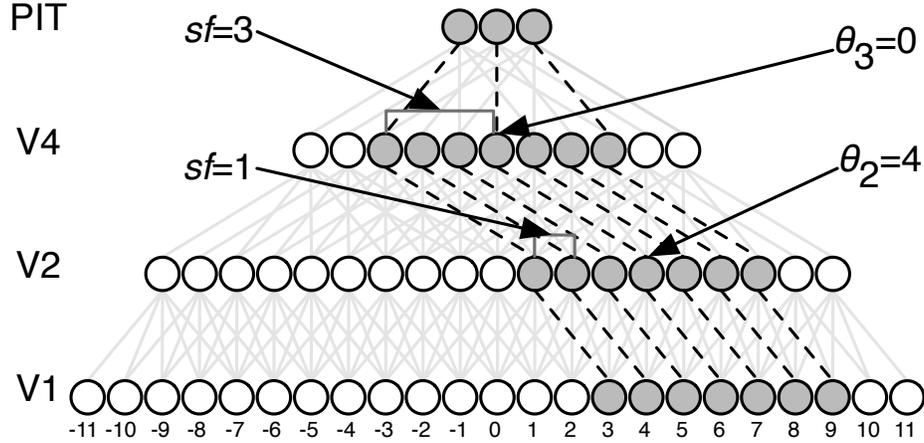


Figure 4.2: Example ARC network showing the position of the target’s representation (θ_l) in each level when attention is directed to the stimulus in V1 shown with gray circles ($A_{pos}=6$ and $A_{len}=7$). Dashed lines indicate the attentional focus of each column, μ , as determined by the local control signals. The relative shift of the target between V2 and V4 is $s_{V4}=-4$, and between V1 and V2, $s_{V2}=-2$.

$$\mu_i = sf \times i + s_{PIT}. \quad (4.6)$$

This value μ_i specifies the centre of a Gaussian-shaped routing function over the receptive field of column i that serves to modulate the gain of visual signals from input columns throughout the receptive field:

$$f(\mu_i, x_j) = e^{\frac{-(\mu_i - x_j)^2}{2\sigma_{att}^2}}, \quad (4.7)$$

where x_j is a constant representing the spatial location of the j^{th} input column in V4. Similar to gain fields models [204], this yields a Gaussian centred on μ_i with width σ_{att} . Figure 4.2 shows that in PIT, two of the columns are selectively processing information from the edge of the target’s representation in V4. In specifying the control signals, the ARC ensures that at least one column is selectively processing information from the edge of the target’s representation at each level, as in biological and computational vision systems, edges are known to play an important role in object recognition [94, 104, 212].

Figure 4.3 illustrates the routing function (Equation 4.7) in a single column $i=0$ in PIT receiving inpts from three columns $x_j=[-1, 0, 1]$. The control signals specify that column i should selectively process information from $x_j = 0$ ($\mu = 0$), shaded in red. Input signals from each column are multiplied by the gain term computed by the routing function for

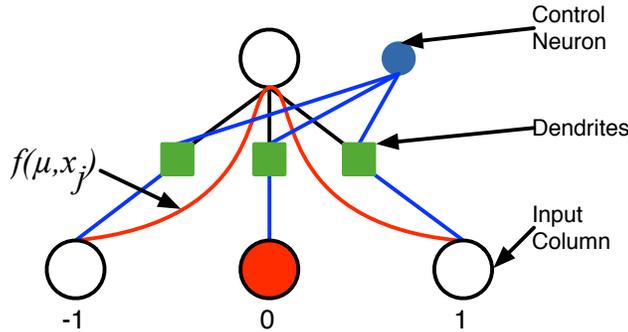


Figure 4.3: Illustration of the routing function in the ARC. The control neuron is sending μ , and the centre of the Gaussian routing function is focused on the input column at position $x_j = 0$ as $\mu = 0$. Visual signals encoded by each of the three input columns ($x_j = [-1, 0, 1]$) are multiplied by the gain term computed by the routing function for each input column's position.

each column, with signals from $x_j = [-1, 1]$ being attenuated, and the signal from $x_j = 0$ being multiplied by 1.

If the number of columns representing the attentional target in a given level exceeds the total number of columns in the next higher level, then sf will be greater than 1. This results in adjacent columns in the upper level having values of μ that skip one or more columns in the previous level that are encoding information from the target. Using the example shown in Figure 4.2, columns in PIT have values of $\mu = [-3, 0, 3]$. If a narrow routing function were used such that all columns except for those at μ_i were significantly attenuated, information from V4 columns $[-2, -1, 1, 2]$ would not be routed to PIT, resulting in this information being lost. In such cases, it is desirable that all information within the target's representation still be routed through, thereby requiring each column in the higher level to pool the visual signals from several columns in the previous level. This is accomplished by adjusting the width of the attention function, σ_{att} in Equation 4.7 by dividing sf by the full width at half maximum:

$$\sigma_{att} = sf/2.35. \quad (4.8)$$

With larger values of sf , this will introduce some blurring into the target's representation in higher levels, although doing so prevents the target's representation in PIT from containing sharp discontinuities between adjacent PIT columns.

Figure 4.4 illustrates the need for changing the width of the routing function depending on the size of the attentional target. Thick blue lines above each level represent the decoded

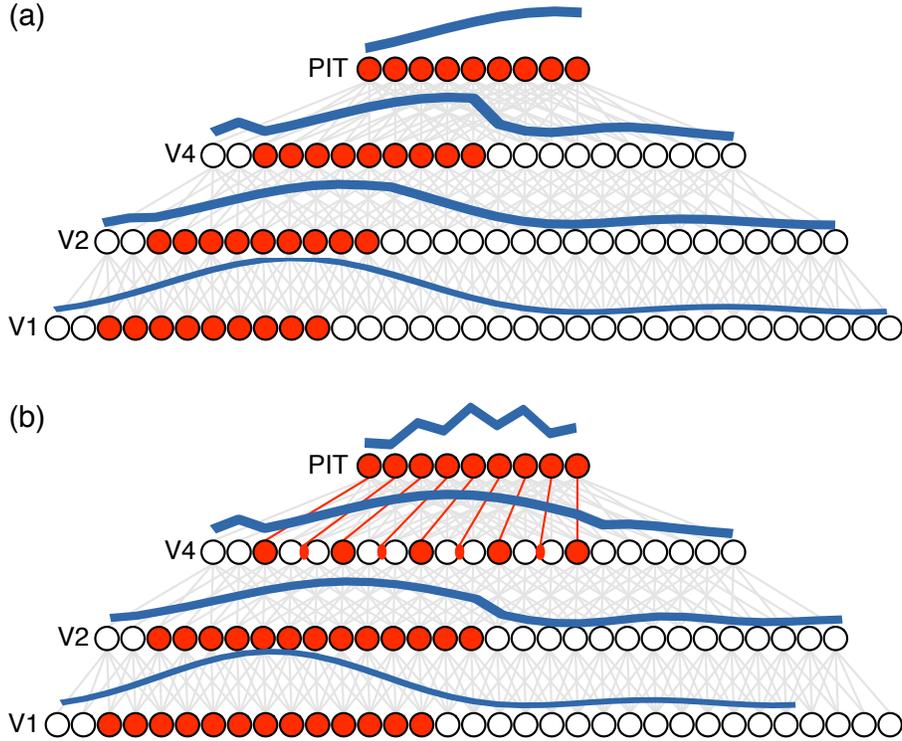


Figure 4.4: Effect of using a routing function with a fixed width (σ_{att}). (a) When σ_{att} is fixed at the appropriate value for all levels performing a shift of the FOA, selective routing proceeds as expected. (b) When the same routing function width is used for routing a larger target that requires both shifting and scaling, some columns in PIT have values of μ that land between two columns and thus the adjacent columns are attenuated, resulting in distortion of the target's representation.

visual signal at each level. In Figure 4.4a, $A_{len} = 9$ and at each level $sf = 1$. The width of the routing function is computed as $sf/2.35$, and information from the attentional target is routed to PIT as expected. In Figure 4.4b, the size of the attentional target is increased to $A_{len} = 13$, but σ_{att} is kept at $1/2.35$. By not modifying σ_{att} to accommodate the selective routing of the larger target, some of the PIT columns have values for μ that fall between two V4 columns, shown as small red ellipses in V4. The result is that the Gaussian routing function produces a value of $f(\mu_i, x_j) = 1$ at a position in between the two columns which, as a result of the narrow routing function, correspondingly yields a value of 0.36 for the columns between which μ falls. When the routing function's gain term is multiplied by the visual signals in the adjacent columns, their encoded values are significantly attenuated, producing a distortion of the visual signal in some PIT columns.

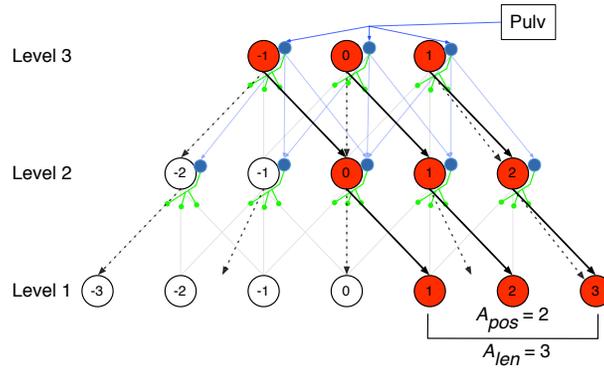


Figure 4.5: Routing example in a three level network. In the default routing state the full visual field is routed to an object-centred reference frame in Level 3. Dashed lines indicate the centre of the routing function for each column in their default routing state. When the target is known, pulvinar projects the global control signals A_{len} and A_{pos} to control neurons in the top level (small blue circles), where the global control signals are used to compute local control signals that specify the centre of the routing function (solid black lines). The global control signals are fed back to level 2, where local control signals are computed using the same mechanisms.

If the size of the attentional target were further increased such that it exceeds the number of columns in V4 or V2, this type of distortion would be similarly added at those levels as well and compounded at higher levels. As a result, it is necessary for σ to be modified for the selective routing of the attentional target as shown in Equation 4.8.

The result of Equation 4.7 is projected along with the feedforward visual signal from input column x_j (\hat{x}_j) to the soma, where the two values are multiplied using the second dendritic nonlinearity. The gated signal is then pooled along with gated signals from other neurons in that column to serve as feedforward input to the next higher cortical level. In total, the pyramidal cells with nonlinear dendrites receive three input signals, s , sf and \hat{x}_j , which interact according to Equations 4.6 and 4.7. The consequence is that the signals encoded by V4 columns at the centre of the Gaussian (i.e. $x_j = \mu_i$) are multiplied by 1 and thus contribute most to the activity of x_i in PIT, whereas columns that are more distant from μ_i are attenuated and contribute less.

Figure 4.5 shows an example of how this routing proceeds in a small three-level network. The network parameters, which are fixed for the network, are summarized in Table 4.2. In the default routing state, the entire visual field is routed to the top level, with the centre of the routing function for each column indicated by dashed lines. Once the target is known, pulvinar projects the global control signals indicating the size and position of the target

Table 4.2: Routing example - network parameter values for the network shown in Figure 4.5.

Parameter	Level 1	Level 2	Level 3
<i>size</i>	7	5	3
<i>a</i>	-	3	3
<i>m</i>	-	1	2

Table 4.3: Routing example - computed parameter values for the network shown in Figure 4.5.

		Level 2	Level 3
μ	default	[-3, -1.5, 0, 1.5, 3]	[-2, 0, 2]
μ	selective	[-3, -1.5, 1, 2, 3]	[0, 1, 2]
<i>sf</i>	default	1.5	2
<i>sf</i>	selective	1	1
<i>s</i>	default	0	0
<i>s</i>	selective	-1	-1
θ		1	0

($A_{len} = 3$ and $A_{pos} = 2$) to control neurons in the top level. The control neurons use the global control signals to compute local control signals for performing selective routing. The centre of the routing function computed by each column is indicated by solid black lines. The result of the routing function is then multiplied by the visual signals arriving in each column

The global control signals are then fed back to all columns in level 2, where each column determines if it will switch from its default routing state to selectively process the attended stimulus (Equation 4.5). For this target, level 2 columns $x_j = [-2, -1]$ will remain in their default state, and the other columns in that level switch to perform selective routing. The parameter values in the default routing state, as well as the local control signals computed for selective routing are summarized in Table 4.3. For columns having switched to a selective routing state, the values of μ are shown in bold, with non-bold values indicating that the corresponding columns remain in their default state.

As was discussed earlier, attention has been shown to affect the activity of neurons in higher cortical levels earlier than in lower levels [26, 32, 86, 138, 196], and the ARC proposes that this effect results from attentional control signals being fed back through the hierarchy, where they affect neuronal activity at each area along the way. As the control signals are computed, applied and fed back at each level, then for some time, neurons in PIT will be receiving a partially routed image. As time proceeds, the lower levels will

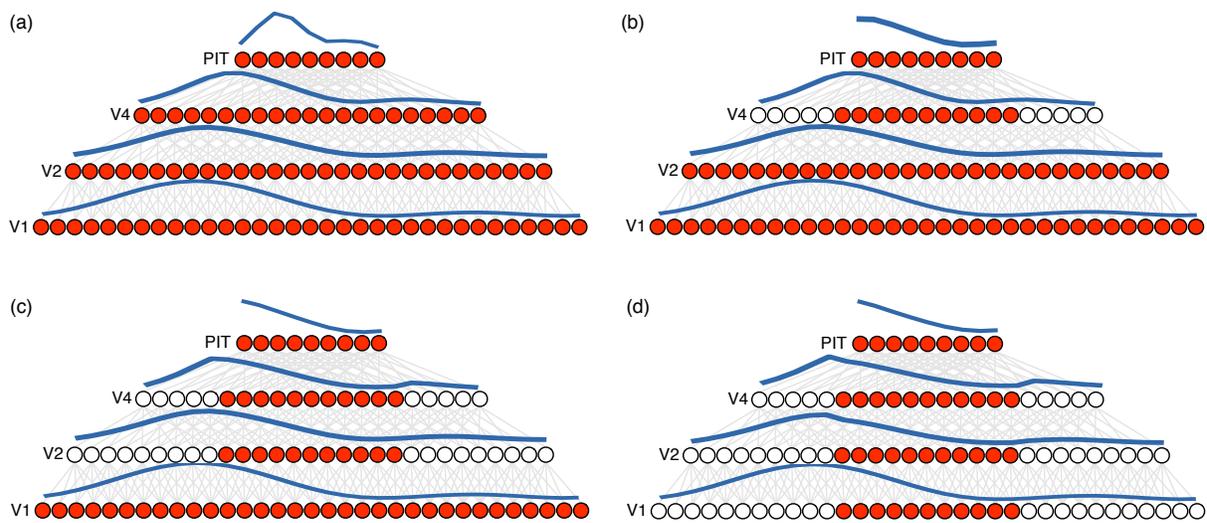


Figure 4.6: Evolution of visual signals in the ventral stream as attentional control signals affect routing in each area in a top down manner. (a) When the attentional target is unknown, the system operates in its default routing state and the entire visual field is routed to PIT. (b) The global control signal from pulvinar arrives in PIT, indicating that the attentional target covers the centre 11 columns in V1. At this point, neurons in PIT selectively route information from the target's representation in V4, while V4 and V2 remain in the default state. (c) The global control signal is fed back to V4, where columns encoding the target switch to a selective routing state, while columns not encoding information from the attentional target remain in their default state. (d) The global control signal is received in V2, which now selectively routes information from the target.

receive the feedback control signals and switch from their default routing state to one in which the target is selectively processed.

Figure 4.6 shows the evolution of the visual signal's representation in each level. At first (Figure 4.6a), all four levels are operating in their default state. Following the arrival of the global control signals in PIT, local control signals are computed in PIT which allows them to selectively route information from the V4 columns that will encode information from the target. This process continues through each lower level until the entire network is selectively processing information from the attentional target.

Examining the representation in PIT at each step in this example, it can be seen that the decoded visual signal does not significantly change once local control signals have been applied in PIT. This suggests that, depending on the amount of spatial detail required, successful recognition of the target may be performed without needing to modulate the routing performed below V4. This observation may explain the lack of attentional effects seen in lower areas when simple tasks are used [118, 131, 147].

4.3 Dendritic Nonlinearities

Pyramidal neurons are the most common type of neuron in cortex and constitute 70-90% of the neurons in cerebral cortex [56, 144]. The classical view of neurons used in most computational models is that neurons compute a single nonlinearity of the weighted sum of their dendritic inputs. For years, theorists have proposed the existence of neurons with nonlinear dendritic processing capabilities [106, 139, 141], and several recent studies have reported such neurons in the deep layers of rat neocortex [182], sensorimotor cortex [160] and hippocampus [117, 122]. When combined with the axosomatic nonlinearity, such neurons are able to compute two nonlinear functions of their dendritic inputs, thereby significantly increasing their computational power. The ARC predicts that pyramidal cells with dendritic nonlinearities will be found in layer-IV of visual cortex, and the experimental evidence that supports this prediction is presented by comparing layer-IV cortical pyramidal cells with those shown to have dendritic nonlinearities in area CA1 of hippocampus.

Pyramidal cells in both CA1 and visual cortex typically contain a single axon, several short basal dendrites, and one or two apical dendrites that branch several times to form the apical tuft [223]. Pyramidal cells in both areas also contain thousands of dendritic spines, and these membranous extensions of the dendritic surface represent the primary postsynaptic site for excitatory glutamatergic synapses [13]. Further, pyramidal neurons in both areas have a relatively high and constant density of Na^+ channels throughout the dendritic tree, increasing density of I_h channels at greater distances from the soma, a narrow window for temporal summation, backpropagating action potentials, dendritic Na^+ , K_A and Ca^{2+} channels, and somatic Na^+ channels [142, 223]. Neurons in both areas

are also capable of acting as coincidence detectors, which refers to the ability to discern the simultaneous occurrence of synaptic inputs at different locations on the same cell [223, 231]. Finally, both types of pyramidal cells can propagate dendritic spikes toward the soma [223].

In order for the neurons to exhibit dendritic nonlinearities, several conditions must be satisfied. For superlinear summation in the dendrites to occur, synaptic inputs to the terminal dendritic branches must arrive within 40 milliseconds and 40 μm of each other, and occur within the same dendritic branch [182]. A simplifying assumption of the neuron model used in this thesis is that synaptic inputs to these dendrites are positioned within 40 μm .

In this neuron model, the first stage of processing reflects the computation occurring in the thin terminal dendrites which receive approximately 85% of total input to the cell [117, 181, 182, 224]. Each thin terminal dendritic branch or dendritic subunit, computes a nonlinear function of its inputs. The second processing stage corresponds to that occurring at the soma, where the dendritic activities are integrated and a spike is generated using the standard LIF axosomatic nonlinearity when the firing threshold has been reached.

4.4 Coarse to Fine Processing

The ARC proposes that in conditions where the attentional target is not known, the visual system routes information according to a default routing state in which the entire visual field is processed. In this state, the visual information encoded by columns in PIT is a coarse representation of the entire visual scene. The representation in PIT lacks fine spatial detail primarily because significantly fewer neurons at each level are representing the visual information. To accommodate the reduced capacity at each level, each column pools the visual signals from multiple input columns within their receptive field. This pooling yields an increasingly blurred representation at higher levels, resulting in a loss of high spatial frequency information, while the low frequency information is better preserved. Once the attentional target is known however, routing may then proceed to allow finer detailed information of the target to be routed through the hierarchy.

The proposal that perceptual processing occurs at both coarse and fine spatial scales, with a first pass of the scene being at a coarse spatial resolution, and subsequent assessment at finer resolutions dates back several decades [158, 159, 172], and is included in several theoretical models [27, 94, 123]. This is closely related to the global advantage effect described in psychology literature, which suggests that the global structure and the spatial relationships between components are processed before the analysis of local details [166]. For a detailed review of psychophysical studies showing this effect, see [99].

Several studies have demonstrated that with rapid presentation of complex scenes, subjects tend to describe the scene based on its low frequency components [164, 165].

Schyns and Oliva [210] investigated the role of spatial frequency in perceptual processing by presenting subjects with “hybrid” images constructed by overlaying the low frequency components extracted from one image with the high frequency components extracted from a second image. Subjects were primed with a brief presentation of the hybrid image for 30 ms, after which subjects had to match the image with a normal picture. With short presentation times, subjects seemed to perform matching based on the coarse, low spatial frequency components, although with longer presentations of 150 ms, this effect was reversed, with subjects matching based on the high frequency components, with similar effects reported by Sripathi and Olson [226]. Consistent with these findings, a study by Parker et al. [175] investigated the effect of presentation sequence on discrimination performance using high-pass and low-pass spatially filtered versions of each image. They found that when the filtered images were presented in a low-to-high order, subjects were significantly better able to discriminate the stimuli than when the stimuli were presented in a high-to-low order.

A recent fMRI study by Goffaux and colleagues [72] provides further evidence of coarse-to-fine processing. In that study, images of faces that were filtered to preserve low, medium or high spatial frequency components were flashed for 75, 150 or 300 ms. In the fusiform face area, a high level cortical region often characterized by the presence of face selective neurons, there was a robust response to low frequency stimuli with short exposures that decayed rapidly with longer exposure durations. Conversely, the responses to face images containing only high spatial frequency components became more robust over time with longer exposures, suggesting that the processing of coarse, low spatial frequency information precedes that of detailed high spatial frequency information.

At the neuronal level, a correlate of the global advantage effect has been reported in macaque area IT. Sripathi and Olson [226] presented shape stimuli that were formed with either solid lines, or as hierarchical shapes composed of smaller shapes. The hierarchical stimuli were either congruent (e.g. a square composed of several small squares) or incongruent (e.g. a square composed of small circles). They found that the responses of IT neurons form a coarse representation of the shape approximately 30 ms before their activity reflects the local shape information, consistent with the coarse-to-fine model. That study also reports evidence of scale-invariance, such that neurons preferentially responding to a particular larger solid shape also tend to prefer the corresponding small shape when presented in isolation.

Taken together, these results indicate that in some conditions, visual processing first performs a rapid analysis of the visual scene based on the coarse properties, with subsequent processing being performed at a finer resolution. Both physiologically and computationally, there is good reason for such an arrangement. In the magnocellular layers of LGN, neurons are sensitive to stimuli with high temporal frequency and low spatial frequency information, and have heavily myelinated axonal projections that provide a rapid conduc-

tance of this information to higher cortical areas. This rapid processing is seen in neurons in dorsal stream areas showing activation significantly earlier than neurons receiving signals from the parvocellular layers of LGN. As noted by Bullier [27], this is consistent with the dorsal stream’s involvement in visuomotor processing, where rapid visual processing is an important factor in guiding fast motor movements.

From a computational perspective, the rapid first pass of the visual scene allows the coarse estimate of the information to guide the expectation of its contents [27, 94, 123] or to activate schemas in memory [210]. Depending on the task, the brief analysis of the scene can be used to guide subsequent processing of salient or task-relevant areas or objects in more detail.

4.5 Default Routing

The ARC proposes that neurons that are not involved in encoding information from the attentional target remain in their default routing state. This proposal suggests that such neurons can be identified by three characteristics. First, their receptive field will generally be larger when they do not contain the attentional target. This effect was demonstrated in a study by Womelsdorf et al. [272] and is addressed in greater detail in the simulations of that experiment in Chapter 6.

The second characteristic of neurons operating in their default routing state proposed by the ARC is that neurons with peripheral receptive fields will not necessarily be most sensitive to stimuli at the centre of their receptive field when it does not contain an attentional target. As seen in the example of attentional routing in Figure 4.5, in the default routing state, the centre of the routing function is not necessarily centred in the receptive field. Rather, in the default routing state, neurons in the periphery have values of μ that are biased away from the fovea, which allows the entire visual field to be routed. This arrangement may explain the receptive field shift asymmetry reported by Connor et al. [36], which found a slight, but significant deviation in the receptive field shift toward the fovea.

This prediction could be tested by having the animal attend to a non-receptive field target while small probe stimuli are displayed in the receptive fields of neurons in V2, V4, and IT. The responses to probe stimuli across the receptive field would then be fit with a Gaussian, similar to the method used by Womelsdorf et al. [272]. The animal would then covertly attend to a non-preferred stimulus placed within the receptive field while probe stimuli are displayed throughout the receptive field.

By systematically comparing the receptive field fits in the attend-out condition with the attend-in condition as a function of receptive field eccentricity, the model’s prediction can be tested. Specifically, it is predicted that as the receptive field eccentricity increases, in

the attend-out case, the peak of the Gaussian will be displaced further from the receptive field centre, in the direction away from the fovea than for neurons having receptive fields in the foveal and parafoveal regions. The model also predicts that when attention is directed to the receptive field stimulus, the magnitude of the shift will be larger when the target is positioned on the side of the receptive field that is closer to the fovea than when the target is on the opposite side of the receptive field, away from the fovea.

Additionally, the ARC predicts that in higher areas, there will be a greater number of cells showing this effect, and that the shift magnitude will also be greater. Following the proposal of the global signals being fed back through the visual hierarchy, the model predicts that the differential timing of attentional modulation occurring earlier in higher cortical areas results from neurons switching from their default routing state to a selective state. This is closely related to the findings of Buffalo et al. [26], but allows for the specific hypotheses of the underlying functional mechanisms producing this effect to be tested.

The third characteristic of neurons operating in their default state proposed by the ARC is that they will encode a coarse and less detailed representation of the visual information in their receptive field. This proposal is supported by a study by Alvarez and Oliva [4] of human subjects performing a multiple object tracking task. The display contained eight moving objects, and subjects were to attentively track a subset of four items that were indicated as targets by flashing on and off at the start of the trial. At a random time in each trial, all items disappeared briefly, after which either one item disappeared (a randomly selected target or distractor), or four items disappeared (all of the targets or all of the distractors). Subjects were then asked to estimate the location of the missing item when one item disappeared, or to estimate the centroid of the group of missing items. When a single item disappeared, the subjects' estimate of the missing item's position was significantly worse when the missing item was a distractor than when it was a target. However, in both single item conditions, this produced a significantly larger error than on trials in which they estimated the centroid of the four missing items, when either all four distractors disappeared or all four targets. The authors suggest that subjects can make accurate judgements about the positions of distractors by pooling information from the individual distractors, even though detailed information about the individual distractors is not represented accurately. Further, they conclude that information outside of the focus of attention is represented at a coarse level that carries an accurate statistical summary of that information, but lacks precise details about individual items.

The default routing state may also be sufficient for simple tasks that do not require fine spatial detail about the stimulus structure. Measuring receptive fields of IT neurons, Rolls et al. [197] found that when monkeys were presented with a stimulus on a blank background, the receptive fields were large (39°) and they were able to rapidly locate the target. However, when the same stimuli were presented with a complex background from a natural scene, the size of the receptive fields shrunk considerably to 11° , and their reaction

times significantly increased. In the ARC, the former case is proposed to correspond to neurons operating in their default routing state, as the simple task does not require fine discrimination. In the latter condition, in which the receptive fields contain a significant amount of irrelevant information, corresponds to neurons operating in a selective routing state. It is presumed that when an animal is under anaesthesia, the attentional system does not guide the selective processing of information and has little influence on stimulus processing. In such conditions, a default routing state would be used, wherein the entire visual field is processed, thereby resulting in the width of the routing function, and thus the receptive field size being larger. This presumption is consistent with reports of IT neurons having significantly larger receptive fields under anaesthesia [42, 77]. Further, in non-anaesthetized animals, IT receptive fields tend to be smaller and restricted to the immediate area the animal is attending [147], indicating that effective inputs are dynamically modulated with changes in attention.

4.6 Object-centred Reference Frame

Being able to recognize an object under varying lighting conditions and viewpoints is a central problem in computer vision, yet is one that is well handled by human vision. We are able to identify a particular object whether it is directly in front of our face or several metres away, despite these two viewpoints forming markedly different images on the retina. Similarly, when overtly attending to an object by looking directly at it, or when covertly attending to the object while directing the eyes elsewhere, a very different image is presented to the retina, which stimulates different retinal ganglion cells, and in turn those in LGN and striate cortex. However, in both cases, we are able to recognize the object despite the change in its size (scale-invariance) and location in the visual field (position-invariance).

The ARC seeks to address how this invariant recognition may be performed, by proposing that attended objects are routed to a common object-centred reference frame. In this thesis, an object-centred reference frame refers a representation of an object that is formed in higher level cortex, invariant to its retinal size and position, and in a manner that is relatively independent of the visual information surrounding that object.

The notion of object representations being transformed to an object-centred coordinate system or reference frame has been around for several decades [87, 124, 168, 169, 170, 173] and this assumption is supported by physiological and psychophysical findings. From a computational perspective, remapping the retinal image of an attended stimulus to a reference frame centred on that object greatly reduces the complexity of recognition by largely discarding visual information that is unrelated to the object being identified [235].

The shifter circuit model [168] proposes that the object-centred reference frame is of

a fixed size, which they estimate to contain 25-30 sampling elements in each spatial dimension. One line of evidence supporting this hypothesis comes from a study by Sperling et al. [221] which investigated the influence of image pixel density on the intelligibility of American Sign Language symbols. They found that when compared to control images with a size of 96×64 pixels, images containing 24×16 pixels were reported to be 86% as intelligible as the control images. Although the study did not examine this effect on images below this size, the proposal of a fixed size reference frame predicts that the performance should deteriorate rapidly once the image resolution is reduced below that of the object-centred reference frame. Olshausen and colleagues [168, 257] suggest that the intelligibility of the images with 24 sample points in one spatial dimension is consistent with their estimate of the size of the reference frame.

In a study by Toet et al. [243], subjects were presented with a display containing three blobs with Gaussian spatial contrast profiles. The blobs were vertically aligned and had an equal spacing between them. The task was a two-alternative forced choice design in which subjects had to report whether there was a left or right displacement of the middle blob relative to the outer blobs. By systematically varying the spatial extent of the blurring and separation between stimuli, they were able to measure subjects' displacement discrimination thresholds. The study found that the discrimination thresholds increased linearly with increasing blur, while for a given separation, the threshold was reduced as the amount of blurring, and thus stimulus size, was reduced. Once the blob sizes were reduced beyond a certain point however, performance plateaued when the separation of the blobs was a constant multiple (~ 25) of the amount of blur. Van Essen et al. [257] suggest that this reflects a fixed-size reference frame that is consistent with their estimate of 25-30 sampling elements in each spatial dimension.

For an object-centred reference frame to be formed at some level in the visual hierarchy, neurons in lower areas providing inputs would also need to contribute some degree of scale and position invariance or coordinate transformation. Such effects have been reported in studies showing receptive field shifts [35, 36, 272], as well as coordinate transformations seen in parietal cortex [5, 6, 217] (see the discussion of gain fields in Section 3.6).

From monkey physiology studies, supporting evidence of an object-centred reference frame comes from observations of increasing scale- and position-invariance in higher ventral stream areas. Recording from superior temporal sulcus (STS), which receives inputs from IT and other areas, Perrett et al. [178] found that presenting images of faces ranging in size from 20 cm to more than 2 metres failed to show any effects of size in the neural activity. Similarly, Rolls and Baylis [198] found that the majority of STS neurons gave relatively invariant responses to faces, responding at rates greater than half of their maximal rate to image sizes changing by a factor of ~ 12 . A study by Ito et al. [93] recorded from neurons in IT to assess the degree to which cells were invariant to size and position, finding that 43% of the recorded cells were maintained their selectivity over a 2-fold change in stimulus

size, while 21% of the cells were responsive to a 4-fold change in the size of the stimulus.

Desimone et al. [42] recorded from IT cortex while images of complex objects, shapes, patterns and faces were presented. Many neurons responded strongly to the majority of presented stimuli, without clear preference for size and other stimulus attributes. Other cells that were more selective for the shape, color or texture of the stimulus maintained their selectivity across positions within their large receptive fields. Further, a small population of neurons were found that were selective for faces. These cells responded selectively over a wide range of image sizes and orientations in the receptive field, but did not respond invariantly to different configurations of facial features, as is found in higher levels such as STS and FFA [72, 251].

A study by Pasupathy and Connor [177] demonstrated that neurons in V4 responded selectively to certain shape characteristics at an object-centred location throughout their receptive field. This finding of cells that are selective for shape elements within a restricted subregion of a larger object suggests that they may be participating in the formation of an invariant representation of the object in higher cortical areas. This finding was confirmed in a subsequent study by Brincat and Connor [23] which found consistent results in PIT when more complex stimuli were used.

As was discussed in the previous section, Sripathi and Olson [226] found that neurons in IT respond strongly to a given shape whether it is constructed from solid lines, or is a hierarchical shape formed from several small shapes. Further, they report that neurons preferring a given large shape tend to also prefer the corresponding small version of that shape, indicating a tendency for shape selectivity to be scale-invariant.

The physiological studies presented above all support the proposal of an invariant object-centred representation in higher level cortex. The varying amounts of scale and position invariance reported in IT in different studies may be due to the influence of projections from other areas. IT receives projections from a multitude of areas including early ventral stream areas, FEF, STS, amygdala, hippocampus, and entorhinal cortex [59, 114, 202], which may affect the neural activity depending on the task requirements and the animal's cognitive state.

Results from several psychophysical studies have been interpreted as reflecting the mental representation of objects in object-centred coordinates, based on increasing reaction times for rotated objects [237, 238] and object-centred deficits in neglect patients [14, 133]. However, the mapping of human performance to physiological effects is challenging, as there are numerous processing stages following the intermediate cortical levels from which recordings were made that can play an important role in generating responses. Despite uncertainty of the relationship between physiological and behavioural effects, the accumulating evidence of object-centred representations provides sufficient support to justify the hypothesis.

At present, it is difficult to determine whether an object-centred reference frame would be formed in PIT, or in a higher cortical area beyond PIT. The current model proposes that PIT is a likely candidate, as in higher visual and association cortical areas, the large receptive fields cause the retinotopic organization of each area to quickly erode. In areas beyond PIT, there is still a columnar organization [62, 203, 202, 233], although it is arranged according to columnar groupings of neurons tuned to similar high-dimensional and complex stimuli.

Whether an object-centred reference frame is formed in PIT or in a higher area, the routing mechanisms proposed by the ARC still apply, and suggest that all columns in the level in which the object-centred reference frame is formed attempt to selectively process information from the target. Due to the specificity of their tuning, not all columns will be responsive to the target stimulus, which will result in the activation patterns in that area showing “hotspots” of activity [104], corresponding to strong responses from columns that are selective for stimuli that are similar to the target. However, the columns showing strong responses to the target will not necessarily be spatially adjacent on the cortical sheet. This issue is discussed further in Section 7.1.2.

4.7 Comparison to Existing Models

In Chapter 3, several previous models of attentional processing were discussed. Having introduced the ARC earlier in this chapter, this section compares the reviewed models with the ARC by highlighting their similarities and limitations.

The principles of gain fields for performing coordinate transformations is well supported by experimental work, and have been shown to provide a solid explanation of receptive field modulation with attention, both in previous models and in the ARC. However, existing gain field models by Salinas and Abbott [204] and Womelsdorf et al. [272] reviewed in Chapter 3 lack details of how such attentional effects fit into the larger context of selective attentional processing across multiple cortical areas. Specifically, neither model suggests how the control signals that define the gain fields are computed or how these effects are related to those in other areas, as they only address receptive field modulation between two adjacent cortical regions.

In contrast, the ARC details a mechanism for generating the local control signals that is physiologically and anatomically consistent, and specifies how the control signals are communicated and transformed between cortical areas, within an individual area, and between laminae of individual columns. By using a multilevel hierarchy, through which a global attention signal is fed back, the ARC is able to provide a qualitative explanation of the differential timing of attentional modulation between areas, with testable predictions of the mechanisms that produce a temporal delay in more peripheral areas. Also, as

both the Salinas and Abbott [204] and Womelsdorf et al. [272] models focus on receptive field modulation for neurons in a single column, they do not address how local spatial relationships are maintained between columns representing the target. The ARC proposes that topography is preserved by having control neurons in each cortical area compute control signals using a common mechanism on the same input signals (i.e. A_{len} and A_{pos}), and locally projecting these signals within the column and to columns within their receptive field in the previous level. Finally, neither model suggests how attention affects neurons that are not encoding information from the attentional target.

The shifter circuit [168, 169] provides a strong starting point for modelling attentional routing in cortex as it describes the relationship of attentional modulation across multiple cortical areas, the computations sufficient for generating the control signals, and provides a functional account of how selective routing may be performed. However, it has several limitations that remain to be addressed (see Section 3.3.2).

First, it is a purely mathematical model that lacks details of how the calculations and representations may be performed in spiking neurons, which preempts assessment of the biological plausibility of the model’s computations. By first presenting the ARC in a purely mathematical instantiation allows it to functionally demonstrate selective routing in a large-scale network, and be subjected to a similar level of analysis as the shifter circuit. By then implementing the ARC in spiking neurons, the plausibility of the control calculations can be explicitly tested. The model can also be used to simulate physiological experiments, which further demonstrate the biological plausibility of the proposed mechanisms.

The second limitation of the shifter circuit is that it has limited scalability. The shifter circuit proposes that specific control signals (similar to the local control signals used in the ARC) are computed in pulvinar, from where they are distributed to ventral stream areas. However, if the model is scaled to approximate the size of visual cortex, the neuronal resources required by the model for computing control signals exceeds that available in the pulvinar (see Section 3.3.2). Since the shifter circuit proposes that these control signals are computed by pulvinar and not simply relayed, thereby significantly increasing the computational requirements of pulvinar, this suggests that pulvinar may not be capable of performing routing with sufficient precision.

The ARC overcomes this second limitation by proposing that pulvinar encodes a coarse representation of the size and position of the attentional target, and projects this signal to cortical control neurons in PIT. By proposing that pulvinar represents a low dimensional control signal, rather than unique control signals for all levels as in the shifter circuit, the ARC significantly reduces the neural requirements of the nucleus. The generation of specific local control signals is then performed by cortical control neurons in each level, which relay the signal from pulvinar through successively more peripheral visual areas. Control neurons in all levels employ a common mechanism that operates on the same signal from pulvinar, without need for precisely targeted and wide reaching projections

from pulvinar.

The third limitation of the shifter circuit is that selective routing is performed by having the control signals dynamically change the synaptic connection weights, which is unlikely to occur at the time scale required for selective routing. The ARC overcomes this limitation by using an approach similar to gain field models, wherein the local control signals and visual signals interact multiplicatively in the dendrites.

The normalization model [189] is closely related to the biased competition model [188], and provides a compelling account of selective processing in cortex. Several aspects of the model are consistent with the ARC, although it also has several limitations. The first similarity is that both models propose that the many forms of attentional modulation result from a single common attentional mechanism, and are able to address such effects within a single model. Second, both models presume that attention affects each processing stage similarly, although in the normalization model, it is unclear how the attention field is propagated and transformed in each level. The third similarity is that both models do not define the processes or mechanisms for selecting the attentional target, but assume that it is provided to the ventral stream areas. Finally, both models utilize a multiplicative interaction between the feedforward visual signals and a control signal, although in the normalization model, the attention field is specified for all individual neurons, whereas the ARC uses a coarse, low-dimensional global control signal that is transformed in each cortical area to generate local control signals.

As with the models using gain fields [204], synchrony and oscillations [162, 163, 204], the normalization model only describes attentional modulation in a single cortical area, without addressing the relationship to other ventral stream areas. Similar to the shifter circuit, the model is a purely mathematical implementation, which limits its predictive power in neurobiology. The implementation also prevents the detailed study of physiological effects such as changes in responses at the single cell and population levels, and the distribution of these effects across multiple repetitions and in different simulated primate brains. Further, the plausibility of the computations being performed in actual neurons is unclear.

The normalization model also suggests that size of the attentional target will differently affect the attentional modulation for low and medium contrasts. Specifically, the model predicts that attending to small stimuli will produce a contrast gain effect, while attending to larger stimuli will produce a response gain effect. However, the results of several studies cast doubt on this hypothesis. First, Lee and Maunsell [112] did not observe a significant positive correlation between attentional modulation and stimulus size. Further, the Lee and Maunsell study used stimuli with an average size of 0.5° which produced a similar amount of modulation to that reported in another study in MT by Treue and Maunsell [247] which used stimuli covering only 0.3° . Finally, a separate study by Lee and Maunsell [111] increased the stimulated area of the MT neurons' receptive field by placing two stimuli that moved in the same direction within the receptive field, finding that the response

modulation with both stimuli was not significantly different than when only one receptive field stimulus was used. In contrast, the ARC is consistent with the results of these studies, which are discussed further in the simulations presented in Chapter 6.

Chapter 5

Attentional Routing Circuit - Computational Specification

To this point, a high-level overview of the mathematical basis for performing selective routing in the ARC has been presented. This chapter introduces the methods for extending the model’s principles to a more biologically detailed instantiation using spiking neurons in the Neural Engineering Framework (NEF). Details are given of the single neuron models, the organization of the columns and their relationship within the hierarchy, and of the specific laminar circuitry for generating and applying control signals.

5.1 Neural Engineering Framework

A core concept of the Neural Engineering Framework (NEF) [55] is that the activity of a neural population “represents” information about external stimuli or internal processing. The framework describes this information and its transformation mathematically, as representations of scalar values, vectors and functions. The NEF proposes that the neural representation of a stimulus or variable is defined by the combination of nonlinear encoding (e.g. neuron tuning curves and spiking) and weighted linear decoding over neural populations and time. The tuning curve determines the cell’s encoding of inputs and can be expressed as:

$$a_i(x) = G_i \left[\alpha_i \left\langle x_i \cdot \tilde{\phi}_i \right\rangle_n + J_i^{bias} \right] + \eta_i(t), \quad (5.1)$$

where α_i is gain, x_i is the input, $\tilde{\phi}_i$ is the preferred direction vector, J_i^{bias} is the bias term corresponding to background activity, $\langle \cdot \rangle$ is the inner product of the n dimensional vectors, G_i is the output nonlinearity that transforms somatic current to spiking activity, and η is

a noise term. In all of the simulations presented in this thesis, G_i represents the standard leaky integrate-and-fire (LIF) nonlinearity [103], although other neuron models may be used. The LIF neuron model is used as it provides a suitable trade-off between biological realism and computational efficiency. The particular choice of neuron model that is used (e.g. adapting-LIF, Hodgkin-Huxley, Morris-Lecar, θ -neuron, etc.), does not affect the principles of the NEF. Nor does it influence the neural coding in the model, as the NEF does not depend on how the spikes are generated, but only on the statistics of the spike generation [55, pp. 89].

The output activity $a_i(x)$ represents the neural response and the simulations presented in this thesis were conducted with spike trains. This value is greatest when the input x is aligned with the neuron's preferred direction vector, $\tilde{\phi}_i$. With spiking neurons, the activity of neuron i is defined as a sum of action potentials:

$$a_i(\mathbf{x}, t) = \sum_n \delta(t - t_{in}), \quad (5.2)$$

where δ is an impulse representing a spike, t represents time, and t_{in} is the time of the n^{th} spike produced by neuron i .

With LIF neurons, the firing rate of neuron i is computed as:

$$a_i(\mathbf{x}) = \frac{1}{\tau_{ref} - \tau_{RC} \ln(1 - \frac{J_i^{thresh}}{a_i \mathbf{x} + J_i^{bias}})}, \quad (5.3)$$

where $\tau_{ref} = 2\text{ms}$ is the cell's refractory period, $\tau_{RC} = 20\text{ms}$ is the membrane time constant, and J_i^{thresh} is the current threshold. This rate representation is used to compute optimal decoders in the NEF.

Decoding in the NEF is a linear operation, determined as a weighted sum of the neuronal activity:

$$\hat{x} = \sum_i a_i(\mathbf{x}) \phi_i, \quad (5.4)$$

where ϕ_i is the decoding vector and \hat{x} is the decoded estimate of x . The decoding vectors ϕ_i can be analytically derived. To derive ϕ_i , we minimize the error between the decoded estimate and the actual value of x as:

$$\begin{aligned} E &= \frac{1}{2} \int_{-1}^1 \left[\mathbf{x} - \sum_{i=1}^N a_i(\mathbf{x}) \phi_i \right]^2 dx \\ &= \frac{1}{2} \left\langle \left[\mathbf{x} - \sum_{i=1}^N a_i(\mathbf{x}) \phi_i \right]^2 \right\rangle_x, \end{aligned} \quad (5.5)$$

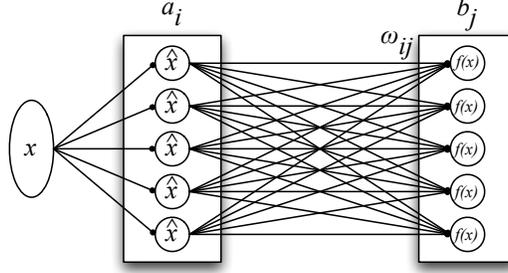


Figure 5.1: A communication channel in the NEF. Input signal x is projected to population a , which then relays this signal to population b . The connection weights between a_i and b_j , ω_{ij} , are analytically derived to approximate the communication function, $f(x) = x$.

where $\langle \cdot \rangle_x$ is the integral over x . Solving for the decoders ϕ_i gives:

$$\phi = \Gamma^{-1}\Upsilon \quad (5.6)$$

$$\Gamma_{ij} = \langle a_i(\mathbf{x})a_j(\mathbf{x}) \rangle_x \quad (5.7)$$

$$\Upsilon_j = \langle a_j(\mathbf{x})\mathbf{x} \rangle_x \quad (5.8)$$

Having specified the methods for representing a value in a neural population in the NEF through encoding and decoding, we now describe how such representations can be communicated and transformed between populations by determining their connection weights. Figure 5.1 shows a simple example of a communication channel between population A and B, where A receives input \mathbf{x} and simply communicates this value to B without transforming the information. The connection weights of a communication channel between presynaptic neuron i in population A and postsynaptic neuron j in population B are determined as the product of their encoding and decoding vectors, scaled by a gain term:

$$\omega_{ij} = \alpha_j \left\langle \tilde{\phi}_j \phi_i \right\rangle_n, \quad (5.9)$$

where $\tilde{\phi}_j$ is the encoding vector of neuron j specifying its preferred direction, ϕ_i is the decoding vector of neuron a_i and α_j is the gain. These weights can be computed by substituting the decoding equation (Equation 5.4) into the encoding equation for population B (Equation 5.1). Thus, the activity of neuron b_j is computed as:

$$b_j(\mathbf{x}) = G_j \left[\sum_{i=1}^N \omega_{ij} a_i(\mathbf{x}) + J_j^{bias} \right]. \quad (5.10)$$

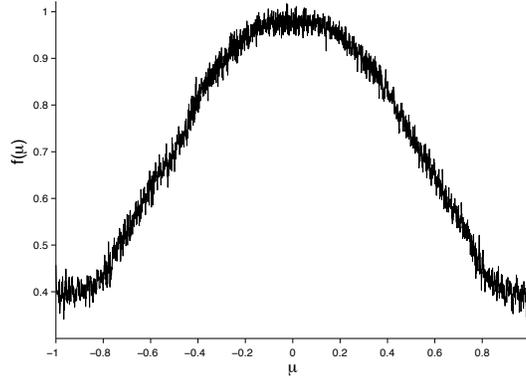


Figure 5.2: NEF ensemble computing the ARC’s routing function for an input column at position $x_j = 0$, with $\sigma_{att} = 0.75$. The control signal μ is varied continuously across the receptive field. As μ becomes increasingly distant from 0, the value computed by the routing function decreases, thereby increasingly attenuating the visual signal carried by neurons in column $x_j = 0$.

This describes how a simple communication channel may be implemented using the methods of the NEF, and by deriving different connection weights, the synaptic strengths between A and B can serve to compute arbitrary transformations of the values represented in A. Using the same methods as when deriving the decoders ϕ_j for a communication channel, the decoders $\phi^{f(x)}$ can be determined to approximate an arbitrary function $f(x)$. Substituting these decoders into Equation 5.9 provides the weights for estimating the representation of $f(x)$ in B.

Figure 5.2 shows an example for a population computing the routing function $f(\mu, x)$ used in the ARC (Equation 4.7). A control signal μ is provided as input to the population x_j where $j = 0$, and the synaptic weights approximate the function $f(\mu) = e^{\frac{-(\mu - x_j)^2}{2\sigma_a^2}}$ by solving for $\phi_j^{f(\mu)}$. In this example, the population is receiving inputs from a column at position $x_j = 0$, and μ is varied across the neuron’s receptive field which spans the range $[-1, 1]$. In the ARC, the result of this function for a given μ is used as a multiplicative gain term for gating the input signals. When $\mu = 0$, inputs from $x_j = 0$ are multiplied by a value of 1, and as μ moves away from 0, the gain term increasingly attenuates input signals from those parts of the receptive field.

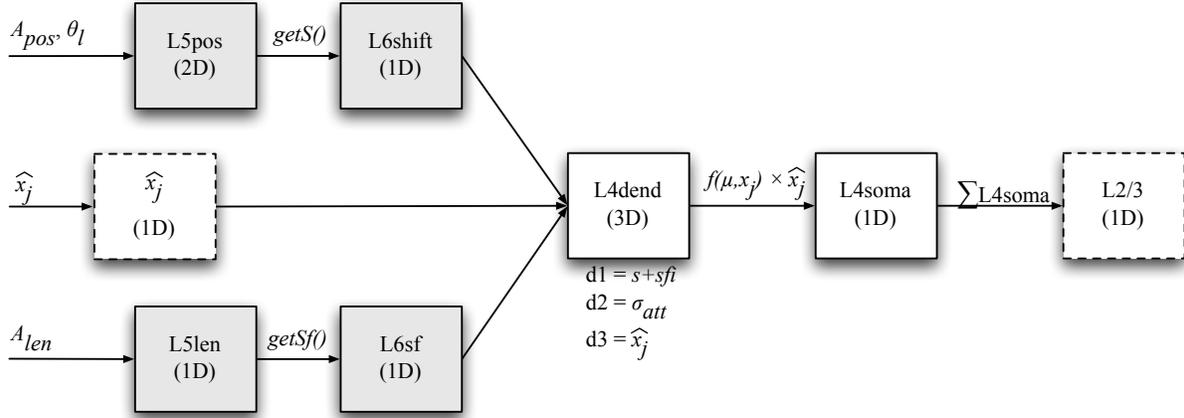


Figure 5.3: Visual depiction of neural circuitry for generating and applying local control signals in the ARC. Each box corresponds to a neural population: dashed boxes indicate populations encoding feedforward visual signals, shaded boxes are control neurons, and white solid line boxes together depict the layer-IV pyramidal cells with nonlinear dendrites. (nD) indicates the dimensionality of the encoding vectors, and d_n indicates the signals encoded by each dimension of the nonlinear dendrites.

5.2 Spiking Neuron Implementation of the Attentional Routing Circuit

To this point, the general architecture of the ARC has been defined and shown to be capable of performing selective routing of attended visual information. Further, the requisite components have been presented for constructing and connecting neural populations that can encode and transform information in spiking neurons using the NEF. In this section, the NEF is used to construct a spiking neuron implementation of the ARC. It is presumed that the same computations for selective routing are performed in each level of the visual hierarchy, and as such, the description and simulations presented here are performed using only a subset of the columns from the large-scale model. Given the model's proposed consistency of function within and across areas, it is presumed that if the model can be shown to perform selective routing as predicted by the mathematical model, then extending the spiking implementation with other columns and levels will also be consistent.

Figure 5.3 depicts the mapping of the mathematical model presented in the previous chapter onto populations of neurons that are found in a single column. The column contains three functional types of neural ensembles: those involved in computing local control signals (shaded boxes), the layer-IV pyramidal neurons with nonlinear dendrites that selectively

gate the visual signals, and those providing feedforward visual signals (dashed boxes).

Input signals to the column are composed of both feedforward and feedback projections: the feedforward signals encode the visual signal \hat{x}_j from column j in the previous level, while the feedback signals encode the global control signals A_{pos} and A_{len} , as well as the relative shift of the attentional target, θ .

The two signals related to the position of the attentional target, A_{pos} and θ are projected to the same population of neurons, L5pos. To aide the transition from describing these populations in functional terms to laminar terms, the nomenclature presently used indicates the cortical layer in which the population is found (e.g. layer-V) and its functional role (e.g. encoding A_{pos}). Further details of the biological mapping of these populations on to specific cortical laminae is presented in Section 5.4.

We can begin by considering in detail the function of a small population of L5pos neurons. The 100 neurons in this population, along with all other neurons in the model unless indicated otherwise, have maximum firing rates drawn from a uniform distribution over the range [60, 120] Hz, a membrane time constant of 20 ms, and a refractory period of 2 ms. As neurons in this population receive two signals, A_{pos} and θ , this can be treated as a two dimensional vector, where each input dimension can have values in the range [-1, 1]. If both input signals have a value of 1, this produces a vector of length $\sqrt{2}$. Thus, in order for this population to be able to accurately represent values in this range, when constructing the population, their preferred direction vectors are drawn from a circle having a radius of $\sqrt{2}$, with the first dimension encoding A_{pos} and the second dimension encoding θ . This distinction of the dimensions in which each signal is encoded is made to lay the groundwork for the discussion of optimizing the encoders and decoders of the dendrites later in this section.

The output projections from L5pos are to a population of 100 neurons, L6shift, that use the two dimensional signal from L5pos to compute the amount by which the target's representation is shifted from the level below (Equations 4.2, 4.3, and 4.4).

The second type of feedback signal arriving in the column is the global control signal that specifies the size of the attentional target, A_{len} . For columns in PIT, both A_{pos} and A_{len} originate in pulvinar, while in levels below PIT, it is fed back from the next higher level. This signal is received by an ensemble of 100 neurons, L5len, that encode A_{len} as a one dimensional vector. The L5len population relay the global control signal to a population of 100 neurons, L6sf, where the synaptic weights between L5len and L6sf approximate Equation 4.1, which specifies the sampling factor (sf), or the number of input level columns by which values of μ are separated for two adjacent output level columns (Figure 4.2). As with L6shift, neurons in L6sf project to the dendrites of layer-IV neurons. The output of L6shift s and L6sf sf , is then scaled to the range $[\frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}}]$. The reason this value is scaled is because neurons in this population project to the nonlinear dendrites

Table 5.1: Summary of representation and transformations in the nonlinear dendrites of layer-IV pyramidal cells.

Source	Signal	L4dend dimension	Transformation
L6shift	s	1	$sf i + s$
L6sf	sf	1	$sf i + s$
L6sf	sf	2	$sf/2.35$
x_j	\hat{x}_j	3	\hat{x}_j

of the layer-IV neurons (L4dend), which encode three dimensional signals; scaling each of the input dimensions to $\frac{1}{\sqrt{3}}$ allows the layer-IV dendrites to have encoding vectors drawn from the unit sphere, and normalizing each dimension ensures that each dimension can be represented with a similar accuracy.

Depending on the size of the attentional target, the width of the routing function σ_{att} can take on different values at each level. Equation 4.8 shows that σ_{att} is a function of sf : $\sigma_{att} = sf/2.35$. This function is approximated by the synaptic weights between L6sf and L4dend, along with the communication channel relaying the value of sf .

The dendrites of the layer-IV pyramidal cells are the point at which much of the interesting processing occurs in the ARC. As was discussed in the previous chapter, they represent the confluence of the feedforward visual signals \hat{x}_j and the local control signals in all levels of the visual hierarchy. Thus, the layer-IV dendrites of neurons in column i receive three signals, two from neurons within the column (s and sf), and the feedforward signal (\hat{x}_j). The first two signals from within the column, s and sf have both been scaled to the range $[\frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}}]$, and are added together in the first dimension of L4dend to specify the attentional target as $\mu_i = s + sf i$ (Equation 4.6). Since i is fixed for each column, its value is not explicitly represented, but is encoded in the synaptic weights between L6sf and L4dend.

The projection from the L6sf population carrying sf is also transformed by the synaptic weights between L6sf and L4dend to compute the width of the routing function, $\sigma_{att} = sf/2.35$, which constitutes the value in the second dimension of L4dend. Finally, the feedforward signal \hat{x}_j is also scaled to the range $[\frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}}]$ by the synaptic weights, and is represented in the third dimension in L4dend.

To summarize, the dendrites receive three distinct input signals, the processing of which is given in Table 5.1. From left to right, the columns in Table 5.1 specify the following: 1) the population providing the input, 2) the signal being projected to the dendrites, 3) the dimension in which it is encoded in the dendrites, and 4) the transformation that is performed on the input.

Dendritic nonlinearities endow the pyramidal cells with more computational power, as they allow the neurons to compute a second nonlinear function, similar to a two-layer artificial neural network (ANN). However, this model differs from a two layer ANN because of the constraint that any dendritic branch projects to only one neuron, as compared with an ANN, in which input layer nodes can project to multiple or all hidden units. Estimates of the number of terminal dendritic branches with nonlinearities per pyramidal cell are on the order of 50-100 [137], and the model presented in this section uses 50 per neuron.

The output of L4dend is to the soma (L4soma). In the ARC, input connections to the soma and also to their dendrites, are constrained such that each layer-IV neuron in column i receives inputs from a population of layer-II/III neurons in single column j in the preceding level. The basis for this is that it eliminates the need for the synaptic weights of the inputs to L4dend to encode the identity of the column providing their feedforward input, and for this information to also be carried to the soma where the routing function is computed for a fixed x_j . The consequence of this arrangement, wherein a given layer-II/III neuron pools inputs from multiple layer-IV pyramidal neurons receiving visual signals from a single column, is that neurons in layer-II/III will exhibit larger receptive fields than the neurons providing their input. This has indeed been reported for neurons in primary visual cortex (area 17) of the cat [70], where layer-IV receptive fields were the smallest, layer-II/III were intermediate, and layer-VI were the largest. Similar findings have also been reported in somatosensory areas 3b and 1 in macaque monkeys, where neurons in layer-IV were found to have smaller receptive fields than those of neurons above or below layer-IV [232].

It is possible that in actual cortex, different dendritic branches of a given layer-IV neuron will receive signals from multiple input columns. If so, this may be incorporated into the current model by specifying different values of x_j to different dendritic branches of the same neuron and having them carry this value to the soma. This modification does not affect the principles of routing proposed by the model, although for simplicity, the implementation presented here follows the former assumption.

The layer-IV dendrites send to the soma, a three dimensional vector containing the values specified in the rightmost column of Table 5.1. The second nonlinearity of the two-layer neuron model [181] used for the layer-IV neurons, takes the three dimensional signal projected from the dendrites, scales the values back to the range $[-1, 1]$, and computes the routing function (Equation 4.7), which provides a multiplicative gain term that is applied to the feedforward visual signal as $f(\mu_i, x_j) \times \hat{x}_j$. This produces a one dimensional value that is then projected to the L2/3 population, where the outputs from multiple layer-IV neurons are pooled. Neurons in L2/3 in turn send the gated visual signals as feedforward projections to layer-IV neurons in the next level of the hierarchy. In the simulations presented in Chapter 6, many of the neurons recorded in experimental studies can be well described as being “on” neurons, and thus are used to model layer-II/III neurons (Figure 5.4(a)). Figure 5.4(b) shows the tuning curves of a population containing both

“on” and “off” neurons.

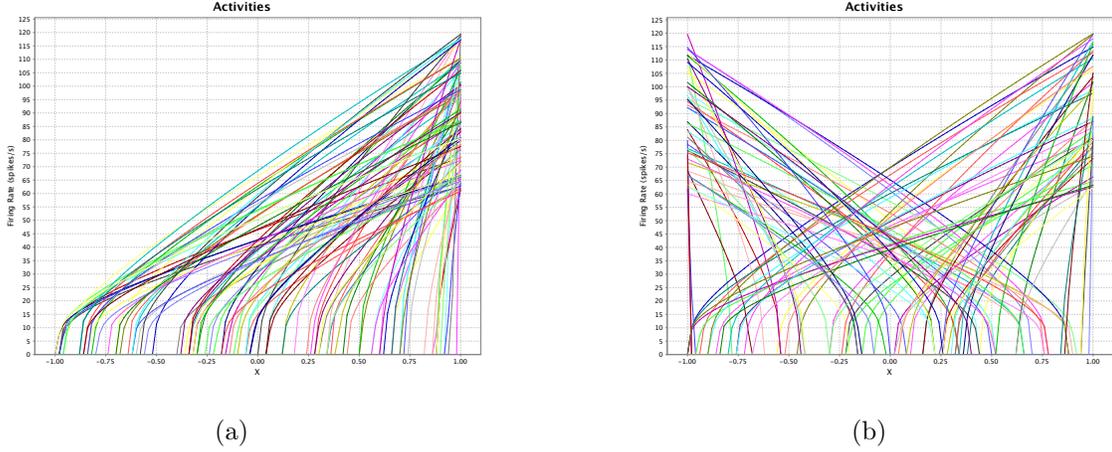


Figure 5.4: Response functions for a population of 100 neurons with maximum firing rates in the range $[60, 120]$ Hz. (a) These neurons have encoding vectors of 1 (“on” neurons), and thus respond increasingly to larger positive values. (b) Neurons with one dimensional encoding vectors of -1 (“off” neurons) and 1 (“on” neurons).

In most cases, solving for the decoding vectors in the NEF (Equation 5.8) is performed by minimizing the error of the decoded estimate over a distribution of possible values, with each sample value having equal probability within a given range. However, if distribution over the range is non-uniform, and is known in advance, this structure can instead be used when solving for the decoders.

For the dendrites in this model, the distribution of input values they will receive can be well characterized for a given network architecture. The first dimension represented in the dendrites receives the signals sf and s and uses them to compute $\mu = sfi + s$. For this column, it is most likely that μ will take on values of -0.25, 0, or 0.25, indicating that the routing function is centred on one of the input columns. Since the values in each dimension in the dendrites is scaled, the expected scaled value of $\mu = \frac{-1}{\sqrt{3}}, 0, \frac{1}{\sqrt{3}}$ in this dimension can be used when solving for the decoders. The second dimension uses sf to compute the width of the routing function σ_{att} , which is scaled to the range $[\frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}}]$. It is assumed that in this network architecture, the attentional target can span 3, 4 or 5 columns, which results in the width of the routing function having three possible scaled values: $\frac{-1}{\sqrt{3}}$ for targets of length 3, 0 for targets of length 4, and $\frac{1}{\sqrt{3}}$ for targets of length 5. The third dimension encodes the visual signal \hat{x}_j , which after scaling, can take on a value anywhere in the range $[\frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}}]$.

The above characterization of values likely to be represented in the dendrites are used

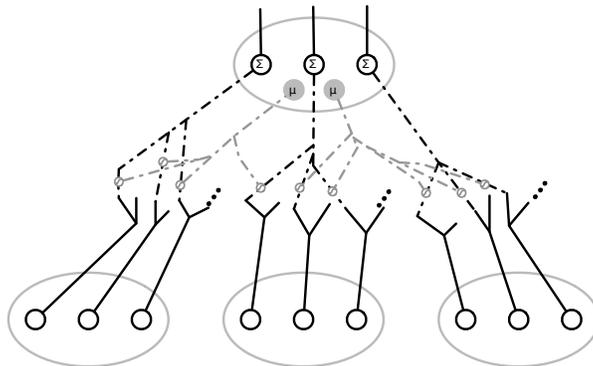


Figure 5.5: Arrangement and connectivity of control and sensory neurons in the ARC. Gray ellipses denote a column containing three visual processing neurons (unfilled circles); gray circles: control neurons; gray dash-dotted lines: axonal projections from the control neurons; black dash-dotted lines: pyramidal cell dendrites; black lines: pyramidal cell axons.

when solving for their decoders by specifying the points in 3D space from which samples should be taken. Specifically, 1000 evaluation points were used in solving for the decoders, where the value in each dimension of each sample point was selected from a Gaussian cloud around a randomly selected point consistent with the distribution of values just described. Specifically, a random value from each of the sets $(\frac{-1}{\sqrt{3}}, 0, \frac{1}{\sqrt{3}})$, $(\frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$, and $(\frac{-1}{\sqrt{3}}, 0, \frac{1}{\sqrt{3}})$ was selected giving a 3D point. This point was then used as the mean of a Gaussian cloud with a variance of 0.1 in every direction, from which the final 3D point was selected.

Given that the values in each dimension are well characterized, and the computations performed by the dendrites are known, this information can be used to further improve the representational accuracy of the dendrites by biasing the distribution of encoding vectors toward the first two dimensions which receive sf and s . Specifically, for $\frac{2}{3}$ of the dendrites, the encoders are assigned random values from a uniform distribution in the range $[\frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}}]$ for the first two dimensions, and zero for the third dimension. The remaining $\frac{1}{3}$ of the dendrites have encoders that are $[0, 0, \pm\frac{1}{\sqrt{3}}]$. This arrangement results in $\frac{2}{3}$ of the dendrites being able to better encode the first two dimensions, which allows for the routing function to be more accurately computed. These encoder and decoder selection methods are not essential for the functioning of the model, but instead help ensure that few neurons are able to sufficiently accurately compute the desired function. The same result could be obtained with many more cells using the standard NEF methods (i.e., random even distribution of encoders, and optimization of decoders over an even distribution on the range of values).

Figure 5.5 illustrates how the control signals interact with the feedforward signals. The bottom row of the figure contains three input columns that project to a single column.

Black lines indicate axonal projections from the input columns carrying the visual signals, black dash-dotted lines represent dendrites of the layer-IV neurons in the upper level, and gray dash-dotted lines indicate the axonal projections of control neurons. Control neurons project local control signals s and sf to the terminal dendrites of the layer-IV neurons, where feedforward projections from the input columns also terminate. This figure illustrates the scalability of the control calculations, which do not require additional control neurons to be included if more input columns are added, since the same control signals (s and sf) are sent to all layer-IV neurons in that column.

5.2.1 Simulation Results from Spiking Neuron Implementation of the ARC

Figure 5.6 illustrates the organization of the model used to demonstrate selective routing in spiking neurons. The shaded block at the top of the figure shows five input columns projecting to three columns in the upper level, where the numbers inside the circles indicate each column’s spatial position. In these simulations, the columns within the non-shaded triangle were implemented. The lower portion of the figure shows an expanded view of the connectivity. The model is similar to the functional diagram shown in Figure 5.3, but with additional input columns providing feedforward visual signals. Feedback signals specifying the global control signals A_{pos} and A_{len} are received by neural ensembles in layer-V (L5pos and L5len), which project to layer-VI populations L6s and L6sf. From layer-VI, local control signals s and sf are sent to the dendrites of layer-IV pyramidal cells (D), which also receive feedforward visual signals from a single column in the previous level ($x_{-0.25}$, x_0 , $x_{0.25}$). The dendrites of the layer-IV neurons then project to their somata (L4), and the output of the layer-IV cells is pooled by neurons in layer-II/III. The properties of each ensemble and their neurons are as specified in the previous section.

To demonstrate that the spiking neuron implementation of the ARC performs routing as predicted from the mathematical model, the model shown in Figure 5.6 was tested for its ability to selectively route visual information with different attentional targets and visual input signals. Figure 5.7(a) depicts the first configuration that was tested. The columns that are implemented in this simulation are outlined in black, and for clarity, column $i = 0$ in the upper level is expanded to illustrate that subpopulations of layer-IV neurons in the column each receive feedforward visual signals from input columns $x_{-0.25}$, x_0 , and $x_{0.25}$. The numbers inside the circles indicate each column’s spatial position. For illustrative purposes, the routing function is depicted by the red line.

For the first 750 ms, the attentional target covers all 5 input columns, with each input column encoding a signal with a value of 1. After 750 ms, the size of the target (A_{len}) is reduced to span only the centre three columns, resulting in a reduction of σ_{att} and thus a

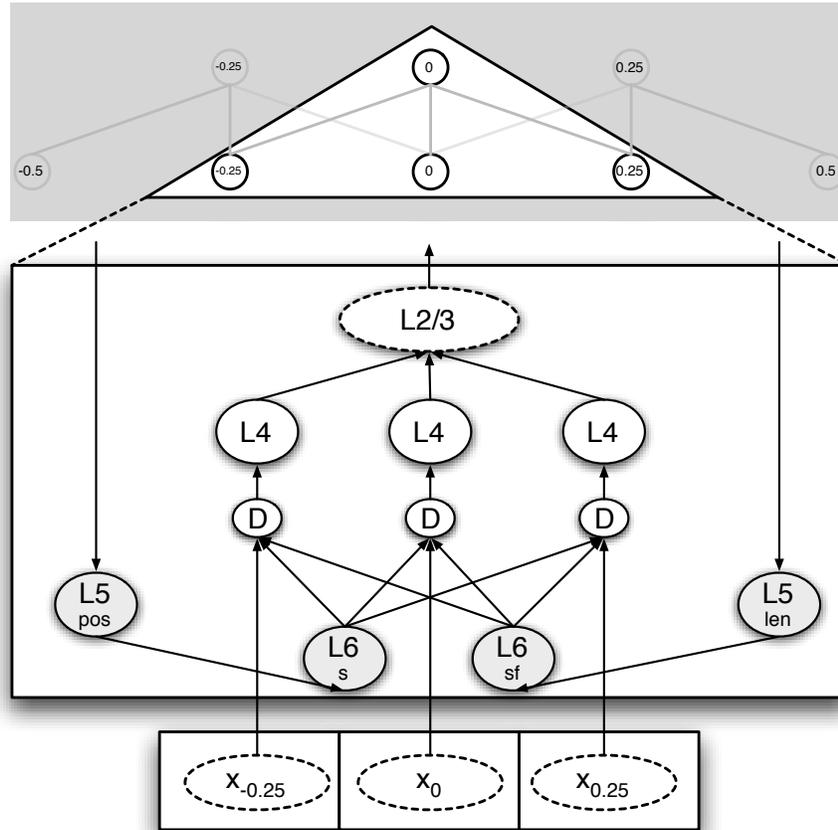


Figure 5.6: Circuitry for selective routing in a cortical column in the ARC. The shaded top block of the figure situates the modelled column among others in each level. The non-shaded region, containing one column that receives inputs from three columns in the previous level, is expanded in the lower portion of the figure. L2/3 – layer-II/III neurons. L4 – layer-IV pyramidal cell somata. D – nonlinear dendrites of layer-IV neurons. L5pos and L5len – layer-V control neurons receiving global control signals. L6sf and L6s – layer-VI control neurons computing local control signals. $x_{-0.25}$, x_0 , $x_{0.25}$ – layer-II/III neurons providing feedforward visual signals.

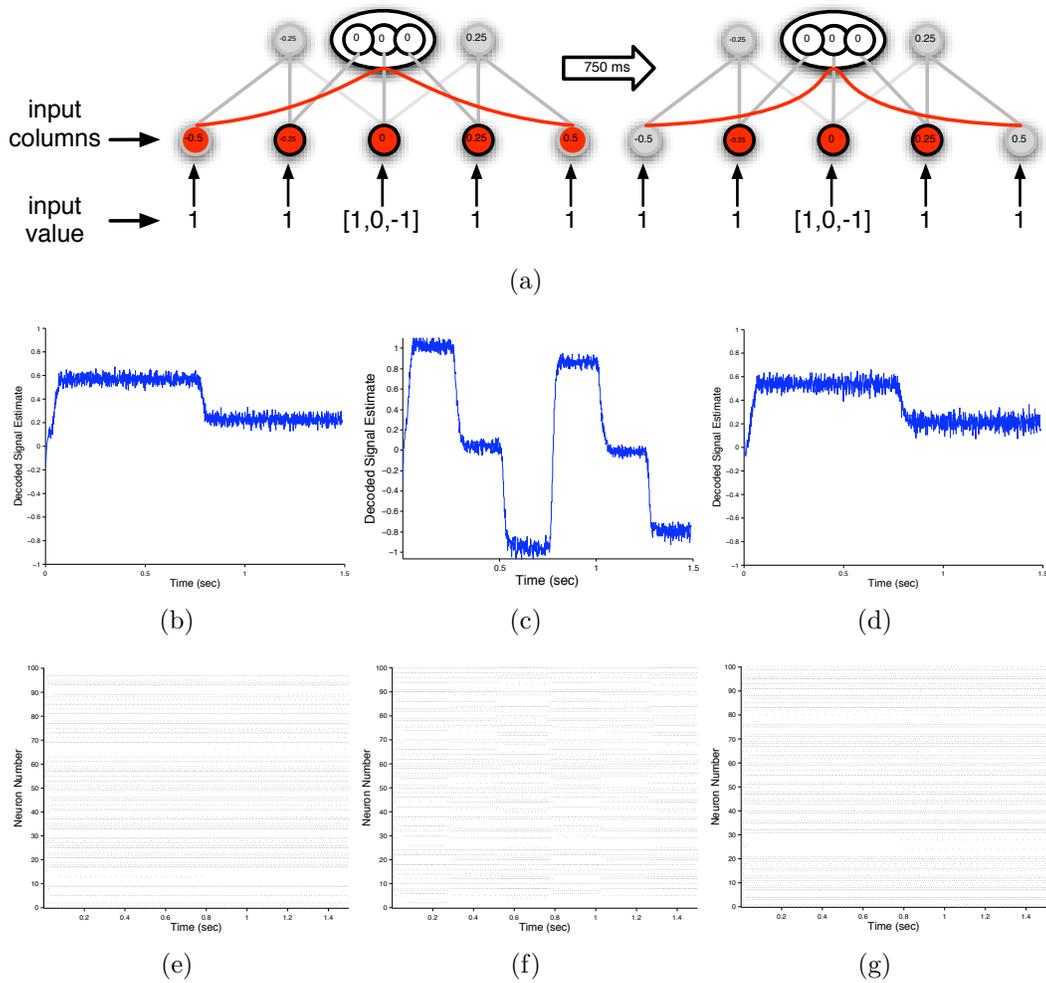


Figure 5.7: Simulation results of selective routing in the ARC. (a) Network architecture used in the simulations. Columns that were implemented are outlined in black, the number inside each circle indicates the column’s spatial position, and red circles indicate the attentional target. Column $i = 0$ in the upper level is expanded to reflect its connectivity. Each input column projects a value of 1 except for column x_0 , which changed every 250 ms from 1 to 0 to -1. For the first 750 ms, the attentional target (red circles) covered five input columns, and after 750 ms, its size was reduced to three columns. Decoded estimates of the gated signals are shown for columns receiving inputs from: (b) $x_{-0.25}$, (c) x_0 , and (d) $x_{0.25}$. (e-g) Spike rasters from layer-IV neurons receiving inputs from (e) $x_{-0.25}$, (f) x_0 , and (g) $x_{0.25}$. Notice that when the target’s size is reduced, the gain of the signals in (b) and (d) is also reduced, while (c) is relatively unchanged.

narrowing of the routing function. To demonstrate the effect of the routing function’s gain term on the visual signals in layer-IV, for each size of attentional target, the magnitude of the input signal in the centre column was decreased every 250 ms from 1 to 0 and then to -1.

Figure 5.7(b) shows the decoded estimate of the gated visual signal for layer-IV neurons that receive feedforward visual signals from input column $x_{-0.25}$. For the first 750 ms, the larger attentional target causes the routing function to be wider, resulting in a moderate attenuation of the visual signal from $x_{-0.25}$. When the size of the attentional target is reduced after 750 ms, the width of the routing function decreases, which results in the signals from $x_{-0.25}$ being significantly more attenuated, as seen in the reduction of the decoded estimate of the gated signal.

Figure 5.7(e) shows the spike raster from this population, where each dot represents a single action potential, and each row of dots represents the spike train from a single layer-IV neuron. The spike raster illustrates that, following the change of the target’s size at 750 ms, the activity patterns reflect the change in the signals being encoded by these neurons. This effect is most pronounced in some cells that stop responding when the target changes size. For example, the activity of neuron 15 in Figure 5.7(e) indicates that it is an “on” neuron with an x -intercept that is greater than 0.2. This property can be deduced because the neuron responds to the larger encoded value in the first 750 ms (~ 0.6), but does not respond when the encoded value is ~ 0.2 during the period 750-1500 ms. Conversely, the activity of neuron 65 indicates that it is an “off” neuron with an x -intercept that is less than 0.6, based on its lack of activity for the larger encoded value during the first 750 ms, although it responds when the encoded value is ~ 0.2 during the second half of the simulation. Finally, some cells do not exhibit any activity with either attentional target as a result of them being “off” neurons with x -intercepts that are less than the encoded value in both conditions.

Figure 5.7(c) shows the decoded estimate of the gated visual signal from layer-IV neurons receiving feedforward signals from input column x_o . With each attentional target size, the magnitude of the visual signal encoded by input column x_o was decreased every 250 ms from 1 to 0 to -1. Figure 5.7(a) shows that for both target sizes, the value computed by the routing function should not change for these neurons, as $\mu=0$ for both target sizes. However, the decoded estimate of the gated signal is slightly attenuated when the size of the target is reduced. This effect occurs because of an interaction between the encoded dimensions in the dendrites, which results in a moderate reduction of the routing function. This issue is discussed in more detail later in this section.

Figure 5.7(f) shows the spike rasters from the layer-IV neurons receiving signals from input column x_o . In this figure, the activity of “on” and “off” neurons is more apparent, with some neurons only responding when the encoded signal has a large positive value (e.g. cell 53 in Figure 5.7(f)), and others that only respond when the encoded signal is negative

(e.g. cell 72).

Finally, Figure 5.7(d) shows the decoded estimate of the gated visual signal from the layer-IV neurons receiving feedforward visual signals from input column $x_{0.25}$. Comparing this signal with that shown in Figure 5.7(c), the decoded values are similar as is expected, since the position of the target does not change and the gain term computed by the routing function is the same for neurons receiving inputs from $x_{-0.25}$ and $x_{0.25}$ for each target size. The spike rasters for neurons receiving inputs from $x_{0.25}$ are shown in Figure 5.7(g). Comparing this figure with the spike raster of neurons receiving inputs from input column $x_{-0.25}$ (Figure 5.7(e)) shows that the two rasters have distinctly different activity patterns, yet both populations are able to similarly encode the signals.

To verify that the spiking neuron implementation of the ARC performs routing as predicted by the mathematical model when the target’s position changes, an additional simulation was conducted with the network presented in Figure 5.6, using the same model as was used for the simulations presented in Figure 5.7. In this simulation, the attentional target covered three input columns, and every 250 ms, the position of the target moved from being centred on $x_{0.25}$ ($A_{pos} = 0.25$) to x_0 and then to $x_{-0.25}$. Figure 5.8 shows the decoded estimates of the gated visual signals as the attentional target changes position, and the stimulus configuration of this simulation is shown in Figure 5.8(a).

Figure 5.8(b) shows the decoded estimate of the gated visual signal from layer-IV neurons receiving feedforward visual signals from input column $x_{-0.25}$. For the first 250 ms, the value of μ computed by the control neurons is 0.25, which results in the signals from input column $x_{-0.25}$ being almost entirely attenuated. However, at 250 ms, the target’s position shifted to x_0 causing the centre of the routing function to be closer to $x_{-0.25}$, which reduces the attenuation of signals from that column. At 500 ms, the target’s position is centred on $x_{-0.25}$, and thus the visual signal is multiplied by 1 and passes through with minimal attenuation. The corresponding spike raster for this population is shown in Figure 5.8(e).

Figure 5.8(c) plots the decoded estimate of the gated visual signal from layer-IV neurons receiving feedforward projections from input column x_0 . For the period 250-500 ms, μ was centred on x_0 , resulting in the visual signals from that column passing through with minimal attenuation. For the first and last 250 ms intervals of the simulation, μ was adjacent to x_0 at positions $x_{-0.25}$ and $x_{0.25}$, resulting in the routing function attenuating the input signals by the same amount in both conditions. Figure 5.8(f) shows the corresponding spike raster from this population as the target’s position shifted.

The decoded estimate from the layer-IV neurons receiving visual signals from input column $x_{0.25}$ is shown in Figure 5.8(d). As expected, the decoded estimate decreases as μ moves away from $x_{0.25}$. Since this column’s position is symmetric to $x_{-0.25}$ in the receptive field, the signal is isomorphic by mirror symmetry to the values shown in Figure 5.8(b),

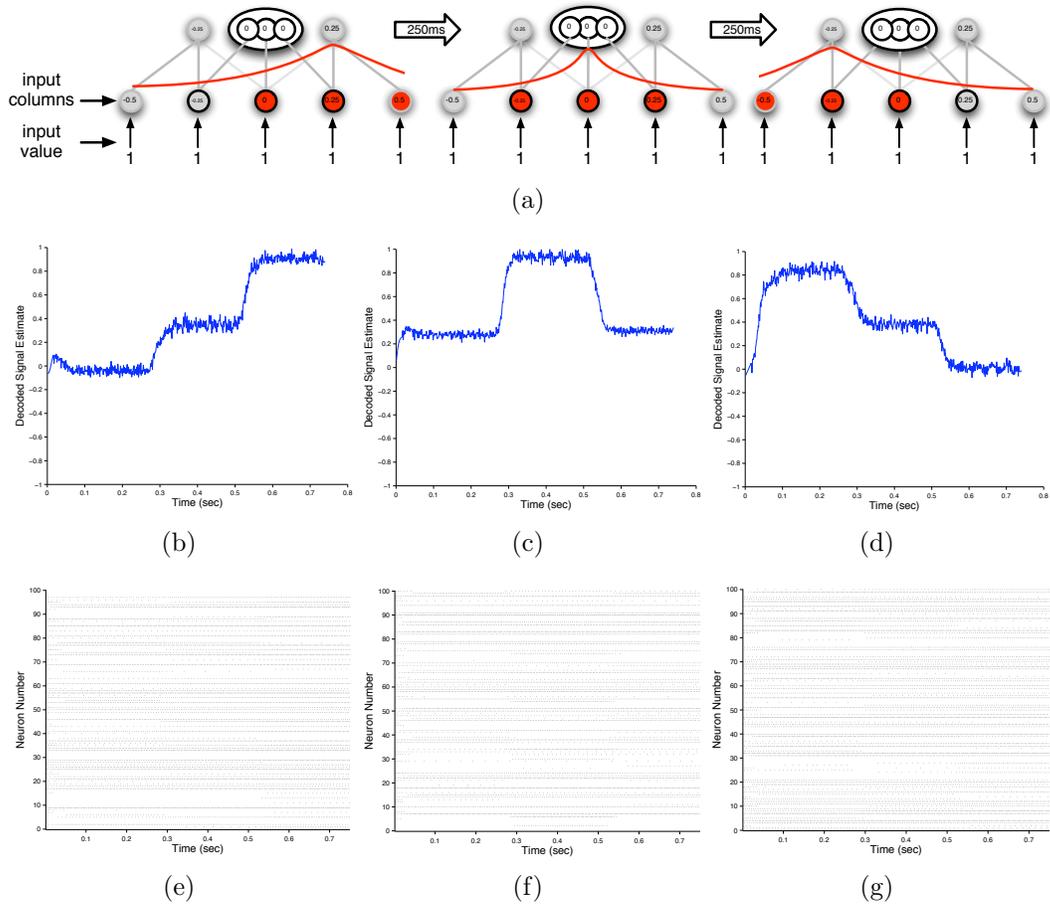


Figure 5.8: Simulation results of selective routing in the ARC. The stimulus configuration is shown in (a), where the target (red circles) covers three input columns, but changes positions every 250 ms. Columns outlined in black were implemented, numbers inside each circle indicates the column’s spatial position, and column $i = 0$ in the upper level is expanded to show that subpopulations of layer-IV neurons receive inputs from different input columns. (b-d) Decoded estimates of the gated visual signals are shown for neurons receiving inputs from input columns (b) $x_{-0.25}$, (c) x_0 , and (d) $x_{0.25}$. Spike rasters are shown for layer-IV neurons receiving inputs from (e) $x_{-0.25}$, (f) x_0 , and (g) $x_{0.25}$. Notice that in (b) and (d), the decoded estimates of the signals are symmetric since the columns providing their input signals are positioned symmetrically in the receptive field, while in (c), the decoded estimate is the same when the routing function is centred on either column adjacent to input column x_0 .

with some variation arising from the different neuron tuning in each population, as can be seen in the spike rasters shown in Figure 5.8(g).

As is expected from the mathematical model, the decoded estimates of the signal are similar in all three populations of layer-IV neurons when the routing function is centred on the column providing their input and when it is centred on the adjacent column. The visual input signals have values of 1, and thus the decoded estimate of the gated signal is equivalent to the value computed by the routing function. With each target position relative to the column providing input to each layer-IV population, a similar decoded estimate of the gated signal is produced, demonstrating that the routing function is computing a similar gain term in each column.

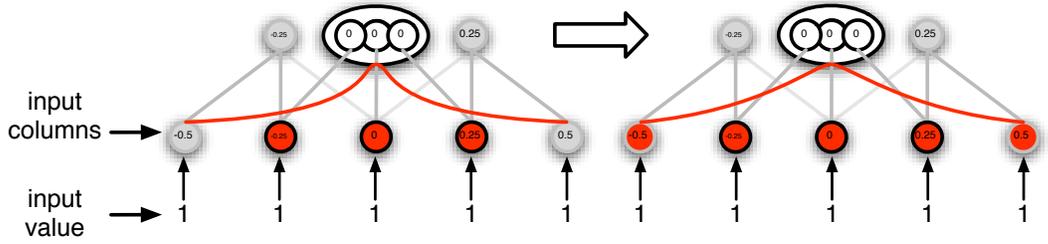
To provide further insight into the processing and representation in the ARC, Figure 5.9 shows the decoded estimates of the visual signals and local control signals from the dendrites and somata of the layer-IV neurons. In this simulation, the attentional target was centred on x_0 . Initially, the target covered the middle three columns, and then its size was increased to cover all five input columns (Figure 5.9(a)).

As with the previous simulations, the neurons in layer-IV are organized such that each column in the previous level projects to a different ensemble of neurons. This arrangement is reflected in Figure 5.9(b), where the plots in the leftmost column are for neurons receiving input signals from column $x_{-0.25}$, the centre column is for neurons receiving inputs from x_0 , and the rightmost column for $x_{0.25}$. The change in the decoded signal that occurs in the middle of each of the nine plots corresponds to the point at which the size of the target A_{len} was increased from spanning the centre three columns to cover all five input columns.

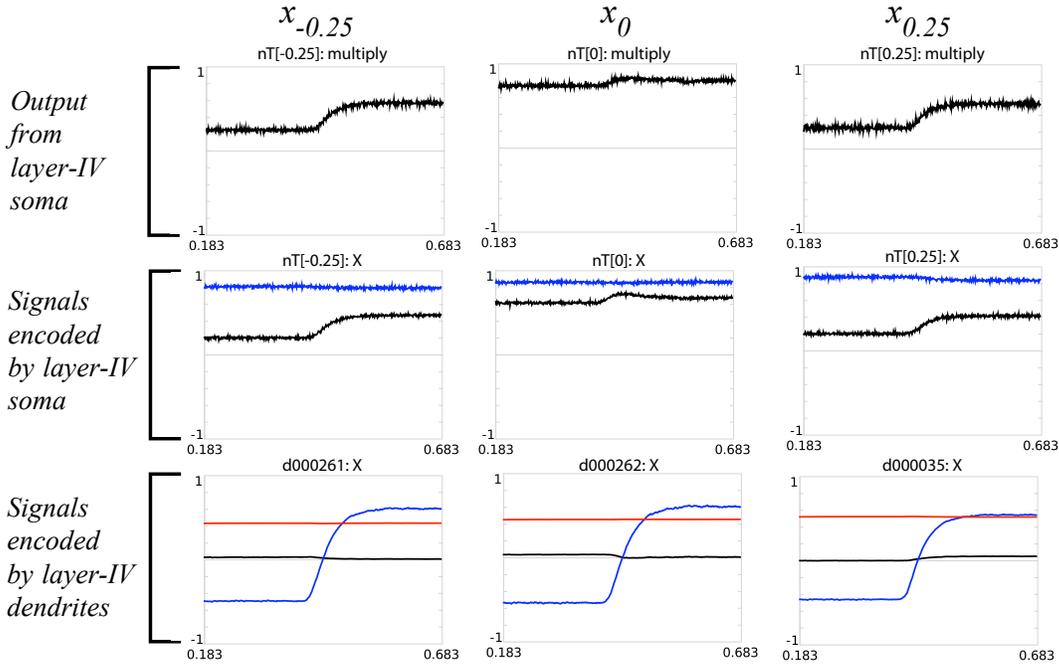
The top row of Figure 5.9(b) shows the decoded value of the gated signal that is projected out of each ensemble of layer-IV neurons, where the y -axis indicates the signal magnitude and the x -axis is time (seconds). These plots show that as the size of the target A_{len} is increased, the width of the routing function σ_{att} also increases, thereby reducing the attenuation of signals from input columns $x_{-0.25}$ and $x_{0.25}$. This is reflected in the decoded estimate of the signal increasing in the left and right columns when the target size is increased, and is similar to the simulation shown in Figure 5.7, but with constant input values, and the target's size increasing rather than decreasing.

The second row of Figure 5.9 shows the decoded estimates of the signals arriving at the soma in layer-IV, where the visual signal is shown in blue, and the result of the routing function is shown in black. Examination of the values in `nT[0]:X` (centre column) reveals that as the size of the target is increased, there is a slight increase in the result of the routing function. This effect is a result of the scaled value of σ_{att} temporarily moving away from the outer limits of the encoding vectors.

The third row of Figure 5.9 shows the scaled values encoded in the dendrites, $sfi + s$ (black), σ_{att} (blue) and \hat{x}_J (red). These estimates are from the 50 dendritic subunits of



(a)



(b)

Figure 5.9: (a) Network architecture, with the attentional target (red circles) covering the middle three input columns before being expanded to cover all five columns. The columns that are implemented are outlined in black, with the number inside each circle indicating the column's spatial position. Column $i = 0$ in the upper level is expanded to illustrate its connectivity. (b) Decoded estimates of the signals from the layer-IV somata and their dendrites. The three columns in the figure correspond to neurons receiving visual signals from input columns $x_{-0.25}$, x_0 , and $x_{0.25}$. Top row – value projected to layer-II/III, which is the product of the feedforward visual signal and result of the routing function. Second row – decoded estimate of the visual signal (blue) and result of the routing function (black). Bottom row – decoded estimate from the dendrites of a single neuron: visual signal (red), $sfi + s$ (black), and σ_{att} (blue), scaled to the range $[\frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}}]$.

a single layer-IV neuron. Ideally, each collection of the dendritic subunits shown in each column should be encoding the same value, as they are all receiving the same local control signals (sf and s) and the feedforward visual signal from each input column is 1. However, due to the limited processing and representational capacity of these dendrites, there is some variation in the encoded values across different collections of dendritic subunits. By using 50 layer-IV neurons, each having different dendritic subunits, this effect is largely ameliorated.

The plots in the bottom row of in Figure 5.9 show the signals that are encoded by the dendrites, with a moderate interaction between dimensions in the decoded estimates. As σ_{att} changes from $\frac{-1}{\sqrt{3}}$ to $\frac{1}{\sqrt{3}}$, the representation of $sfi + s$ (black lines) improves and moves closer to the correct value of zero. This interaction is present in the majority of dendrite “ensembles” (i.e. the dendrites of multiple layer-IV neurons), and is significantly stronger when random encoding vectors are used rather than the approach taken here in which $\frac{2}{3}$ of the dendritic subunits of each neuron have encoders biased toward the first two dimensions. This point is further demonstrated by observing that the decoded estimate of \hat{x}_j , for which $\frac{1}{3}$ of the encoders were specifically biased, is not significantly affected by changes in the other two dimensions.

5.2.2 Influence of the Number of Neurons and Dendrites in each Column

The network shown in Figure 5.6 illustrates the organization of the model used to demonstrate selective routing in spiking neurons. In this section, this architecture is used to examine how the accuracy of the gated signal is affected by the number of dendritic subunits per neuron and the number of neurons receiving feedforward visual signals from each input column. In these simulations, the control signals to output column i specified that attention should be directed to the centre of its receptive field. The input signals had a magnitude of 1 and were presented for 250 ms, with the first 50 ms discarded. Accuracy was measured by comparing the decoded estimate of the gated signal with the value produced in the math model using the root mean square error (RMSE). Ten repetitions of each simulation were performed, using 10, 20, 25, 50, 75 and 85 dendritic branches per neuron, and 25, 30, 35, 50 neurons in each layer-IV population.

Figure 5.10 shows the average RMSE for each pair of parameter values across 10 repetitions. Increasing the number of dendritic subunits per neuron significantly improves the accuracy of the signal, although there is little improvement in the decoded estimate beyond 50 terminal branches per neuron. Increasing the number of neurons receiving signals from each input column also improved the accuracy of the decoded signal, but in a less pronounced manner. Not surprisingly, the inclusion of additional dendritic subunits and

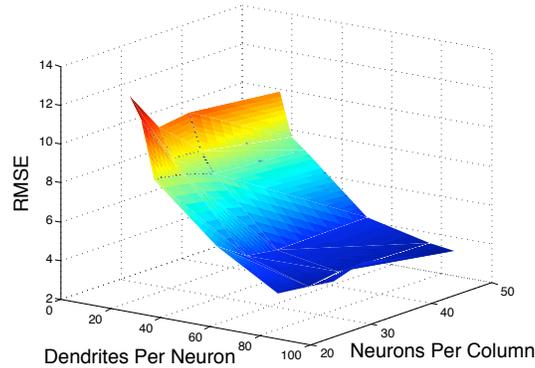


Figure 5.10: Effect of varying the number of dendritic subunits per neuron and the number of neurons receiving the visual signal from each input column on the accuracy of the gated signal. With additional dendritic subunits for each neuron, the error of the encoded signal (RMSE) significantly decreases. The error also decreases with additional neurons in each column, although the improvement is less pronounced.

neurons reduces the error in the decoded estimate of the signal, but significantly increases the computational demands. In the simulations presented in Chapter 6, the models use 50 neurons per column, each having 30 dendritic subunits, as this arrangement seems to provide a suitable trade-off between representational accuracy and computational efficiency.

5.3 Separating Inhibitory and Excitatory Connections

In Section 5.1, the methods for computing connection weights in the NEF were presented (Equation 5.9). Using that approach, the synaptic weight from neuron a_j to postsynaptic neurons in b_j could be negative (inhibitory) to some neurons in b_j and positive (excitatory) to others, as there were no constraints placed on the sign of the individual weights. However, in cortical neurons, synaptic connections from a given neuron are either excitatory or inhibitory. In many models of neural systems, this constraint is ignored, thereby limiting the models' biological plausibility. To overcome this limitation, Parisien, Anderson and Eliasmith [174] presented an elegant method (the APE transform) for transforming such models into more realistic forms by separating inhibitory and excitatory projections, while still maintaining the model's function and dynamics.

Consider the communication channel between population A and B show in Figure 5.1. To remove the negative weights of projections from A to B, the APE transform involves manipulating the encoding and decoding vectors using a two step process. The first step is to ensure that all weights between A and B are positive (excitatory) by adding a bias

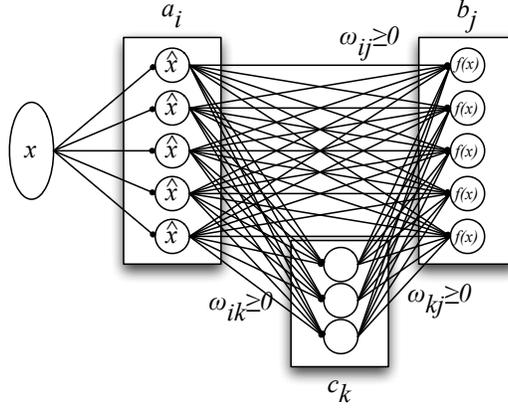


Figure 5.11: Communication channel in the NEF after applying the APE transform to ensure that synaptic connections from each neuron in population A are either inhibitory or excitatory. Population C contains inhibitory neurons to satisfy this constraint. See text for details.

to the synaptic weights of the b_j connections that is equal to the largest negative weight from A. In its original form, the current arriving at neuron b_j is:

$$J_j(\mathbf{x}) = \sum_{i=1}^N \omega_{ij}^{orig} a_i(\mathbf{x}) + J_j^{bias}, \quad (5.11)$$

where ω_{ij}^{orig} contains both excitatory and inhibitory weights. Adding a bias to these weights gives:

$$\begin{aligned} \omega_{ij} &= \omega_{ij}^{orig} + \Delta\omega_{ij} \\ \delta\omega_j &= -\min(\omega_{ij}^{orig}), \end{aligned} \quad (5.12)$$

which results in the current arriving at b_j being given by:

$$J_j(\mathbf{x}) = \sum_{i=1}^N \omega_{ij}^{orig} a_i(\mathbf{x}) + J_j^{bias} + \sum_{i=1}^N \Delta\omega_j a_i(\mathbf{x}) + J_j^{bias}. \quad (5.13)$$

Adding this additional bias $\sum_{i=1}^N \Delta\omega_j a_i(\mathbf{x})$ results in a greater input current being provided to that cell, which alters the function being computed at that connection. The second step of the transform mitigates this effect by adding a set of inhibitory neurons c_k that also receive projections from population A. First, the decoders of neurons in population B,

ϕ^f are scaled so that the new function computed using the augmented weights between A and B, $f(x)$ is constrained to the range $[0, 1]$. With ϕ^f defined, the encoders of neurons in b_j are then computed as:

$$\tilde{\phi}_j^f = \frac{\Delta\omega_{ij}}{\alpha_j\phi^f}. \quad (5.14)$$

As the bias function $f(\mathbf{x})$ will typically be nonzero for the entire domain of \mathbf{x} , separating it into a constant portion (f_1) and variable portion ($f_2(\mathbf{x})$) such that $f(\mathbf{x}) = f_1 + f_2(\mathbf{x})$ reduces the amount of inhibitory current, and thus firing rate, of neurons in the inhibitory population.

The bias current may now be expressed as:

$$J_j^f(\mathbf{x}) = \alpha_j\tilde{\phi}_j^f f_1 + \alpha_j\tilde{\phi}_j^f f_2(\mathbf{x}). \quad (5.15)$$

As the first part of Equation 5.15 is constant, by incorporating it into the background current of b_j we obtain:

$$J_j^{biasnew} = J_j^{bias} - \alpha_j\tilde{\phi}_j^f f_1. \quad (5.16)$$

The inhibitory interneurons in population C are set to respond more strongly to positive inputs (“on” neurons), since they receive only positive inputs as a result of scaling ϕ^f to the range $[0, 1]$. The current arriving at neuron C_k is given by:

$$\begin{aligned} J_k(\mathbf{x}) &= \alpha_k\tilde{\phi}_k f(\mathbf{x}) + J_k^{bias} \\ &= \sum_{i=1}^N \omega_{ik} a_i(\mathbf{x}) + J_k^{bias}, \end{aligned} \quad (5.17)$$

with the weights $\omega_{ik} = \alpha_k\phi^f$ being positive. The background current of C_k is then adjusted similarly as for neurons b_j :

$$J_k^{biasnew} = J_k^{bias} - \alpha_k f_1, \quad (5.18)$$

such that population C will output only the variable component of the bias function, $f_2(\mathbf{x})$. Similar to when solving for the decoders the communication channel (Equation 5.8), the decoders ϕ_k are determined for $f_x(\mathbf{x})$:

$$\hat{f}_2(\mathbf{x}) = \sum_{k=1}^N \phi_k c_k(f(\mathbf{x})). \quad (5.19)$$

The input current to population B, through both excitatory connections from A and in-

hibitory connections from C, is now computed as:

$$J_j(\mathbf{x}) = \sum_{i=1}^N \omega_{ij} a_i(\mathbf{x}) + J_j^{biasnew} - \sum_{k=1}^N \Delta\omega_{kj} c_k(f(\mathbf{x})), \quad (5.20)$$

where $\omega_{kj} = \alpha_j \tilde{\phi}_j^f \phi_k$.

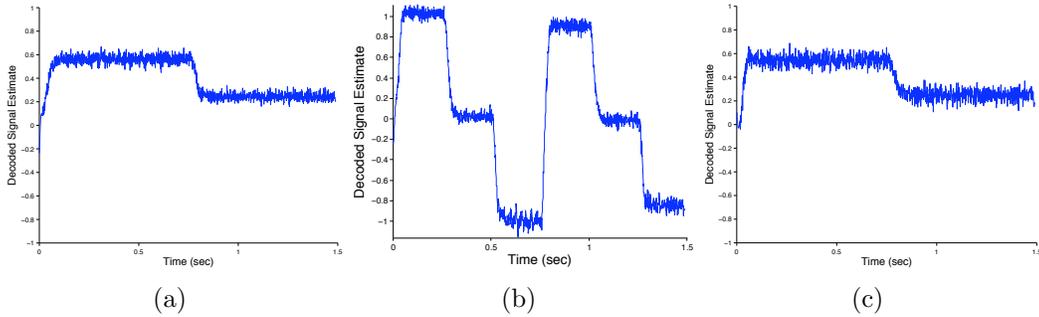


Figure 5.12: Effect of applying the APE transform to the ARC. Decoded estimates of the gated visual signals from three input columns when attention was directed to the receptive field centre. The stimulus configuration is the same as that used in Figure 5.7. Compared with the signal estimates from the model not using the transform, the gated signals are indistinguishable.

To assess the impact of using the APE transform in the ARC, the transform was applied to the model shown in Figure 5.6. Estimates of the proportions of inhibitory and excitatory neurons in visual cortex suggest that in layer-V, 15.8% of all synapses are inhibitory, while 16.2% of the synapses in layer-VI are inhibitory [17, 64], which corresponds to 18% of the cells in layer-V being GABAergic and 17% for layer-VI.

The model shown in Figure 5.6 was thus modified using the APE transform which introduced 22 additional inhibitory neurons in each of the layer-V populations (22/122=18%) and 21 inhibitory neurons in the layer-VI populations.

Figure 5.12 shows an example of routing in the ARC after applying the APE transform. The same stimuli as in Figure 5.7 were used, with attention directed to $x_j = 0$. For the first 750 ms, the attentional target covered five input columns, after which it was reduced to the centre three columns. With each attentional target, the visual signal in column $x_j = 0$ was changed from 1 to 0 and -1, with each stimulus being presented for 250 ms. Comparing these results to those shown in Figure 5.7 demonstrates that there is no appreciable difference in the decoded value in the model when using the APE transform.

5.4 Neurobiological Mapping

In the previous section, an overview was given of the calculations performed in the ARC for generating attentional control signals. This description is now extended by describing an anatomically consistent mapping of the ARC's control calculations to cortical neurons in specific laminae.

Anatomical findings place important constraints on neurobiological models in terms of the number of neurons, their intra- and inter-laminar input and output connections, and placement and properties of their somata, dendritic, and axonal arbours. Combining these anatomical constraints and previously defined computational specifications provides a firm basis for constructing a detailed model of attentional processing. By ascribing a particular function to neurons in a cortical layer places constraints on the possible computations that may be performed due to the anatomical connectivity and neuronal density of that layer. Further, it allows for the assessment of the plausibility of these functions being performed in actual cortical circuits by comparing the number of neurons required by the model with the number found in the corresponding biological area. The problem addressed here is to define a laminar circuit that is consistent with biological constraints and is sufficient for implementing the cortical calculations in the ARC.

In addition to the anatomical distinction of these layers, the ARC follows the idea that the laminar arrangement can be divided along functional lines as well [69, 104, 258]. The general proposal of laminar organization in the ARC is that neurons in layers-V and -VI are involved in computing the local control signals, neurons in layer-IV are involved in selective gating of inputs based on control signals from the deep layers, while neurons in layer-II/III serve to process the gated visual signals. The proposed physical and functional organization of cortical layers within each column is shown in Figure 5.13. This extends the cortical circuit for implementing the control calculations in a single column presented in Figure 5.3. Other plausible variations to this mapping may be found (for a detailed analysis of inter-laminar connectivity, see [50, 240, 241]). However, the arrangement proposed here is consistent with the most salient and well documented inter-laminar connections, and requires minimal control neurons and wiring.

As with previous proposals [33], minimizing the neural requirements and wiring between neurons is a central factor constraining the model's design. The model attempts to minimize the number of neurons required for computing the local control signals for two primary reasons. First, in visual cortex, there are significantly more neurons in layer-IV than in layers-V and -VI combined [17], suggesting that control signals are being distributed to multiple neurons. Second, as layer-V and VI neurons are also involved in other processing via their projections to subcortical structures and the thalamus [71, 97, 240], it is likely that only a fraction of the neurons in these layers are involved in computing local control signals. The ARC satisfies this constraint by having common neural ensembles

compute the local control signals for the entire column. This arrangement also permits the model to be scaled to have significantly more neurons in each column receiving feedforward inputs from many more columns, without requiring any additional control calculations to be made.

To reduce the computational demands of control neurons, the ARC attempts to minimize the number of highly nonlinear functions because these tend to require more neurons to compute in the NEF. Further, in the calculations of local control signals, parts of the routing function are linearly separable, and thus the ARC distributes these functions across different groups of neurons, so that each population is computing an easier function. Combining these components of the routing function together in the dendrites yields the same result, while requiring significantly fewer neurons.

5.4.1 Overview of Cortical Layers

The six layers of homotypical cortex can be distinguished by their patterns of connectivity and the types of neurons they contain. Starting at the outermost layer (i.e. closest to the skull), layer-I primarily contains the apical tufts of pyramidal cell dendrites, axon terminations and relatively sparse GABAergic inhibitory neurons [50]. Layer-I is the primary target of feedback projections from higher cortical areas, pulvinar, and from non-specific thalamus [96, 104]. Layers II and III are often grouped together as part of the superficial layers. Unlike layer-I, these layers are densely populated with small pyramidal, spiny stellate and inhibitory neurons [241]. Generally, corticocortical feedforward projections originate in the superficial layers and terminate in layer-IV along with dense thalamic afferents [50].

Layer-IV is particularly thick in cortical areas that process sensory information [97], and contains various types of pyramidal and inhibitory neurons. This layer is often divided into sublaminae (See Figure 2.1) and is the primary target of thalamocortical projections. In V1, retinal signals transmitted by LGN terminate in various sublaminae of layer-IV, while in extrastriate cortex, feedforward projections from layer-II/III largely terminate in layer-IV, with a minority of feedforward projections targeting layer-VI [84, 90, 119]. Layer-IV neurons have strong projections to nearby cells in layer-II/III that are organized in strong columnar fashion [50, 71, 119].

Layer-V contains many small- and large-bodied pyramidal cells and projects to a variety of subcortical structures including non-specific thalamus, pons and superior colliculus, as well as to other cortical regions [241]. In addition to their subcortical targets in pulvinar and superior colliculus [240], neurons in layer-V have local projections to nearby neurons in layer-VI [50, 240], and send feedback projections that terminate in layer-I [193, 194]. In layer-V, two primary neuron types can be distinguished by their firing patterns and

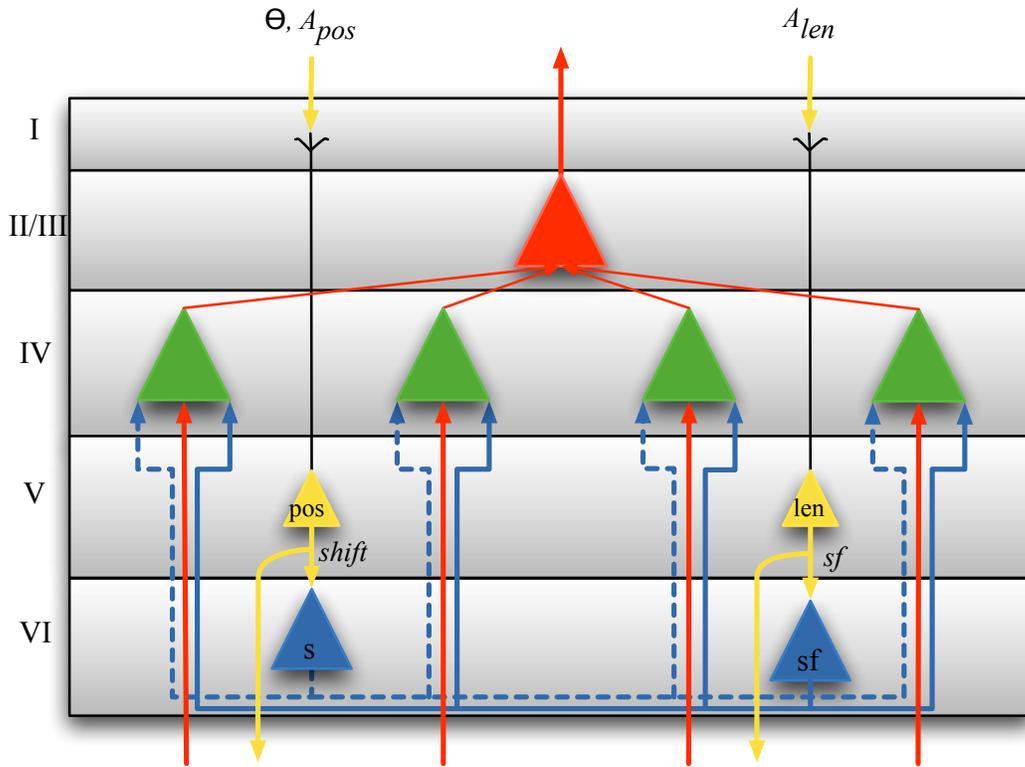


Figure 5.13: Laminar circuitry of attentional control in the ARC for a single column. Global attention signals (A_{len} and A_{pos}) and θ are fed back from the next higher cortical level to layer-I where they ramify on apical dendrites of layer-V cells. The layer-V neurons relay this signal to the next lower area with collaterals projecting to control neurons in layer-VI of that column where sf and s are computed. These signals, along with feedforward visual signals \hat{x}_j are received by layer-IV pyramidal cells where the routing function is computed in the dendrites and multiplied with \hat{x}_j at the soma. Cells in layer-II/III pool the activity of multiple layer-IV neurons and project the gated signal to the next higher level.

dendritic formation. Intrinsically bursting neurons are large-bodied pyramidal cells found in the lower portion of layer-V that have long apical dendrites with tufts located in layer-I, through which they have access to broadly distributed feedback signals from the level above, as well as from calbindin-expressing neurons in pulvinar and thalamus [214, 240, 241]. The second primary type of layer-V neurons are regular spiking cells that are small- or medium-bodied pyramidal cells with apical dendrites that seldom extend beyond layer-II/III.

Finally, layer-VI also contains large-bodied pyramidal neurons that can be divided into at least 8 subtypes [50]. In macaque primary visual cortex, two of these subtypes predominately project locally within layer-VI, while the axons of the remaining 6 subtypes largely arborize in the various sublaminae of layer-IV [22, 268]. Afferents from layer-VI to layer-IV are similar in density to thalamic inputs to layer-IV, with layer-VI pyramidal cells providing $\sim 45\%$ of the excitatory input to layer-IV [119], although the relative strength of these projections is not known [119]. In addition to their corticocortical projections, pyramidal neurons in layer-VI also make strong projections to various thalamic nuclei [71].

5.4.2 Laminar Mapping

Having now described the composition and connectivity of the cortical layers, the description of the ARC given in Section 5.2 is now extended by ascribing the model’s control calculations to different types of cortical neurons, and specifying the laminae in which such neurons may be found. Figure 5.13 outlines the general connectivity of neurons within a column.

Feedback projections from higher cortical areas (yellow lines) terminate in layer-I upon the apical dendrites of layer-V pyramidal cells. This type of layer-V control neuron is an intrinsically bursting (IB) pyramidal cell that receives the signal A_{len} . Such neurons have large cell bodies and long apical dendrites that spread widely in layer-I, providing them with access to the broadly distributed signals from the area above [214, 240, 241]. These neurons relay the signal A_{len} to dendrites in layer-I of the preceding area in the hierarchy [193, 194] and to the dendrites of nearby layer-VI pyramidal cells in the same column [50, 240]. It is proposed that at the synapse between the layer-V axons and layer-VI dendrites, sf (Equation 4.1) is calculated.

Other layer-V IB neurons with dendritic arbours in layer-I are involved in computing the relative shift of the target’s representation from the preceding level (s). These neurons receive feedback projections that encode A_{pos} and θ and have similar connectivity as those relaying A_{len} , except these neurons use the A_{pos} signal to compute s (Equation 4.4). This requires a greater amount of computation to be performed by neurons in layer-VI than in layer-V, as the latter simply relay the feedback signals to layer-VI neurons. This is consistent with reports of layer-VI having a greater neuronal density and relative layer

thickness than layer-V in striate and extrastriate cortex in both humans and non-human primates [46].

The primary efferents of layer-VI neurons are to nearby pyramidal cells in layer-IV [22, 241, 268], and in the dendrites of these layer-IV pyramidal cells, σ_{att} (Equation 4.8) and μ (Equation 4.6) are computed. The product of $f(\mu, x_j)$ and \hat{x}_j is then computed at the soma. The layer-IV neurons project the gated visual signals to neurons in layer-II/III, which pool the gated visual signals from multiple layer-IV neurons in their column. Cells in layer-II/III then project the gated signals to layer-IV cells in the next higher cortical area [50, 71, 119, 194].

The layer-IV neurons in the ARC are modelled using the 2-layer neuron model with nonlinear dendrites, where each layer-IV neuron receives inputs from a single column in the preceding level, and layer-II/III neurons in a given column pool the activity of multiple layer-IV neurons in that column. It is from these layer-II/III neurons that the activity is recorded in the simulations presented in the following chapter. This arrangement also requires considerably more complex computations to be performed by neurons in layer-IV than by those in layers-V and VI, although anatomical studies suggest that the number of neurons in layer-IV outnumbers those in V and VI by a factor of $\sim 20:1$ [129, 153].

Chapter 6

Attentional Routing Circuit - Simulations of Attentional Modulation Experiments

In this chapter, the spiking neuron implementation of the ARC presented in the previous chapter is used to simulate three experiments of attentional modulation in visual cortex. The first experiment by Womelsdorf et al. [272] investigates the effects of attention on spatial receptive field properties that occur when shifting attention between targets outside or inside the receptive field of the recorded cell. The second experiment by Treue and Martinez-Trujillo [245] examines attentional effects on feature tuning that occur when attention is directed to a target outside the receptive field, or to a preferred or anti-preferred stimulus inside the receptive field. The third experiment by Lee and Maunsell [112] investigates the influence of attention on contrast-response functions when attention is switched between receptive field targets. Each of these three studies describe seemingly different forms of attentional modulation, although as is discussed in Section 6.4, the ARC is able to reproduce each effect using the same mechanisms for selective routing described earlier in this thesis. That each of these effects can be demonstrated using the same routing mechanisms in a biologically detailed model, suggests that the ostensibly different forms of modulation may in fact be consequent to a common underlying routing principle.

6.1 Simulations of Womelsdorf et al. (2008)

In Womelsdorf et al. [272], recordings were made from area MT in macaque monkeys performing a spatial attention task. The study had the animal foveate a fixation point, after which a cue stimulus (stationary random dot pattern, RDP) was presented for 440

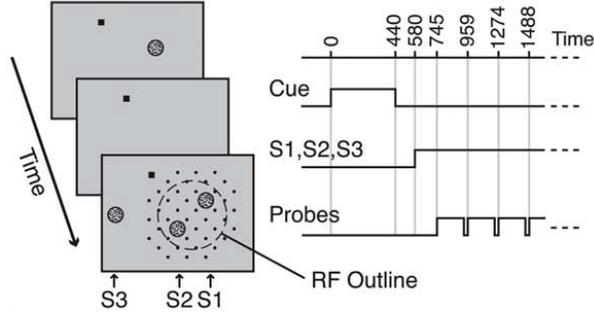


Figure 6.1: Experimental method used in Womelsdorf et al. (2008). (Reproduced from [272]).

ms at one of three target locations indicating where the animal should covertly attend (Figure 6.1). Following a delay, two RDPs were presented inside the isolated cell’s receptive field at equal eccentricity (stimulus S1 and S2), and a third RDP (S3) was placed outside of the receptive field in the opposite hemifield. The task was to detect a small, transient change in the direction of motion of the stimulus at the cued location. As the monkey covertly attended each of the three stimuli, probe stimuli (RDPs with higher contrast) were presented for 190 ms at various locations inside the receptive field, without overlapping the stimuli inside the receptive field, and neural activity was recorded in response to each probe. For each attended stimulus, the recorded activity was used to determine the amount that each probe contributed to the cell’s activity. Receptive field profiles were constructed from responses to each probe stimulus by subtracting the baseline activity and fitting the activity with a Gaussian having four parameters (mean, standard deviation, amplitude, and shift). Directing attention to a target stimulus inside the receptive field resulted in the peak of the Gaussian shifting toward the target while its width shrunk, without a significant change in peak activity.

This experiment was simulated using the ARC implemented in the network shown in Figure 6.2. The two level network is composed of a single MT column containing 100 control neurons, 100 layer-II/III neurons, and 9×50 layer-IV pyramidal neurons, each having 30 nonlinear terminal dendrites. Visual input signals come from nine V1 columns containing 100 neurons. Several studies have demonstrated that the sensitivity of visual cortical neurons across their receptive field has a Gaussian profile [35, 36, 171], and thus connection strengths between V1 and MT are determined by projecting a Gaussian with $\mu_\omega = 0$ and $\sigma_\omega = 1$ onto the connection weights using the methods of the NEF. Specifically, the Gaussian connection strengths were computed as:

$$\omega = e^{-\frac{(\mu_\omega - x_j)^2}{2\sigma_\omega^2}}. \quad (6.1)$$

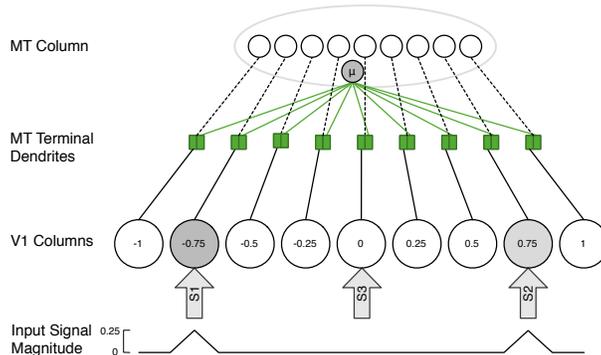


Figure 6.2: ARC with spiking neurons used for simulations of Womelsdorf et al. (2008). MT has a single column containing 450 layer-IV pyramidal neurons, subsets of which receive input from a single V1 column. Control neurons in the MT column (gray circle) project a local control signal to the terminal dendrites of MT neurons to enable selective routing. The spatial position of V1 columns is indicated inside the V1 columns and the magnitude of the visual signals encoded by each column is shown at the bottom of the figure. Attentional targets are shown as arrows (S1, S2, S3), with S3 corresponding to the attend-out/default routing condition.

Each layer-II/III MT neuron has a maximum firing rate drawn from a uniform distribution in the range $[90, 120]$ Hz, and responds more strongly to larger positive input values. 100 cortical control neurons also project to the terminal dendrites of MT cells and send the control signal μ , indicating the location within the receptive field from which the neurons should selectively process. In the dendrites of the layer-IV MT pyramidal cells, the feedforward visual signals (\hat{x}_j) and control signal (μ) interact according to Equation 4.7. Layer-IV neurons then project to 100 layer-II/III pyramidal cells, from which recordings were made of the spiking activity that is used in the subsequent analysis.

The two non-preferred reference stimuli were simulated using an input signal of 0.25, placed symmetrically at positions $x_j = -0.75$ and $x_j = 0.75$, with all other inputs set to zero (Figure 6.2). For the “attend-out” condition, the attentional control signal for the MT column is set to $\mu = 0$, such that attention is directed to the centre of the receptive field (S3 in Figure 6.2). This value is based on the data of Womelsdorf et al. [272] which show that in the attend-out condition, the peak of the receptive field occurs at a point that is roughly equidistant between S1 and S2 and corresponds to the neurons being in their default state in which their receptive fields are large, with the width of the routing function $\sigma_{att} = 1$. When attention is directed to either receptive field stimulus, the width of the routing function is set to 0.75.

Preferred probe stimuli with a magnitude of 0.5 were then presented at each of the

Table 6.1: Qualitative summary of simulation data from the ARC with experimental data reported by Womelsdorf et al. (2008).

		ARC		Womelsdorf et al. (2008)	
		Change Direction	Statistically Significant	Change Direction	Statistically Significant
Gain	Entire sample	-	no	+	no
	Selected pairs	-	no	+	no
Shrink	Entire sample	-	yes	-	yes
	Selected pairs	-	yes	-	yes
Shift	Entire sample	+	yes	+	yes
	Selected pairs	+	yes	+	yes

seven remaining V1 columns. After stimulating the MT neurons with the reference stimuli alone to establish baseline activity, each probe was displayed for 190 ms, during which time the activity of all 100 layer-II/III MT neurons was recorded. The same process was then repeated with attention focused on the left stimulus (S1; $\mu = -0.75$) and right stimulus (S2; $\mu = 0.75$) in the receptive field.

The value for σ_{att} in the attend-in condition was selected not by parameter fitting, but as being some smaller value than that used in the attend-out condition. With these simple values for σ_{ω} and σ_{att} , it is demonstrated that the model can account for the experimental results without taking recourse to parameter fitting. It may be possible to have the model more closely match the experimental data by adjusting the value of σ_{ω} and σ_{att} , although as is shown below, the model is able to reproduce their results without doing so.

To analyse the neuronal activity for each attentional condition, the methods of Womelsdorf et al. [272] were followed. Specifically, the analysis used the recorded data from 60 to 200 ms after probe onset for each probe condition and attentional condition, and the average response to each of the seven probes was calculated with the baseline activity subtracted. These data were then fit by a Gaussian having four free parameters, and the goodness of fit was evaluated by computing the r^2 value for each fit. Any neuron having one or more parameter values greater than 4 SDs from the group mean were excluded from subsequent analysis. For completeness, curve fitting was performed with Matlab’s curve fitting toolbox, using the nonlinear least squares method with a maximum of 600 function evaluations. The gain, width and baseline shift had a lower bound of zero and no upper bound, and the mean parameter was bound to the size of the receptive field [-1, 1].

For statistical power and to demonstrate the robustness of the general architecture, 100 different simulated monkeys were used and the neural responses to probe stimuli were recorded with attention directed at each of the three targets. In this analysis, the methods

Table 6.2: Quantitative results - Aggregate data from 100 simulated monkeys in the ARC compared with experimental data from two monkeys reported in Womelsdorf et al. (2008). Mean indicates the average change in the receptive field property measured, and 95% CI Low and High represent the lower and upper 95% confidence intervals. The Womelsdorf et al. (2008) study reports standard error of the mean (SE%), from which lower and upper 95% confidence intervals were estimated.

		ARC			Womelsdorf et al. (2008)			
		Mean (%)	95% CI Low	95% CI High	Mean (%)	SE (%)	95% CI Low	95% CI High
Gain	Entire sample	-0.343	-1.485	0.797	4.1	4.3	-4.328	12.528
	Selected pairs	-0.583	-1.733	0.577	5.0	3.0	-0.88	10.88
Shift	Entire sample	33.231	32.864	33.620	31.4	2.8	25.912	36.888
	Selected pairs	33.407	33.037	33.776	25.3	4.6	16.284	34.316
Shrink	Entire sample	-13.567	-14.425	-12.694	-12.1	1.9	-15.824	-8.376
	Selected pairs	-16.193	-17.190	-15.183	-11.2	2.3	-15.708	-6.692

of Womelsdorf et al. [272] were used to investigate three effects on receptive field properties when attention is directed toward a stimulus inside the receptive field compared with a stimulus outside the receptive field: a change in the peak firing rate (gain of the Gaussian fits); a shift of the receptive field centre (mean of the Gaussian fit); and a change in the receptive field size (standard deviation of the Gaussian fit). These data were then compared with the reported experimental results. For all statistical analyses, the data were bootstrapped ($n = 3000$) for amplitude, shift and shrink using the values from the Gaussian fits, both for individual neurons and for the aggregate values of each simulated monkey. The bootstrapped data were then used to calculate the 95% confidence interval for each test. Table 6.1 summarizes the changes in the three attentional effects for the simulation and experimental data, and quantitative results are summarized in Table 6.2.

Across the 100 simulated monkeys, the average number of neurons discarded in each monkey was 30.34 of 100 neurons, with a 95% confidence interval of [29.41, 31.10], which is larger than the 20.2% discarded in Womelsdorf et al. [272]. However, 65.54% of the discarded cells were excluded because they exhibited little or no activity in response to the reference and probe stimuli as a result of them having had x -intercepts near -1.0 or 1.0 rather than as a result of the fitting process. Combining all neurons from all simulated monkeys, the median r^2 value of the fit was 96.43%. Figure 6.3 shows the distribution of r^2 values for all neurons (blue bars) and selected pairs (red bars) at each attentional focus. Following Womelsdorf et al. [272], those neurons having an r^2 value from the Gaussian fit greater than the median for each of the three attentional targets are referred to as “selected pairs.”

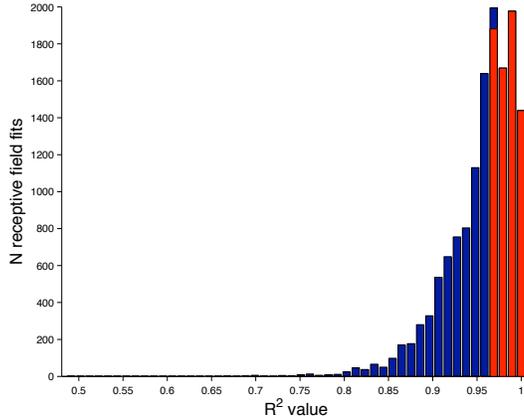


Figure 6.3: r^2 values for individual neuron receptive fields fit with a Gaussian in different attentional conditions for all simulated monkeys. Red bars indicate neurons with above median r^2 values (“selected pairs”), and blue bars indicate the fits of the entire sample.

6.1.1 Changes in Receptive Field Amplitude

For the change in receptive field amplitude, a significant decrease in activity was observed in the attend-in condition compared with attend-out ($p < 0.001$, one sample t-test, $n = 13932$) when all neuron pairs from the 100 simulated monkeys were considered (Figure 6.4(a)). The change in peak activity was calculated as $\frac{in-out}{out}$. This change had a mean decrease of -0.359% (CI = [-0.597%, -0.132%]). For just the selected pairs, there was also a significant decrease in peak activity ($p < 0.001$, one sample t-test, $n = 6914$), with a mean decrease of -0.597% (CI=[-0.943%, -0.255%]).

To test whether the data were normally distributed, the Jarque-Berra test was applied to the data. The distribution of activity changes of the individual neurons is highly non-normal for both the selected pairs and entire sample ($p < 0.001$ Jarque-Berra test), so an additional analysis was performed to examine the mean amplitude changes from each simulated monkey. Figure 6.5 (blue bars) shows a histogram of the aggregated data composed of the mean change in peak activity between the attend-out and attend-in conditions for the ensemble of neurons recorded in each of the 100 simulated monkeys. A non-significant decrease in peak activity was observed ($p > 0.001$, one sample t-test, $n = 200$), with an average decrease of -0.343% (CI=[-1.485%, 0.797%]). In this test, the $n=200$ corresponds to 100 measurements from when attention was directed to S1 versus attend-out, and 100 from attend-S2 versus attend-out. An estimate of the upper and lower confidence interval of the experimental data (R_u and R_l) can be obtained as $R_l = R_{mean} - (1.96 * SE)$ and $R_u = R_{mean} + (1.96 * SE)$. The simulation data is within the 95% confidence interval of the mean reported in Womelsdorf et al. [272], which found an average change of $4.7\% \pm 4\%SE$,

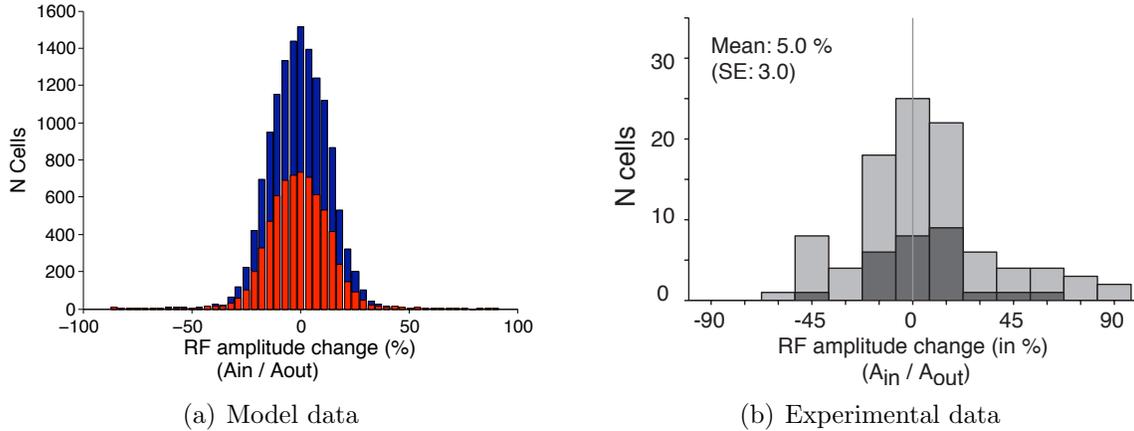


Figure 6.4: Change in receptive field amplitude of the Gaussian fits between the attend-out and attend-in conditions. (a) Simulation data for neurons from 100 simulated monkeys. Blue bars indicate the change in the entire sample of neuron pairs (mean = -0.359% , $95\% \text{ CI} = [-0.597\%, -0.132\%]$). For the selected pairs (red bars), the mean change was -0.597% ($95\% \text{ CI} = [-0.943\%, -0.255\%]$). (b) Experimental data reported by Womelsdorf et al. (2008). For the entire sample (light gray bars), the mean change was 4.7% ($95\% \text{ CI} = [-4.33\%, 12.54\%]$). For the selected pairs (dark gray bars), the mean change was 5% ($95\% \text{ CI} = [-0.88\%, 10.88\%]$). (Adapted from [272]).

corresponding to a $95\% \text{ CI}$ of $[-4.328\%, 12.538\%]$. Similarly, for just the selected pairs (Figure 6.5, red bars) there was a non-significant decrease ($p > 0.001$, one sample t-test, $n = 200$), with a mean of -0.583% ($\text{CI} = [-1.733\%, 0.577\%]$), which is consistent with the results of Womelsdorf et al. [272] ($5\% \pm 3.0\% \text{ SE}$), corresponding to a $95\% \text{ CI}$ of $[-0.88\%, 10.88\%]$ (Figure 6.4(b)). The slight reduction in amplitude can be attributed to the influence of the Gaussian connection strengths over the receptive field (σ_ω). These feedforward weights have the effect of moderately attenuating inputs from the edges of the layer-IV neurons' receptive fields.

Comparing the experimental data Figure 6.4(b) with the simulation data for all neurons in all simulated monkeys (Figure 6.4(a)) and with the distribution of the average change in each simulated monkey (Figure 6.5) illustrates that the simulation data closely matches the experimental data. These figures show that the model captures the overall effect that the receptive field amplitude does not change. Not surprisingly, by conducting the simulations with more neurons and simulated monkeys, the distribution of changes in receptive field amplitude is considerably more smooth than the experimental data.

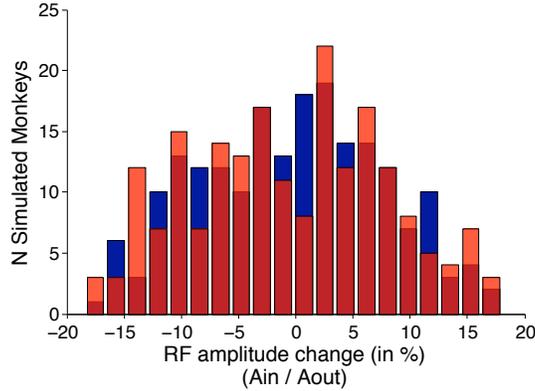


Figure 6.5: Average population change in receptive field amplitude from 100 simulated monkeys between attend-out and attend-in conditions. There was not a significant change between conditions, with the average change for the entire sample (blue bars) being -0.343% (95% CI=[-1.485%, 0.797%]), and for selected pairs (red bars) = -0.583% (95% CI = [-1.733%, 0.577%]).

6.1.2 Changes in Receptive Field Position

For each neuron, the Gaussian fit receptive fields were compared between each of the attend-in conditions versus the attend-out condition to determine if their receptive fields had shifted, and by how much. The shift in receptive field with attention was calculated as $(x_R^{attended} - x_R)/(x_{Att} - x_R)$, where $x_R^{attended}$ represents the mean of the Gaussian fit with attention directed to a stimulus inside the receptive field, x_R is the mean of the Gaussian fit in the attend-out condition, and x_{Att} is the position of the control signal, μ . For individual neurons from all simulated monkeys, a significant shift in the receptive field centre toward the stimulus was observed ($p < 0.001$, one sample t-test, $n = 13932$). The average shift for all neuron pairs was 33.26%, with a 95% confidence interval of [33.197%, 33.325%] (Figure 6.6(a), blue bars). For the selected pairs (Figure 6.6(a), red bars), there was also a significant shift of the receptive field centre toward the attended stimulus ($p < 0.001$, one sample t-test, $n=6914$), with an average shift of 33.436% (CI=[33.347%, 33.532%]). As with the amplitude change, the data were not normally distributed ($p < 0.001$, Jarque-Berra test).

Figure 6.7 (blue bars) shows the average population change in receptive field position for the entire sample between attentional conditions for 100 simulated monkeys, where the average shift was 33.231%, with a 95% confidence interval of [32.864%, 33.620%]. This shift was statistically significant ($p < 0.001$, one sample t-test, $n = 200$), and the mean change of the simulation data is also within the 95% confidence interval of Womelsdorf et

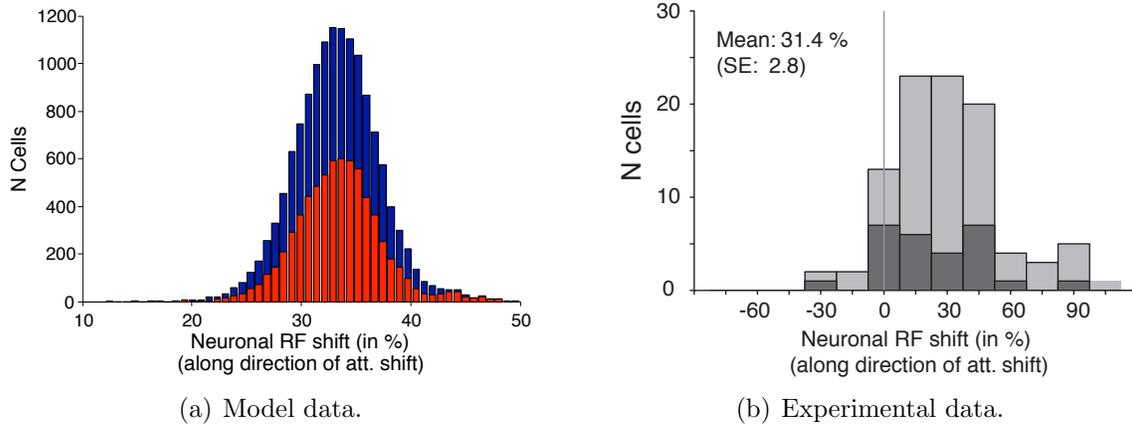


Figure 6.6: Shift of receptive field position between attentional conditions. (a) Simulation data from individual neurons of 100 simulated monkeys. For the entire sample (blue bars), the mean shift was 33.26% (CI=[33.197%, 33.325%]), and for the selected neurons (red bars), the mean shift was 33.436% (CI = [33.347%, 33.532%]). (b) Experimental data reported by Womelsdorf et al. (2008). For the entire sample (light gray bars), the mean shift was 31.4% (95% CI = [25.01%, 36.89%]), while for the selected pairs (dark gray bars), the mean shift was 25.3% (95% CI = [16.28%, 34.32%]). (Adapted from [272]).

al. [272] ($31.4 \pm 2.8\%$ SE; 95% CI = [25.913%, 36.888%]) (Figure 6.6(b)). For the selected pairs from each simulated monkey (Figure 6.7, red bars), a significant shift of receptive field centres toward the target was observed across simulated monkeys ($p < 0.001$, one sample t-test, $n = 200$), with an average of 33.407% (95% confidence interval=[33.037%, 33.776%]). This shift is also within the 95% confidence interval of the mean found in Womelsdorf et al. [272] ($25.3\% \pm 4.6\%$ SE), which corresponds to a 95% CI of [16.284%, 34.316%].

As with receptive field amplitude, the significantly larger number of neurons in the simulation data produces distributions of changes (Figures 6.6(a), 6.7) that are considerably smoother than the experimental data (Figure 6.6(b)), while still demonstrating a consistent average change in receptive field size.

6.1.3 Changes in Receptive Field Size

Finally, the changes in the receptive field size with attention were also quantified by measuring the change of the width of the Gaussian fits between the attend-out (σ_R) and attend-in ($\sigma_R^{Attended}$) conditions, as $(\sigma_R^{Attended}/\sigma_R) - 1$. Figure 6.8(a) shows that when using

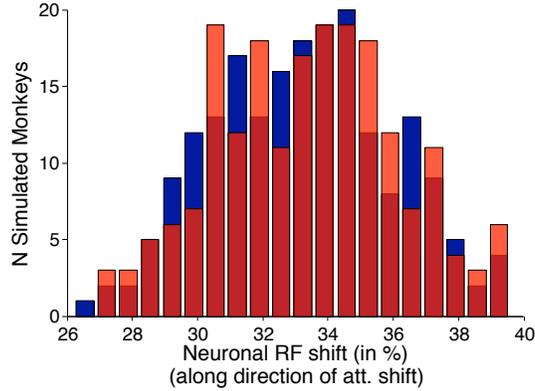


Figure 6.7: Average population receptive field shift between attend-out and attend-in conditions from 100 simulated monkeys. The average shift when using the entire sample (blue bars) was 33.231% (95% CI = [32.864%, 33.620%]), and for selected pairs (red bars) = 33.407% (95% confidence interval=[33.037%, 33.776%]).

the entire sample from all simulated monkeys (blue bars), there was a significant reduction in the size of the receptive field in the attend-in condition ($p < 0.001$, one sample t-test, $n = 13932$), with an average reduction of -13.615% (CI=[-13.875%, -13.382%]) for the entire sample. With just the selected pairs (Figure 6.8(a), red bars), there was a significant decrease in receptive field size ($p < 0.001$, one sample t-test, $n=6914$), with an average change of -16.237% (CI=[-16.660%, -15.806%]). As with amplitude change and shift, for both the selected pairs and entire sample, the distribution of receptive field size changes was non-normal ($p < 0.001$, Jarque-Berra test), so this effect was assessed using the mean change across the recorded neurons in each simulated monkey. Using the mean of the entire sample from each simulated monkey (Figure 6.9, blue bars), the receptive field size significantly decreased ($p < 0.001$, one sample t-test, $n=200$) on average by -13.567% (CI=[-14.425%, -12.694%]). As with the previous two attentional effects, these results are consistent with the experimental data, which report an average change in receptive field size with the entire sample of $-12.1\% \pm 1.9\%$ SE, corresponding to a 95% CI of [-15.824%, -8.376%] (Figure 6.8(b)). The attentional effect on receptive field size was similar when using the mean change of the selected pairs across all simulated monkeys (Figure 6.9, red bars), where there was a significant decrease in receptive field size ($p < 0.001$, one sample t-test), with an average change of -16.193% (CI=[-17.190%, -15.183%]). This effect is also consistent with the experimental observation of $-11.2\% \pm 2.3\%$ SE (95% CI=[-15.708%, -6.692%]).

Again, the distribution of changes in receptive field size is considerably smoother in the simulation data with ~ 14000 neurons (Figure 6.8(a)) and with the population average

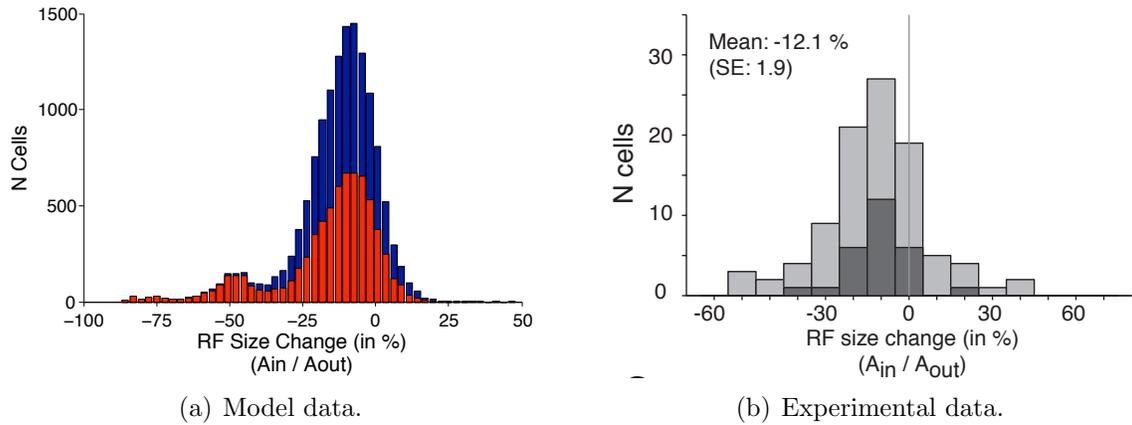


Figure 6.8: Changes of receptive field sizes between attentional conditions. (a) Simulation data for individual neurons from 100 simulated monkeys. For the entire sample (blue bars), there was an average decrease of -13.615% (CI= $[-13.875\%, -13.382\%]$), and for the selected pairs (red bars), the average was -16.237% (CI = $[-16.660\%, -15.806\%]$). (b) Experimental data by Womelsdorf et al. (2008). For the entire sample (light gray bars), the mean shrink was -12.1% (95% CI = $[-15.82\%, -8.38\%]$) and for the selected pairs, the average shrink was -11.2% (95% CI = $[-15.71\%, -6.69\%]$). (Adapted from [272]).

from 100 simulated monkeys (Figure 6.9) than in the experimental data (Figure 6.8(b)), although the simulation data are still consistent with the experimental reports.

6.1.4 Results Summary

Figure 6.10 summarizes the changes in receptive field properties from the simulation data (dashed lines) and experimental data (solid lines) for both the selected pairs (red lines) and entire sample (blue lines). The mean change of each attentional effect is shown with an asterisk and error bars indicate the 95% confidence intervals of the mean.

As shown in Figure 6.10, across these three measurements, there is no statistical difference between the simulation data and the experimental data. However, this figure also illustrates that the distribution of each attentional effect is much narrower in the simulation data than in the experimental data, and there are several possible causes of this effect. First, the simulation data are from recordings of ~ 14000 neurons across 100 simulated monkeys, whereas the experimental data recorded from 78 neurons in two monkeys ($n = 57$ and $n = 21$ for the two monkeys). As the lines in Figure 6.10 are for the 95% confidence intervals and not standard deviation, it is expected that the confidence intervals would shrink with increasingly large datasets. The significant variance in the experimental

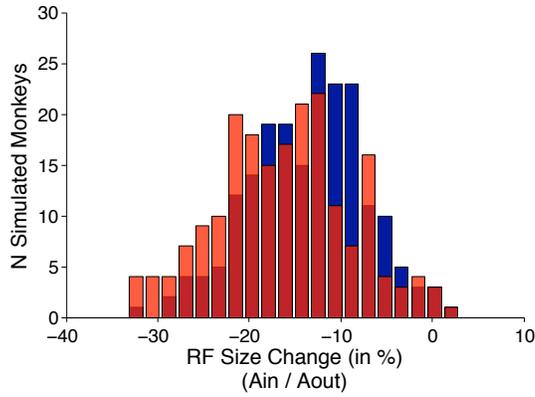


Figure 6.9: Average population change in receptive field size between attend-out and attend-in conditions from 100 simulated monkeys. Average size change for the entire sample (blue bars) = -13.567% (CI=[-14.425%,-12.694%]), and for selected pairs (red bars) = -16.193% (CI=[-17.190%, -15.183%]).

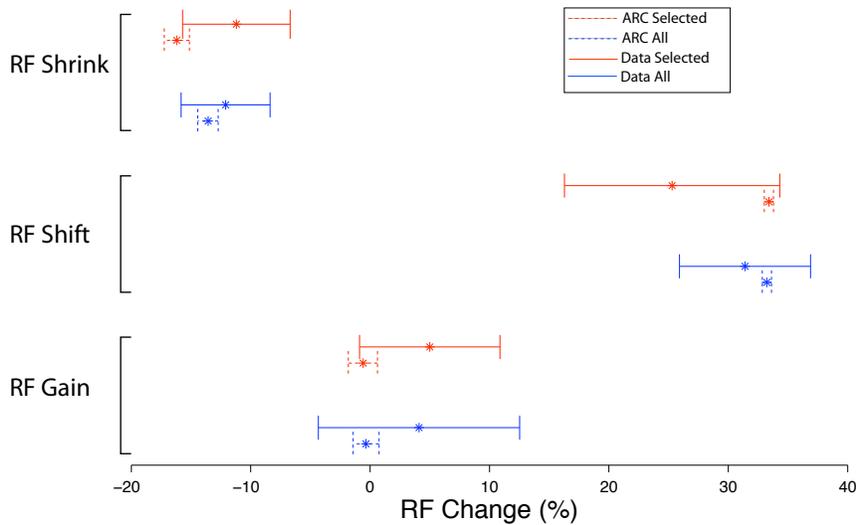


Figure 6.10: Summary of attentional effects (mean and 95% confidence intervals) for 100 simulated monkeys (dashed lines) and results reported by Womelsdorf et al. (2008) (solid lines). Blue bars are from the entire sample, and red bars are selected pairs. See text for details.

data may also arise from measurement noise, or the myriad forms of noise found in real brains, such as changes in the animal's alertness and cognitive state, stimulus processing effects, or subtle movements in the recorded cell's position.

The Womelsdorf et al. [272] experiment examined the influence of attention on the spatial position, size and amplitude of receptive fields when attention is shifted between targets inside and outside the receptive field. Considering each effect of attention on receptive field profiles in the context of the model presented in Chapter 4, these forms of modulation are consistent with the model's principles. The first effect reported in Womelsdorf et al. [272] is that the amplitude of the receptive field does not significantly change between attentional conditions, an effect that was reproduced by the ARC. Considering the routing mechanisms of the ARC, this effect is predicted by the model. The local control signal μ can have values anywhere within the receptive field, and for each of these positions, the peak of the routing function has a constant value of 1, meaning that since μ was centred on each target, the visual signal from each stimulus was similarly multiplied by 1.

The simulation data showed a slight, but non-significant reduction in the receptive field amplitude when attention was directed to stimuli within the receptive field. This reduction is presumably due to the Gaussian shaped connection strengths between the V1 columns and the MT column. For simplicity, the connection strengths had a Gaussian profile (Equation 6.1) with $\mu_\omega = 0$ and $\sigma_\omega = 1$. If a smaller value for σ_ω were to be used instead, this would cause the inputs from columns at the edges of the receptive field to be more strongly attenuated. With a smaller value for σ_ω , the model predicts that directing attention to a stimulus at the receptive field edge will result in the amplitude of the receptive field fits decreasing, and thus a reduction in gain. This observation also suggests further details of this attentional effect, and predicts that if the two receptive field stimuli are placed closer to the edges of the receptive field, the receptive field gain will decrease when attending either receptive field target as compared with the attend-out condition.

The second effect reported is that when attending to a receptive field stimulus, the receptive fields shift toward the attended target, and that the position of the shifted receptive field varies with the position of the attentional target. The ARC proposes that selective information routing from a particular region of the receptive field results from the routing function attenuating input signals from parts of the receptive field that are distant from the attentional target for neurons in that column (μ). This position is computed by control neurons in each column, based on the global size and position of the attentional target. Consequently, changes in the target's position cause changes in the global control signals, which results in μ changing in each column when the target is within the receptive field. This principle is seen in the receptive fields shifting with different attentional targets.

Comparing the simulation data for receptive field shift with the experimental data

(Figure 6.6), there is a subset of neurons reported experimentally having characteristics not found in the simulation data. Specifically, a minority of cells in the experimental data were found to have receptive fields that shifted away from the target, producing a negative shift, while all of the simulated neurons only showed positive shifts. However, a different study of receptive field modulation with spatial attention in V4 found that 46% of the recorded cells had receptive fields that significantly shifted in the direction of attention, although none of those cells showed a shift in the opposite direction [36]. This suggests that it may be worth repeating this experiment with additional cells to see if these “away” shifts are reproducible.

The third effect reported in Womelsdorf et al. [272] is that the size of the receptive field decreases when attention is directed to a stimulus inside the receptive field, an effect that was also reproduced by the ARC. As was discussed in Chapter 1, attention serves to selectively process attended information and to filter out extraneous information. In higher cortical levels in particular, the large receptive fields provide each neuron with information from a large spatial area, and much of this information may be irrelevant for the current task. The ARC proposes that columns that are not encoding visual information from the attentional target operate in their default routing state, and in this state, the width of the routing function is larger than when neurons are in a selective routing state. The ARC predicts that the reduction in receptive field size observed when attention is directed to a receptive field stimulus results from these neurons switching from their default routing state to one in which the target is being selectively processed.

Overall, considering all three effects together, the width of the routing function was set to 1 for the attend-out condition and $\sigma_{att} = 0.75$ for the attend-in conditions. These values were selected to ensure that for the attend-in condition, the value was lower than for the attend-out condition on a normalized range. With these values, the simulation data were statistically indistinguishable from the experimental data, for all three measurements. Given the considerable variance of the mean in the experimental data, it is plausible that there exist several parameter regimes in which the model and experimental data will be consistent, and perhaps more closely match the mean changes reported experimentally. However, until further data are available to constrain these values, there is no advantage in doing so. Rather, the model’s ability to quantitatively match the experimental results without parameter fitting demonstrates the robustness of the model to reproduce these effects.

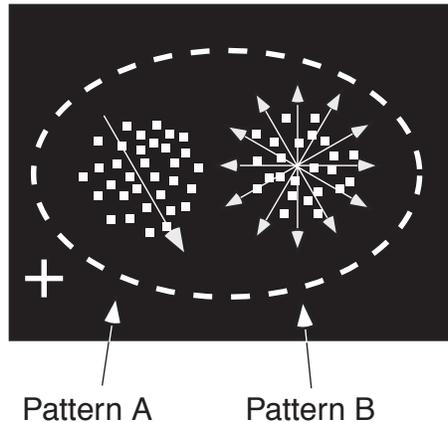


Figure 6.11: Stimulus design used in the Treue and Martinez-Trujillo experiment. The monkey was required to maintain fixation at the fixation point (cross in lower left) while covertly attending to one of two random dot patterns in the receptive field. Pattern A moved in the cell’s anti-preferred direction, while Pattern B could move in one of 12 directions on a given trial. (Reproduced from [245]).

6.2 Simulations of Treue and Martinez-Trujillo (1999)

In the previous section, it was shown that the ARC can account for modulation of spatial receptive fields when attention is shifted between targets within a receptive field. The simulation presented in this section address the effects of attention on feature tuning with shifts of attention. Treue and Martinez-Trujillo [245] recorded from 56 neurons in area MT of two macaque monkeys while the animals covertly attended to one of two stimuli inside the recorded cell’s receptive field or to a stimulus placed at the fixation point. The stimuli used were two random dot patterns, one moving in the cell’s anti-preferred direction (Pattern A), and the other (Pattern B) moving in one of 12 directions (Figure 6.11). On each trial, either stimulus was designated the target by the presentation of a stationary random dot pattern at that location. After the cue was extinguished, the two stimuli were presented and the monkey was to release a lever when the target stimulus changed speed or direction, while ignoring any changes of the distractor. The target could change anywhere between 270 and 4000 ms following stimulus onset, and only trials that were correctly completed without eye movements were subject to analysis. Firing rates were then determined based on the average activity in the 1000 ms following target movement, with the first 200 ms discarded.

To construct tuning curves for each cell, activity was recorded for each of 12 directions that the stimulus moved, with attention directed to Pattern A, B or to fixation. For

each attentional condition, the 12 firing rates were fit with a Gaussian having four free parameters: r_{null} , the cell’s response to the anti-preferred direction, $dirGain$, the maximal response modulation, $prefDir$, the cell’s preferred direction, and $width$, the range of direction to which the neuron responds:

$$r_{null} + dirGain \times \exp(-0.5 \frac{(dir - prefDir)^2}{width^2}). \quad (6.2)$$

They found that attending to the variable motion stimulus (Pattern B) resulted in an increase in response gain ($dirGain$), while attending to Pattern A produced a reduction in gain, with an intermediate response occurring when attending the neutral stimulus at fixation. Further, the width of these curves did not significantly change between attentional conditions, but rather the responses were multiplicatively scaled similarly for all motion directions.

This latter finding is of particular significance for distinguishing between possible mechanisms of attentional routing. Specifically, the biased competition [188], normalization model [189] and ARC all suggest that directing attention to a stimulus will increase its contribution to the neuron’s activity while diminishing the effect of the unattended stimulus. That is, the cell’s response will be more strongly influenced by the properties of the attended stimulus, with the unattended stimulus contributing less. However, both the biased competition and normalization models differ from the ARC by predicting that the range of motion directions to which a neuron responds will increase when attending the anti-preferred stimulus, with an overall decrease in gain. This is seen as an increased width of the Gaussian fit when attending to Pattern A as compared with attending to the variable motion Pattern B. In other words, attending to the variable direction or preferred stimulus will produce a sharpening of the tuning curve. Conversely, the ARC predicts that the range of directions (i.e. the width of the Gaussian fit) will not change when attention is directed to either stimulus, or to the neutral stimulus. For clarity, note that although the spatial width of the receptive field may change with shifts of attention as was shown in the simulations of Womelsdorf et al. [272], the range of features to which the neurons respond should not.

This prediction was tested by simulating the experiment of Treue and Martinez-Trujillo [245], using a similar model to that used in the simulations of the Womelsdorf task (Figure 6.12). 100 simulated monkeys were used, in each of which nine V1 columns provide visual input to a single MT column. The V1 columns each contain 100 layer II/III neurons, and the MT column contains 9×50 layer-IV pyramidal cells, 100 layer-VI control neurons and 100 layer II/III neurons from which recordings were made.

As with the Womelsdorf et al. simulations, the connectivity of the afferents to the MT column was organized such that each V1 column projects to 50 layer-IV MT cells, each having 30 nonlinear dendritic subunits that compute the routing function (Equation 4.7).

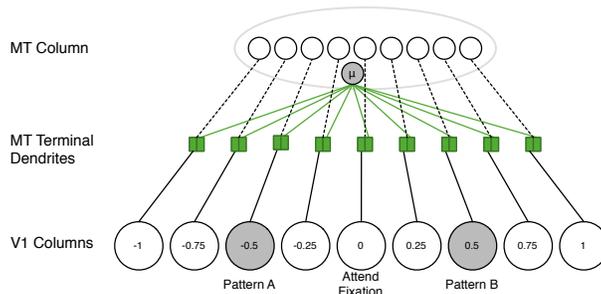


Figure 6.12: Spiking neuron ARC used in the simulations of Treue and Martinez-Trujillo [245]. Pattern A is at $x_{-0.5}$ and Pattern B at $x_{0.5}$. When attention is directed to the fixation point, the network operates in its default routing state with attention at x_0 .

Further, the width of the Gaussian connection strengths was set to $\sigma_\omega = 1$, and the width of the routing function was set to $\sigma_{att} = 1$ for the attend-fixation condition, and $\sigma_{att}=0.75$ when attention was directed to a stimulus inside the receptive field. As in the previous simulation, the value of $\sigma_{att} = 0.75$ was selected not by parameter fitting, but as being some value less than that used in the default state.

The three attentional conditions were modelled by setting the attentional target for Pattern A as $\mu = -0.5$ and Pattern B as $\mu = 0.5$. The neutral condition reflects these neurons being in their default routing state, and for simplicity, the control signal in this state was specified as $\mu=0$, as in the previous simulation. Any neuron that did not respond to two or more motion directions was excluded from analysis. Following Treue and Martinez-Trujillo [245], in each attentional condition, each stimulus direction was presented for 1000 ms, with the first 200 ms being excluded from analysis.

To simplify the analysis, all layer II/III neurons were modelled as “on” neurons, such that they respond more strongly to increasingly large positive values. Direction of motion was modelled by having the V1 neurons project a scalar value in the range $x_{V1} \in [-1, 1]$, corresponding to directions $[-180, 180]$. As the tuning curves for motion direction in MT cells can be well characterized as a Gaussian over the range of directions [8], this is modelled by transforming the direction signals so that the maximum response obtained for a direction of 0° corresponds to the neuron receiving an input value of 1.0. The minimum response to -180° or 180° thus corresponds to an input value of -1.0. This transformation is given by:

$$x_{MT} = 2 \times \exp\left(-\frac{(x_{V1}-0)^2}{2\sigma^2}\right) - 1, \quad (6.3)$$

Thus, Pattern A had a value of -1.0 corresponding to a motion direction of $\pm 180^\circ$ and Pattern B had 12 equally spaced directions in the range $[-1.0, 1.0]$, corresponding to [-

180°, 180°]. This can also be modelled in the NEF by using neurons with two dimensional tuning for direction, although in order to have the stimuli represented with the same accuracy as in the 1D neurons used here, the number of neurons encoding the input signals would need to be increased by a factor of approximately 2. Thus, the present simulations use neurons sensitive to the transformed signals in one dimension, as it produces the same functional effect, while requiring significantly less computational resources.

Although in cortex it would be expected to find neurons responsive to all directions of motion, rather than all neurons having a preferred direction of 0°, this arrangement is equivalent to using neurons tuned to a variety of directions and aligning their preferred directions to 0 before conducting the analysis. Further, to ensure that the recorded neurons were responsive to a range of directions, the x intercepts of their tuning curves (i.e. the input value at which they begin spiking) were selected from a uniform distribution over the range [-1.0, -0.5], corresponding to [-180°, -90°] or [90°,180°]. Maximum firing rates of layer II/III neurons were drawn from a uniform distribution over the range [90, 120] Hz, with refractory periods of 5 ms, and a membrane time constant of 20 ms. These choices were made to match the responses observed in the experiment.

Following Treue and Martinez-Trujillo [245], the change in gain and width when attention is directed to Pattern B is measured as an attentional index:

$$Ai(dirGain) = \frac{(dirGain_b - dirGain_a)}{(dirGain_b + dirGain_a)} \quad (6.4)$$

$$Ai(width) = \frac{(width_b - width_a)}{(width_b + width_a)} \quad (6.5)$$

where X_a and X_b indicate attention being directed to Pattern A or B respectively.

Figure 6.13 shows the responses of a single cell in each of the three attentional conditions. The top curve is the neuron's responses when attention was directed to the variable motion stimulus (Pattern B), the middle curve is the activity when attention was directed to a target outside the receptive field ("sensory" condition), and the lower curve is with attention directed to the anti-preferred stimulus (Pattern A). For this cell, the r^2 value of the fits are 0.99, 0.97, and 0.98 with attention directed to Pattern B, fixation and Pattern A respectively. Across the simulated monkeys and attentional conditions, the average r^2 value was 0.9847, with a 95% confidence interval of [0.9640, 0.9953].

Figure 6.14 shows a histogram of the attentional index for gain from 85 cells in a single simulated monkey, as calculated using the attentional index (Equation 6.4). When attention was directed to Pattern B, there was a significant increase in gain across the population ($p < 0.001$), with an average attentional index of 0.171 (95% CI=[0.164, 0.177]), which corresponds to an average increase in directional gain of 141.5% (95% CI=[139.6%, 143.4%]). Figure 6.15 shows a histogram of the attentional modulation of tuning width

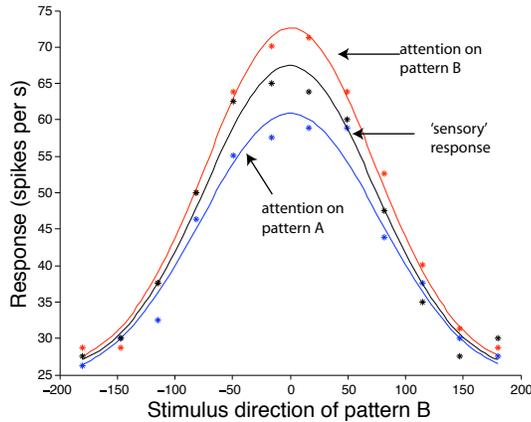


Figure 6.13: Responses of an example cell with attention directed to the anti-preferred stimulus (Pattern A, red), fixation (“sensory” response, black), or to the variable motion stimulus (Pattern B, blue). Asterisks indicate neural responses to each direction of motion and the curves are the best fitting function for responses in each attentional condition.

for the same simulated monkey as shown in Figure 6.14. For this single simulated monkey, there was no significant change in the width between conditions ($p > 0.001$), with an average AI of 0.000 (95% CI=[-0.005, 0.006]), which corresponds to an average ratio of widths for Patterns B and A of 1.002%, (95% CI=[99.1%, 101.4%]).

Grouping the data from all neurons in all simulated monkeys, the attentional modulation of gain (Figure 6.16(a)) showed a significant increase between attentional conditions ($p < 0.001$), with an average attentional index of 0.1564 (95% confidence interval = [0.1558, 0.1570]). This corresponds to an increase in the gain term of 137.39% (95% confidence interval = [137.31%, 137.57%]) when attention was directed to the variable direction stimulus.

As with the simulations of the Womelsdorf et al. [272] experiment, the mean change across the population of cells in each simulated monkey were combined, and the data were subjected to the same analysis as when combining individual cells from all monkeys. Figure 6.16(c) shows the distribution of changes in the attentional index for gain aggregated across simulated monkeys. Again, there was a significant change in the attentional index for gain ($p > 0.001$), with a mean of 0.1564 (95% confidence interval = [0.1529, 0.1597], corresponding to an average change in the ratio of 137.39% (95% confidence interval = [136.41%, 138.32%]). Figures 6.16(e) and 6.16(f) show the distribution of changes reported in Treue and Martinez-Trujillo [245] for gain and width respectively.

Figure 6.16(b) shows a histogram of the attentional modulation of the width of the Gaussian fit when combining all cells from all simulated monkeys. Again, the width of the

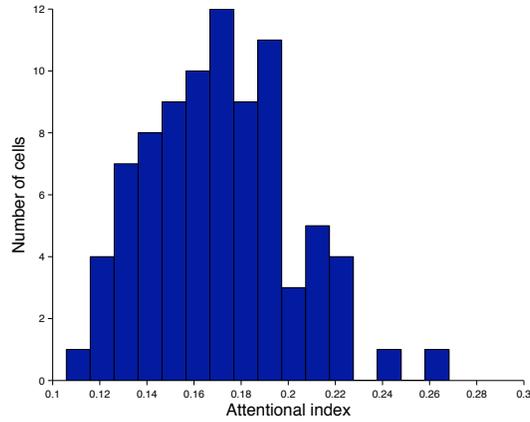


Figure 6.14: Change in the gain attentional index for a single simulated monkey. On average, the change in gain was 0.171 or a geometric mean of 141.5%, with 95% confidence intervals of [0.164, 0.177] and [139.6%,143.4%].

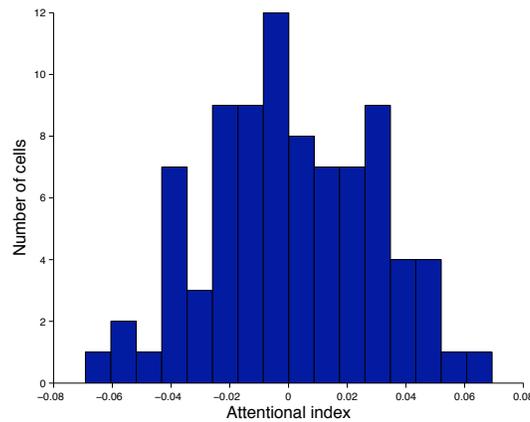
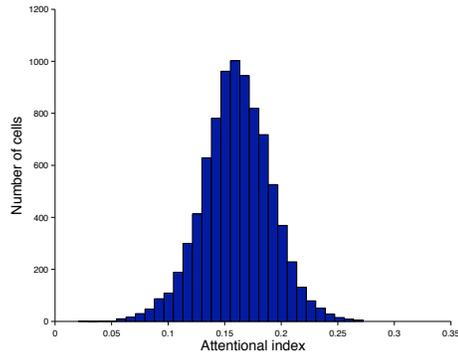
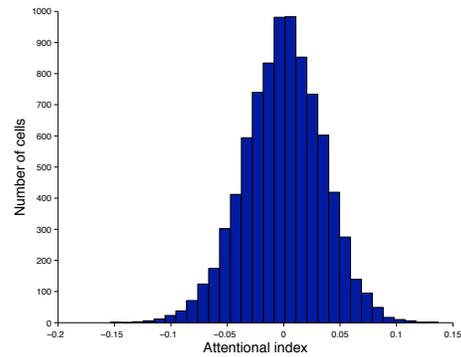


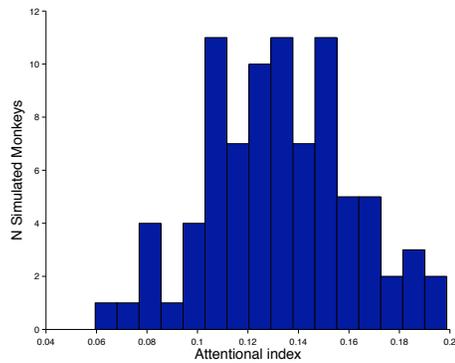
Figure 6.15: Change in the width attentional index for a single simulated monkey, with an average attentional index of 0.000, corresponding to 0.2%, with 95% confidence intervals of [-0.005, 0.006] and [99.1%, 101.4%].



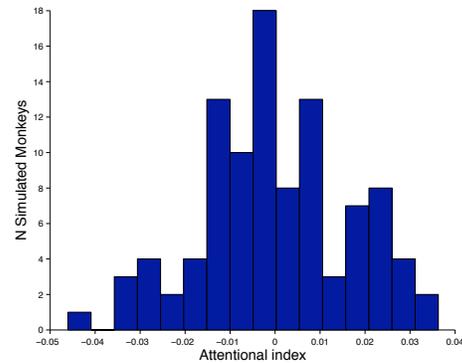
(a) $AI(dirGain)$ for all cells.



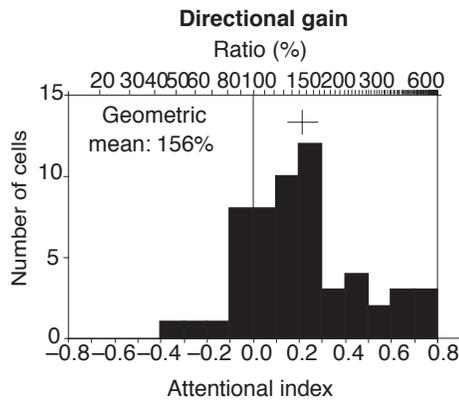
(b) $AI(width)$ for all cells.



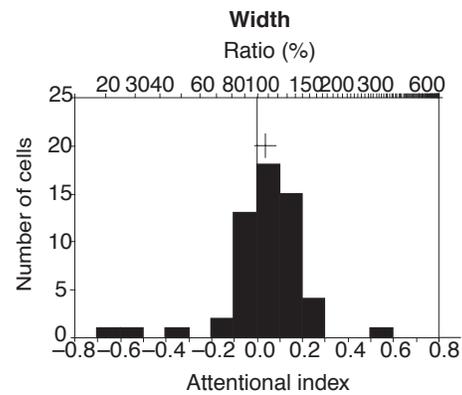
(c) $AI(dirGain)$ aggregated across simulated monkeys.



(d) $AI(width)$ aggregated across simulated monkeys.



(e) Experimental $AI(dirGain)$.



(f) Experimental $AI(width)$.

Figure 6.16: Summary of attentional index values for gain and width. (a) and (b) combine all neurons from all simulated monkeys, (c) and (d) are the average change in each simulated monkey, and (e) and (f) are experimental results reproduced from [245]. The simulation data shown in (a) and (c) and experimental data (e) show that the receptive field gain significantly increases when attending the variable motion stimulus, while the width does not change in the model (b, d) or experimental data (f).

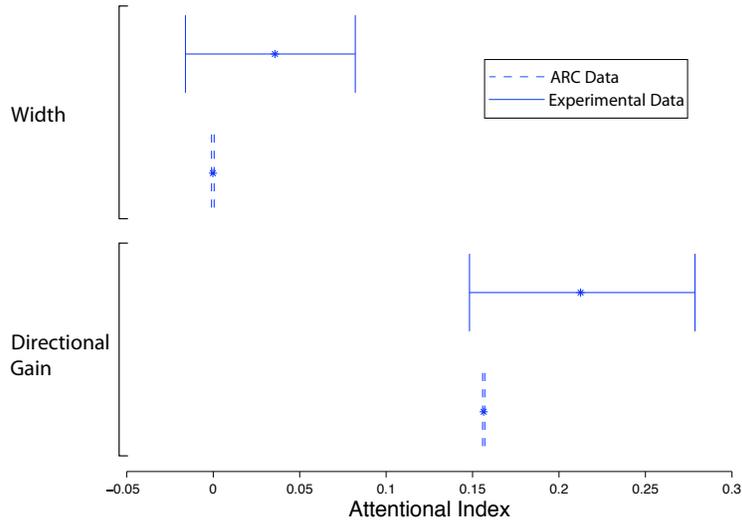


Figure 6.17: Comparison of attentional effects reported by Treue and Martinez-Trujillo (solid lines) and from simulations using the ARC (dashed lines).

Gaussian fit did not change significantly between attentional conditions ($p > 0.001$). The average AI for width across all cells was -0.00018 with a 95% CI of [-0.00094, 0.00053], corresponding to an average ratio of 0.2% (95% confidence interval = [0.05%, 0.3%]).

The corresponding change in the attentional index for width across the 100 simulated monkeys is shown in Figure 6.16(d). As was the case when combining all cells from all simulated monkeys (Figure 6.16(b)), there was not a significant change ($p > 0.001$), with an average attentional index of -0.00018 (95% confidence interval = [-0.0033, 0.0031]), corresponding to a mean change of 100.2% (95% confidence interval = [99.58%, 100.86%]).

From the original experiments, the authors report few details of the data, indicating only the geometric mean of the directional gain being 156% and a non-significant change in width of 8%. However, from the histograms in the original article, (Figure 4 in [245]), confidence intervals of the mean can be estimated by taking the number of cells in each histogram bin, and assigning to these points, the value corresponding to the centre of the histogram bin. This estimated data set was then bootstrapped ($n=3000$) and the bootstrapped data were used to estimate their confidence intervals. For directional gain, this yields an average AI of 0.2125, with an estimated 95% confidence interval of [0.1518, 0.2768], and an average AI for width of 0.0357 (CI=[-0.0179, 0.0821]).

6.2.1 Results Summary

Figure 6.17 summarizes the attentional effects from the simulation data and the experimental data. For each measurement, the confidence intervals of the simulation and experimental data overlap, indicating that there is not a significant difference between the simulation and experimental data. However, it is useful to further assess the consistency of the model with the experimental data by determining the amount by which the model may be incorrect, in the worst-case scenario given the available information. To determine this value, the maximum likely difference, or equivalence test threshold is used [229, 249]:

$$MLD = \max(M_u - R_l, R_u - M_l), \quad (6.6)$$

where M_u and M_l are the upper and lower 95% confidence intervals of the model, and R_u and R_l are the “real-world” or experimental confidence intervals. The MLD specifies the smallest threshold θ for which the null hypothesis that $|R - M| > \theta$ can be reliably rejected, and is defined in the same units as the confidence intervals (attentional index). For the attentional index for width, this measurement yields an MLD value of 0.0457, while for gain, MLD=0.0619; that is, given the available data, in the worst case, the difference between the attentional index produced by the model and that of the experimental data is 0.0457 for width and 0.0619 for gain.

Similar to the comparative summary presented for the previous simulation (Figure 6.10), Figure 6.17 shows that the confidence intervals of the simulation data overlap with the experimental data, but are significantly narrower. In these simulations, recordings were made from 100 neurons in 100 simulated monkeys, while Treue and Martinez-Trujillo [245] recorded from 56 neurons in two monkeys.

The experiment by Treue and Martinez-Trujillo [245] examined two effects when attention was switched between the fixation point, the anti-preferred and variable motion stimulus. The first effect reported is that attending to the variable motion stimulus produced stronger responses (i.e. increased directional gain) than when attending the fixation point or anti-preferred stimulus, an effect that was reproduced by the ARC. The reason this effect is expected in the model is as follows.

In the ARC, the signal that is encoded by the layer-II/III neurons is a weighted sum of the signals from the anti-preferred and variable motion stimuli. The routing function, being centred on the receptive field in the attend-fixation condition, attenuates the gain of each stimulus by a similar amount, since the stimuli were placed symmetrically in the receptive field. Directing attention to either of the stimuli inside the receptive field both shifts the centre of the routing function (μ) toward the target, and causes the model to switch from the default routing state, thereby decreasing the width of the routing function σ_{att} . When attending the anti-preferred stimulus, which on its own produces little activity

in the recorded neurons, the combined effect of shifting μ away from the variable motion stimulus, and reducing σ_{att} , causes the signals from the variable motion stimulus to be more attenuated. This results in the anti-preferred stimulus contributing more to the activity of the recorded cells, thereby reducing the magnitude of the encoded visual signal, and in turn the firing rate. Conversely, attending to the variable motion stimulus results in the anti-preferred stimulus being attenuated by a similar amount as the preferred stimulus was in the previous condition. However, when attending the variable motion stimulus, which generally has directions closer to the preferred direction of the layer-II/III cells, the signals from that stimulus are located at the centre of the routing function, and pass through with little attenuation.

While the simulation data cannot be statistically distinguished from the experimental data, there is a subset of recorded neurons in the experimental data with characteristics not seen in the simulation data. Specifically, in the experimental data, 10 of the 56 cells showed a negative attentional index for directional gain (Figure 6.16(e)), indicating that the fit of these cells had a larger *dirGain* value when attending the anti-preferred stimulus than when attending the variable motion direction stimulus.

As shown in Figure 6.16(a), none of the neuron fits in the simulation data exhibited a reduction of the gain parameter as was found in the experimental data. This discrepancy may arise for several reasons. First, although the equation to which the neuronal responses were fit was given in the original article, there are numerous fitting methods that may be used (e.g. least-squares, weighted least-squares, robust regression, etc.), which may result in the data being fit differently. Further, the number of fit evaluations, bounds for the parameters, and fitting priority of each parameter can result in different fits, and thus parameter values, for the same data. Second, the fit values for the other parameters, particularly r_{null} which shifts the baseline of the curve, can have a significant impact on the values of the gain term (*dirGain*), and can obscure the peak gain of the curve when different fit methods or constraints are used. Finally, measurement noise, changes in the animal's cognitive state or the recorded neurons being involved in processing other information may all contribute to this effect. Whether the few neurons in the experimental data showing decreased gain parameters is an artifact of the fitting methods, or is due to another factor, remains unclear.

The second effect reported by Treue and Martinez-Trujillo [245] is that the range of directions to which the cell is responsive, does not change with attentional shifts, an effect that was also reproduced by the ARC. The ARC proposes that this effect arises because the multiplicative gain term computed in the routing function only depends on the spatial position being attended and similarly affects all feature values from each receptive field position. The routing function is defined over the neuron's spatial receptive field, and as long as the global control signals indicating the size and position of the target does not change, any visual signals from a given column will be modulated by the same amount,

independent of the contents of the visual signal. This property of the model is reflected in the simulation data (Figure 6.16(b), 6.16(d)) which show that the tuning curves do not change width, as all motion directions from a given spatial position are multiplied by the same gain term.

The results of these simulations provide support for the ARC, and provide the model with some explanatory advantage over other models that were discussed in Chapter 3. Specifically, the biased competition [188] and normalization model [189] predict that directing attention to the variable motion stimulus will produce a sharpening of the neuron's orientation tuning curve. However, the experimental data demonstrate that the tuning curve width is unaffected by directing spatial attention to different receptive field targets. In contrast, the ARC predicts that the multiplicative scaling of responses in each attentional condition is a result of the routing function attenuating visual signals from parts of the receptive field that are not being preferentially processed (i.e. columns that are distant from μ). The model predicts that the effect of visuospatial attention to a particular location is largely restricted to the spatial dimensions. That is, visuospatial attention affects the processing of information based on its retinal or receptive field location, and similarly affects all featural information from a given location. This prediction is further tested in the simulations presented in the following section, where the effects of attentional shifts on contrast-response functions are investigated.

6.3 Simulations of Lee and Maunsell (2010)

To this point, we have seen that directing attention to a stimulus in a neuron’s receptive field can influence the response of the cell as compared to the response elicited by directing attention to a stimulus outside the receptive field. Another means of testing attentional mechanisms has been pursued by several studies that have investigated the influence of attention on the contrast-response function, which describes the change in neuronal firing rates as contrast increases, to assess whether the attentional effect is better described as response gain or contrast gain. Due to the possible confounding factors present in several previous studies (see Section 1.3.2), Lee and Maunsell [112] sought to clarify this question in their recent work.

Their study recorded from 56 neurons in area MT of two macaque monkeys performing a speed-change detection task (Figure 6.18). Two Gabor stimuli were presented in the receptive field of each recorded neuron, one moving in the cell’s preferred direction, and the other in a non-preferred direction. The two Gabors had equal contrast, with the contrast value of each presentation randomly drawn from a set of 8 values [0, 1.6, 3.1, 6.2, 12.5, 25, 50, 100%]. The stimuli were placed at locations within the receptive field that were at equal eccentricity from fixation, and that gave strong and roughly equivalent responses. After signalling the location of the target, either with a location cue or by instruction trials, the animal’s task was to detect the appearance of a Gabor at the cued location having a slightly faster drift speed, while ignoring distractor Gabors at the non-cued location having a faster drift speed than the target. When the change was detected, the animal then responded by making a saccade to the target, and was rewarded with a drop of liquid if the response was made within 600 ms of the change.

In their experiment, neurons were held for an average of 36 repetitions of each contrast value and attention condition, and the response was measured as the average activity between 50 and 250 ms following stimulus onset. The mean responses across repetitions from each attentional condition and stimulus were first fit with a hyperbolic ratio function [3]:

$$R = R_{max} \left(\frac{c^n}{c^n + c_{50}^n} \right) + m, \quad (6.7)$$

where R_{max} is the maximum firing rate, c is contrast, n is the steepness of the function, and m is the baseline activity. To further assess whether the attentional modulation was more consistent with a response gain or contrast gain, the mean responses in each attentional condition were fit with modified contrast response functions (Equation 6.7) that model either a pure response gain or pure contrast gain.

The response gain model is given by incorporating an additional term, a , that acts

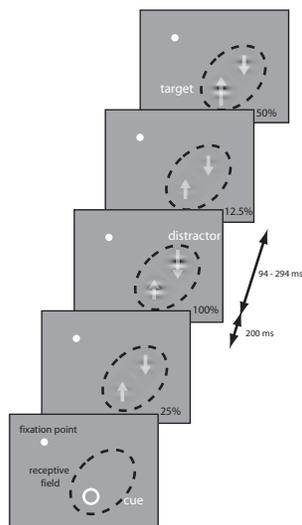


Figure 6.18: Task design for the Lee and Maunsell experiment. Once the animal had held its gaze on the fixation point, the attentional target was signalled either by instruction trials or a location cue, and pairs of Gabors with equal contrast were presented in the receptive field. The animal had to detect a change in the speed of the cued stimulus while ignoring changes in the distractor. (Reproduced from [112]).

multiplicatively on the contrast-response function:

$$R = aR_{max} \left(\frac{c^n}{c^n + c_{50}^n} \right) + m. \quad (6.8)$$

The contrast gain model is given by:

$$R = R_{max} \left(\frac{c^n}{c^n + (ac_{50})^n} \right) + m. \quad (6.9)$$

For both models, a is fixed at 1 when attention is directed to the non-preferred stimulus and varies freely when attention is directed to the preferred stimulus.

After fitting the average responses with the response gain and contrast gain models, the partial correlation was calculated using the correlation between the data and each model's fit. Fisher's r -to- Z transformation was then applied to each partial correlation coefficient to produce a Z score, and the transformed values were divided by the standard error, given as the square root of the degrees of freedom (df) minus 3. In the experiment and simulations presented below, $df=16$ is the number of data points, with 8 values from each attentional condition. To test whether the response modulation was more consistent with a response or contrast gain model, a statistical criterion of 1.645 was used, which corresponds to a p value of 0.05.

In a third form of analysis, the data were fit using Equation 6.7, but with m and n fixed between attentional conditions to permit a closer examination of the attentional influence on R_{max} and c_{50} . In this approach, a neuron would be said to exhibit a pure response gain if R_{max} changes between conditions and c_{50} does not. In a pure contrast gain however, c_{50} would change between attentional conditions while R_{max} would not.

Comparisons of the changes in R_{max} and c_{50} between attentional conditions were made by computing a modulation index:

$$mi(R_{max}) = \left(\frac{R_{max}^p - R_{max}^{np}}{R_{max}^p + R_{max}^{np}} \right) \quad (6.10)$$

$$mi(c_{50}) = \left(\frac{c_{50}^{np} - c_{50}^p}{c_{50}^{np} + c_{50}^p} \right), \quad (6.11)$$

where p and np indicate attention being directed to the preferred and non-preferred stimuli respectively.

To simulate this experiment in the ARC, a similar network to that used in the simulations of the Womelsdorf et al. [272] and Treue and Martinez-Trujillo [245] was implemented, and is shown in Figure 6.19. Nine V1 columns provide input to a single MT column having 9×50 layer-IV neurons, 200 layer-II/III neurons from which recordings were made, and 100 layer-VI control neurons. To ensure that the layer-II/III neurons responded similarly weakly to 0% contrast, their x -intercepts were distributed in the range [-0.1, 0.1]. Further,

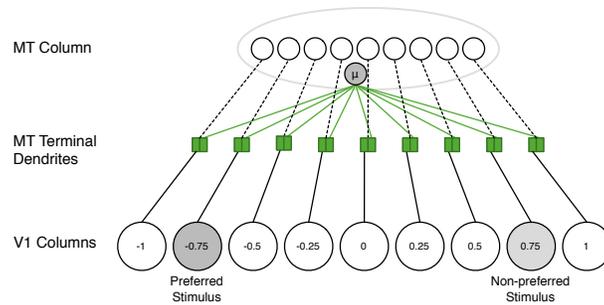


Figure 6.19: ARC with spiking neurons used for simulations of Lee and Maunsell (2010). MT has a single column containing 450 layer-IV pyramidal neurons, subsets of which receive input from a single V1 column. Control neurons in the MT column (gray circle) project a local control signal to the terminal dendrites of MT neurons to enable selective routing. The preferred stimulus is presented at $x_{-0.75}$ and the non-preferred stimulus at $x_{0.75}$.

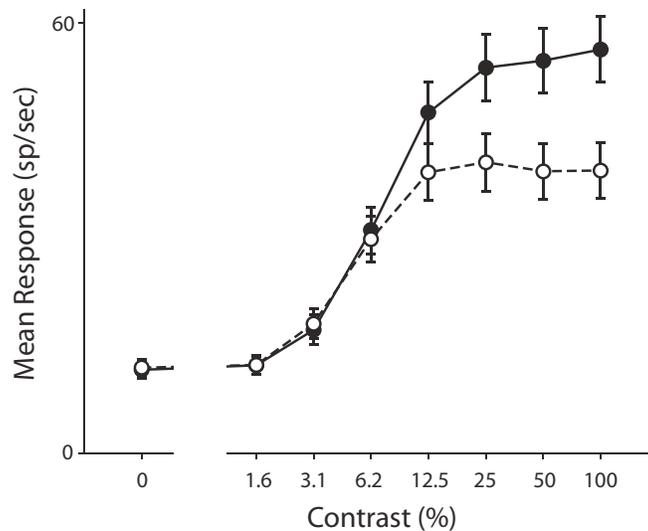


Figure 6.20: Average response of 56 MT neurons reported by Lee and Maunsell (2010) with attention directed to the preferred stimulus (black solid line) or non-preferred stimulus (dashed line). Error bars are SE. (Reproduced from [112]).

as in the previous simulations, the layer-II/III neurons had encoding vectors of 1 such that they respond more strongly to higher contrast stimuli. The activity of all cells was recorded for 36 repetitions at each of the eight contrast values and in both attentional conditions. To add variability in the neural responses across repetitions, noise was simulated by injecting the soma of the layer-II/III neurons with a random amount of current drawn from a uniform distribution in the range $[-22.5\mu\text{A}, 22.5\mu\text{A}]$ at each time step. Further, to mimic the variability introduced by neurons below MT, the input was scaled by a factor randomly chosen from a uniform distribution in the range $[75\%, 125\%]$. Although this has the potential to introduce a considerable amount of noise in the input signals, its effects on firing rate are less apparent, as the activity is averaged across 36 repetitions. This is discussed further in Section 6.4.

Figure 6.20 shows the average response of the 56 cells recorded by Lee and Maunsell [112] as the stimulus contrast and attentional target was varied. Solid black circles are the responses when attending the preferred stimulus, and unfilled circles are the responses when attending the non-preferred stimulus. Although these contrast-response functions are not uncommon for MT and V4 cells [112, 190, 266], their tuning to contrast is distinctly different than for other feature dimensions. Specifically, the recorded cells respond similarly for subthreshold contrasts below 3%, and have considerable sensitivity to a narrow range of contrasts between 3% and 12% before largely saturating at contrasts above 25%. This high sensitivity to contrast, along with preference for low spatial frequency and high temporal frequency stimuli, are among the hallmark characteristics of cells along the pathway extending from the magnocellular layers of LGN [97]. Indeed, similar tuning has been observed in V1 neurons projecting to dorsal stream areas, but with less sensitivity than MT cells, or a higher c_{50} value when fit with the hyperbolic ratio function (Equation 6.7) [3, 266].

In the NEF, this effect of saturation at higher contrasts can be simulated with a population of neurons having encoding vectors that are -0.2 or 0.2 (Figure 6.21(a)). The y -axis of Figure 6.21(b) plots the decoded estimate of a continuously monotonically increasing input value in the range $[0, 1]$. This population is able to well represent signals having values within the range for which their encoders are defined (i.e. $[-0.2, 0.2]$), although values beyond this range largely result in the neural activity saturating, as well as the decoded estimate of the signal carried in their activity.

Given that this simulation aims to investigate the influence of attention rather than the processing of information in lower areas, the computational requirements of the model can be reduced by simply replacing the input population shown in Figure 6.21(a) with 10 direct mode neurons that approximate this saturated value. In the NEF, direct mode neurons are non-spiking neuron-like elements that functionally represent and transform signals similarly to spiking neurons, but with considerably less computational demands. In the present simulations, these input neurons threshold any input value exceeding 0.3 to mimic the high

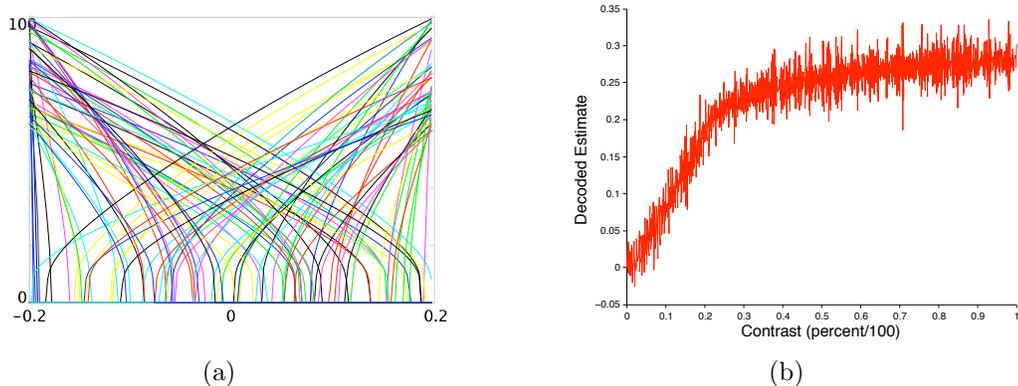


Figure 6.21: (a) Response functions for a population of 100 neurons with encoding vectors of -0.2 and 0.2 . (b) Decoded estimate of a continuously increasing value. The cells have high sensitivity for input signals having values in the range $[0, 0.2]$, but largely saturate at higher contrasts. These responses were used in the model to account for the high contrast sensitivity and saturation effects of neurons projecting to MT.

contrast sensitivity and saturation effects at higher contrasts observed in areas projecting to MT that receive signals originating in the magnocellular layers of LGN [258]. Layer-II/III neurons in the V1 columns are also “on” neurons, with the preferred stimulus moving in the direction 1.0 , and the non-preferred stimulus moving in a direction of 0.5 (i.e. orthogonally). This input is then projected to the layer-IV MT cells as the dot product of the stimulus direction and its contrast.

Figure 6.22 shows the activity of an example neuron from the ARC when attention was directed to the preferred stimulus (filled circles and black solid line) and when attention was directed to the non-preferred stimulus (open circles and dashed line). The fitting for this cell was performed using Equation 6.7 for the average activity across 36 repetitions, with error bars indicating SE. For this neuron, the fits explained 99.8% of the variance of the mean responses.

Figure 6.23(a) shows the average activity of the entire population when fit with the hyperbolic ratio function (Equation 6.7). For all 200 neurons, the median variance explained was 99.7% with attention directed to the preferred stimulus and 99.8% for the non-preferred stimulus. Figure 6.23(b) shows the same data as in Figure 6.23(a), but with each neuron’s activity normalized to its largest response. As with Lee and Maunsell [112], this typically occurred when attending the preferred stimulus at the highest contrast value.

For completeness, and to allow comparison with other studies [266], the distributions of parameter values obtained from fitting the data with Equation 6.7 are shown in Figure 6.24. Plots in the left column indicate parameter values when attention was directed to the

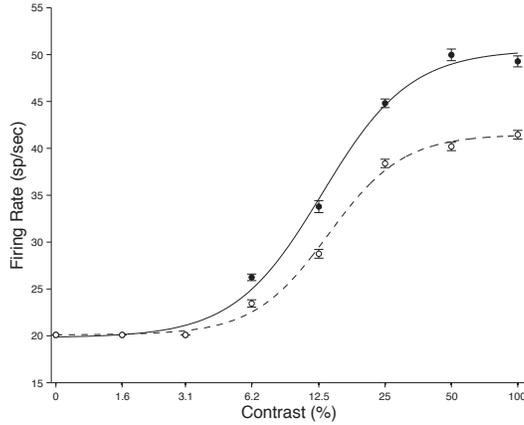


Figure 6.22: Contrast-response functions of a single layer II/III model neuron. The solid and dashed lines are the best fitting function when attention was directed to the preferred and non-preferred stimuli respectively. Error bars are SE values.

preferred stimulus and the right column when attending the non-preferred stimulus (dashed lines indicate the median value). For R_{max} (Figures 6.24(a), 6.24(b)), the mean value when attending the preferred stimulus was 62.1154 (95% CI= [59.8235, 64.3975]), and the mean when attending the non-preferred stimulus was 43.7445 (95% CI=[42.3958, 45.1034]). For n , the mean value when attending the preferred stimulus was 2.3161 (95% CI=[2.2061, 2.4304]), and 2.8574 (95% CI=[2.6799, 3.0419]) for the non-preferred stimulus. For c_{50} , the mean value when attending the preferred stimulus was 0.1297 (95% CI=[0.1272, 0.1322]), and 0.1471 (95% CI=[0.1426, 0.1516]) for the non-preferred stimulus. Finally, for m the mean value when attending the preferred stimulus was 7.3892 (95% CI=[6.2704, 8.5141]) and 6.5354 (95% CI=[5.4741, 7.6785]) for the non-preferred stimulus.

In subsequent analysis, the average responses in the two attentional conditions were fit with the response-gain and contrast-gain models (Equations 6.8 and 6.9) to determine which of these two models better explained the simulation data. Following Lee and Maunsell [112], the partial correlation was computed between the simulation data and the values generated by each fit, and transformed to a Z score. Figure 6.25(a) shows a scatter plot of the Z-transformed correlations for the contrast-gain and response-gain models. Filled circles represent neurons that were significantly better fit by the response-gain model, and open circles are neurons for which neither model provides a statistically significantly better fit. Dotted lines mark the statistical criterion for the test.

Of the 200 neurons that were fit by both the contrast-gain and response-gain models, only one could not be distinguished as being better fit by either model. As seen in Figure 6.25(a), the 199 neurons that were significantly better fit by the response-gain model

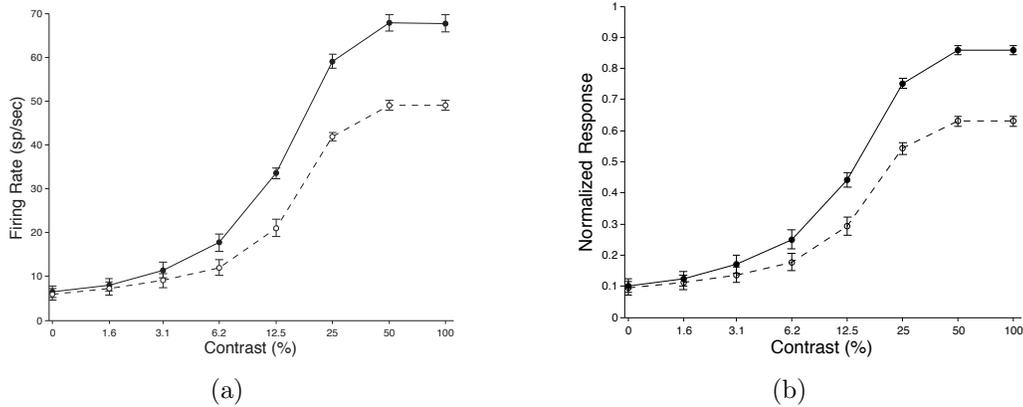
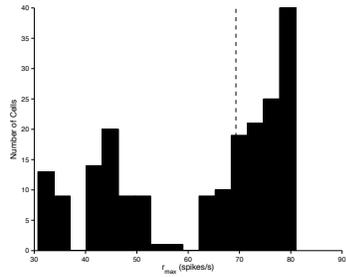


Figure 6.23: Population average contrast-response functions for 200 model neurons. The solid and dashed lines are the best fitting function when attention was directed to the preferred and non-preferred stimuli respectively. Error bars are SE values. (a) Average response across the population. (b) Normalized average population responses, with each neuron’s response is normalized to to its maximum mean firing rate.

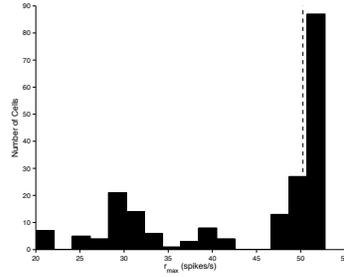
appear to form two distinct clusters. The first cluster, having higher Z -scores for both Z_c and Z_r contains 92 neurons, all of which have x -intercepts greater than zero and lower background activity (J^{bias}). The consequence of their higher x -intercepts is that the majority of cells in this cluster did not respond above baseline to the first two contrast values (0% and 1.6%). The other cluster in which Z_c and Z_r are lower, predominately contains neurons having x -intercepts less than zero and higher background activity (J^{bias}), resulting in them exhibiting stronger responses at the lowest contrast values.

For comparison, Figure 6.25(b) shows the Z -transformed partial correlations reported by Lee and Maunsell [112]. Of the 56 cells recorded, 45 were significantly better fit by the response gain model, 10 were comparably fit by both models, and one cell was significantly better fit by the contrast gain model. As with the simulation data from the ARC, the majority of cells in the experimental data were better explained by a response gain.

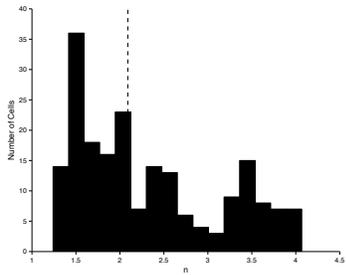
Figure 6.26(a) shows the results from the third form of analysis, in which responses were fit using Equation 6.7, with m and n fixed between attentional conditions. Gray circles indicate neurons in which there was a significant change in the modulation index of R_{max} and c_{50} , and black filled circles indicate a significant change in R_{max} but not in c_{50} . The solid lines indicate the 95% confidence intervals. Of the 200 neurons, 18 showed a significant change in R_{max} , consistent with a pure response gain, and the remaining 182 exhibited a significant change in both R_{max} and c_{50} . This finding is consistent with the increased mean value of c_{50} when attending the non-preferred stimulus (Figure 6.24(e), 6.24(f)), when fit



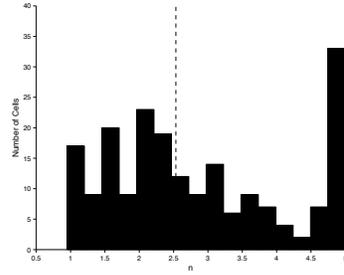
(a) R_{max} - attend preferred



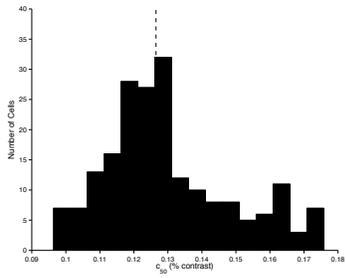
(b) R_{max} - attend non-preferred



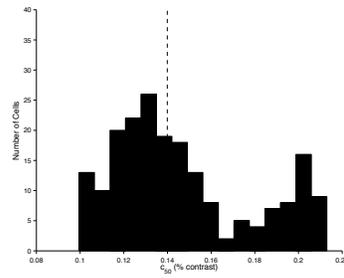
(c) n - attend preferred



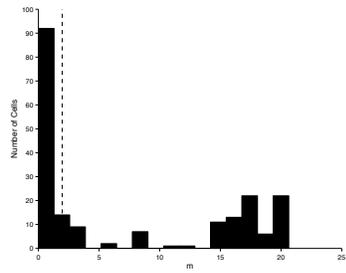
(d) n - attend non-preferred



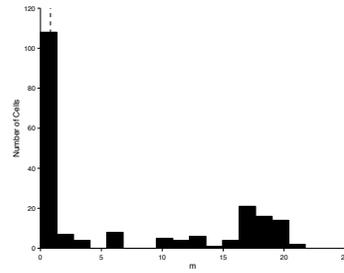
(e) c_{50} - attend preferred



(f) c_{50} - attend non-preferred



(g) m - attend preferred



(h) m - attend non-preferred

Figure 6.24: Distribution of fit parameter values when attending the preferred (left column) or non-preferred stimulus (right column). Dashed lines are the median parameter values from 200 neurons. See text for details.

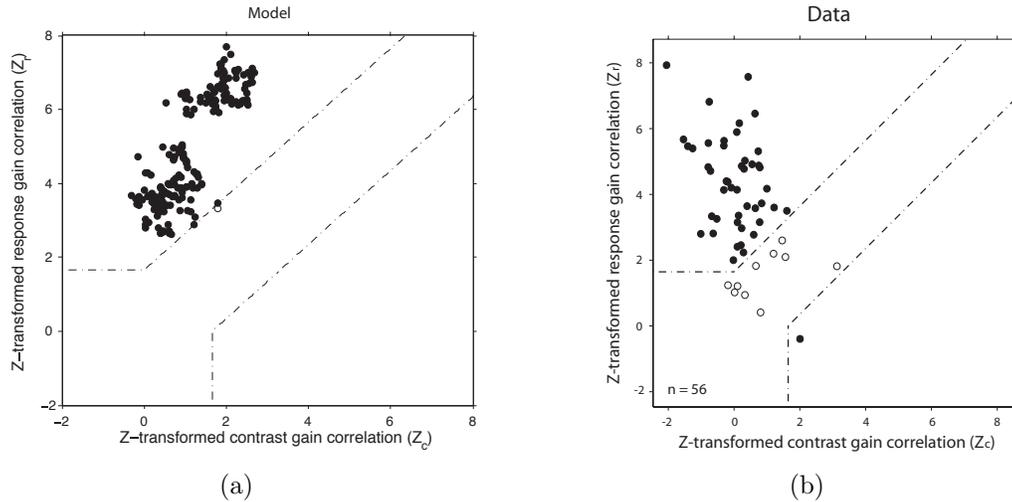


Figure 6.25: Z-transformed partial correlations between simulation data and curve fitting. Black circles are neurons having fits that are significantly better described by the response gain model. Dotted lines indicate the threshold for statistical significance (1.645, corresponding to $p=0.05$). The model (a) and experimental data (b) both indicate that for the majority of neurons, the attentional effect on contrast-response functions is significantly better explained by response gain. (Figure (b) reproduced from [112]).

with the standard hyperbolic ratio function (Equation 6.7). For comparison, the results reported by Lee and Maunsell are shown in Figure 6.26(b). Of the 56 recorded cells, 27 had a significant change in R_{max} but not c_{50} , 21 had a significant change in both, and the remaining 8 had no significant change in either parameter (white circles).

6.3.1 Results Summary

Of the numerous studies that have examined contrast-response functions [115, 125, 190, 266], the experiment by Lee and Maunsell [112] was selected, as their method avoids possible confounds found in earlier studies, and provides a rigorous examination of the data that quantitatively evaluates both contrast gain and response gain.

The simulation data were shown to be consistent with the results of Lee and Maunsell, which report that the dominant effect of attention is a multiplicative scaling of neural sensory responses. Lee and Maunsell found that none of the recorded cells showed a clear attentional effect that was consistent with contrast gain, a finding that was reproduced by the ARC. The results of this simulation provide further insight for addressing the question of whether shifting attention between receptive field stimuli will produce an effect that is more consistent with response-gain or contrast-gain. The simulations demonstrate that,

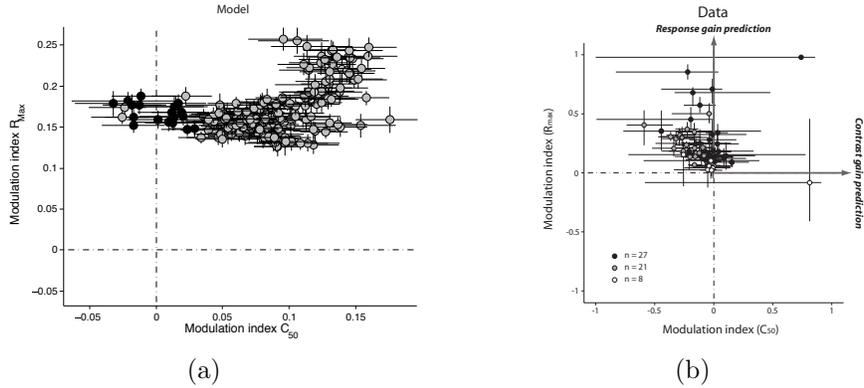


Figure 6.26: Modulation indexes for R_{max} (y -axis) and c_{50} (x -axis). Gray circles – neurons showing a significant change in both R_{max} and c_{50} . Black filled circles – significant change in R_{max} but not in c_{50} . Unfilled circles – neurons with no significant change. Solid lines – 95% confidence intervals. Both the model and experimental data show that the majority of cells had a significant increase in their maximum firing rate (R_{max}) when attending the preferred stimulus, consistent with a response gain effect. (a) Simulation data from the ARC. (b) Experimental data from Lee and Maunsell (2010). (Reproduced from [112]).

were the data only assessed using a contrast gain model, it could be rightfully concluded that contrast gain provides an excellent explanation of the phenomena (see Figures 6.23(a) and 6.25(a)). However, closer inspection of the data by quantitatively comparing both contrast-gain and response-gain models reveals that the data are significantly better explained by a response-gain model. These simulation results demonstrate that the analysis and fitting methods used to characterize the attentional effects can have a significant impact on the interpretation of the data.

The core of the model used to simulate the experiment by Lee and Maunsell [112] was the same as that used in the first two simulations. However, in order to capture the contrast sensitivity characteristics seen in the experimental data, the layer-II/III neurons in V1 and MT were altered. In this simulation, the V1 neurons were constructed to exhibit the high contrast sensitivity seen in the experimental data, as well as the LGN-magnocellular recipient neurons in V1 reported elsewhere [3, 258, 266]. As in the previous simulations with the ARC, the layer-II/III neurons in MT were modelled as “on” neurons, although for the simulations of the Lee and Maunsell experiment, their x -intercepts were drawn from a significantly narrower range to ensure that their responses were similarly weak for low contrasts. However, contrast is but one of the feature dimensions to which neurons in higher cortical areas are sensitive, and as was shown in the simulations, for other stimulus dimensions, a much broader sensitivity can be used.

Consistent with the results reported by Lee and Maunsell, when the simulation data

were fit using a response gain and contrast gain model, it was found that response gain significantly better describes the attentional effects 6.25. However, similar to the results from the previous two simulations in this chapter, the distribution of the simulation data was significantly narrower than the experimental data. For example, in the simulation data (Figure 6.25(a)), only one neuron was equally well described by both contrast gain and response gain models, although Lee and Maunsell report that 10 of the 56 recorded cells were equally well fit by both models (Figure 6.25(b)). Additionally, the experimental data are generally more scattered than are the simulation data, although both sets of results demonstrate that, in the majority of neurons, the attentional effect is significantly better described as response gain.

In comparing the two scatterplots shown in Figure 6.26, the experimental data generally have considerably larger confidence intervals of the mean for both parameters, whereas the simulation data have considerably less variation. For example, the cell in the top right corner of Figure 6.26(b) has a 95% confidence interval for the modulation index of c_{50} that, after being bootstrapped 1000 times, nearly spans the range $[-1, 1]$. For a cell to exhibit such widely varying modulation indexes, its responses would need to be significantly larger when responding to the preferred stimulus on some repetitions, and significantly larger when attending the non-preferred stimulus on other repetitions. However, another study of the attentional effects on contrast-response functions in area MT by Martinez-Trujillo and Treue [125] had to discard 29/63 recorded neurons because the fits produced c_{50} values that exceeded the maximum contrast used in the experiment in at least one attentional condition. If neurons with similar fits to those discarded in [125] were included in the analysis in the Lee and Maunsell study, this may partially explain the considerable variance they report, although it is unknown whether such neurons were included.

The 95% confidence intervals of the simulation data are significantly narrower than the experimental data, despite a similar number of repetitions being performed (the experiment performed 36 on average, and the simulations used exactly 36), and the same analysis methods being used. Even with 25% noise added to the encoded values of the input signals, the recorded cells do not exhibit the considerable variability found in the experimental data. Qualitatively consistent results were obtained without the inclusion of noise at the input level, although there was significantly less variability across repetitions. Several possible sources of trial to trial variability may have influenced the experimental recordings that are not present in the model, such as measurement error, stimulus processing effects, changes in the animal's cognitive state and alertness, and involvement of different cell assemblies. Given the paucity of data that can inform the selection the most appropriate method for modelling this variability, the simple approach used here provides a suitable approximation.

6.4 Summary of Simulation Results

In this chapter, the ARC was used to simulate three experiments of attentional processing by Womelsdorf et al. [272], Treue and Martinez-Trujillo [245], and Lee and Maunsell [112]. In the analysis of each simulation, the data were analysed using the same methods as in the original experiments, and the model was shown to be quantitatively consistent with the experimental results.

The first simulation presented in this chapter (Womelsdorf et al. [272]) demonstrated that directing attention to a target inside the receptive field causes the receptive field to shrink and shift toward the target. This form of attentional modulation occurs as a result of the Gaussian shaped routing function attenuating input signals from neurons located in parts of the receptive field that are distant from where information is being selectively processed by the column (μ).

Using this same model, the findings of Treue and Martinez-Trujillo [245] were also reproduced, which highlights that the allocation of spatial attention does not affect tuning sensitivity. Rather, the model demonstrates that changes in neural responses occurring when attending a preferred or non-preferred stimulus can be attributed to changes in the multiplicative gain term from the routing function that are determined by the location of the attentional target. This finding distinguishes the ARC from the normalization model [189] and biased competition model [188], which predict that directing spatial attention to a preferred receptive field stimulus will cause a sharpening of the tuning curves, which is inconsistent with the experimental data and simulation data from the ARC. Consistent with these effects, the third simulation of the Lee and Maunsell experiment [112] demonstrated that the attentional effect of response gain also can be seen to result from the multiplicative gain term produced by the routing function, and that response gain is equivalent to the increased amplitude and unchanging tuning curve selectivity reported by Treue and Martinez-Trujillo [245].

From a modelling perspective, the configuration of the Treue and Martinez-Trujillo experiment is similar to that of Womelsdorf et al.: two stimuli are placed in the receptive field of the recorded cell, and attention is directed to either of the stimuli, or to the fixation point. However, the Womelsdorf et al. experiment examined the effects of attention on the spatial properties of the receptive field by measuring responses to probe stimuli, while the Treue and Martinez-Trujillo experiment assessed the effects of attention on feature selectivity. The similarity of the experimental configurations allows the model's generality to be tested by using the same model configuration to simulate both experiments, with the primary difference between these two simulations being the stimuli that were used.

In the model, it is not particularly surprising that the peak response activity is consistently higher when attending to the variable direction stimulus than when attending

the anti-preferred stimulus. Since the layer-II/III neurons from which the activity was recorded are “on” neurons, their activity is strongly correlated with the value being represented. This value is based on a weighted summation of the feature values of each stimulus, multiplied by a gain term computed by the routing function. When attention is directed to the anti-preferred stimulus, its value is multiplied by the peak of the routing function, and the other stimulus is attenuated, resulting in the inputs to the recorded cells being predominately from the anti-preferred stimulus. Similarly, attending to the variable motion direction stimulus causes the inputs from the anti-preferred stimulus to be attenuated by the routing function, thereby typically yielding a greater response from the recorded cells. Regardless, these results demonstrate that the ARC is consistent with the experimental observations of increased directional gain, and that the range of stimulus values to which the cells respond does not change with attentional shifts.

The third set of simulations with the ARC modelled the experiment by Lee and Maunsell [112] that investigated the effect of attentional shifts on contrast response functions. As with the previous two simulations, the attentional effects shown by Lee and Maunsell are consistent with that predicted by the model’s principles. When compared with the results of Treue and Martinez-Trujillo [245], the model predicts that a response gain should be expected. The Treue and Martinez-Trujillo study found that spatial shifts of attention produce a multiplicative scaling of neural responses across motion directions. This effect was similarly shown by Lee and Maunsell, except with a multiplicative scaling across contrasts. In this light, the attentional effects reported by Treue and Martinez-Trujillo for motion direction are equivalent to that produced by a response gain model.

Situating the combined results from these three experiments within the context of the large-scale functional model allows the experimental observations of attentional modulation to be interpreted at multiple levels of analysis. This property of the ARC, as discussed in previous chapters, is presently not found in the majority of models of attention. Using a single model of attentional routing also highlights that the seemingly different forms of modulation can be explained through a common mechanism of visuospatial attention. In each of the simulations, the gain term produced by the ARC’s routing function affect the feedforward visual signals similarly, although by using different stimuli and analysis methods on the data, these seemingly different forms of attentional modulation were reproduced using the same mechanisms. The model thus predicts that the various attentional effect observed in feature responses are primarily consequent to the modulation of spatial receptive fields.

These results also highlight several distinctions between the ARC and previous models. First, the simulations were performed using a spiking neuron model that was shown to be consistent with three forms of attentional modulation. The gain field model by Salinas and Abbott [204] was shown to be qualitatively consistent with reports of receptive field shifts, although that model uses rate code neurons and the model results were not quan-

titatively compared with specific physiological data. The synchrony- and oscillation-based models [162, 163] are implemented in spiking neurons, but it remains to be demonstrated that they can account for the attentional effects described in this chapter.

The ARC also has a significant explanatory advantage over models using mathematical implementations. The shifter circuit [168] and SAIM [82] models address the problem of selective processing at a more abstract level than the ARC, which significantly limits the models' ability to address neurophysiological effects.

The results of the simulations presented in this chapter also distinguish the ARC from the biased competition [188] and normalization models [189], which are mathematical implementations. As with other non-spiking models, the neurophysiological explanatory and predictive capabilities of the biased competition and normalization models are limited, as the plausibility of the models being implemented in spiking neurons is unclear.

Implementation concerns aside, the biased competition and normalization models differ from the ARC in their predictions about the forms of attentional modulation. Specifically, the biased competition and normalization models predict that in the Treue and Martinez-Trujillo experiment [245], a neuron's tuning selectivity will become sharper when attending the preferred stimulus, although experimental results report that the tuning selectivity does not change. In contrast, the ARC's principles predict that selectivity will not change with spatial shifts of attention, and this prediction was shown to be quantitatively consistent with the experimental data.

Chapter 7

Future Work and Conclusions

The ARC provides a detailed account of how the selective routing of attended visual information may be performed in cortex. Extending this high level functional model with a biophysically detailed and anatomically consistent spiking neuron implementation allows the principles of the model to be quantitatively tested against specific neurophysiological recordings. The design of the model permits it to be significantly scaled up, in terms of additional hierarchical levels, the number of neurons in each column, the visual features they encode, and the number of columns in each level. This scalability results from the use of common ensembles of control neurons in each column that compute local control signals for all visually responsive neurons in the column. The local control signals are computed using minimally complex functions on low dimensional signals, which allows them to be computed accurately and efficiently with plausible neural requirements.

7.1 Summary of Results and Predictions

The presentation of the ARC in this thesis began with a mathematical description of selective attentional routing in cortex (Chapter 4). The mathematical model was shown to perform routing similar to a shifter circuit [168, 169], but in a manner that overcomes several limitations of that model. Specifically, the ARC performs selective routing without needing to change synaptic weights, without exceeding the neural capacity of pulvinar, and without need for specific long-range projections from pulvinar.

The mathematical implementation of the ARC provides a strong functional account of performing selective attentional routing in cortex. However, the plausibility of its computations being performed in neurons remained to be tested. To verify the neural feasibility of the model, the ARC was implemented in spiking neurons using the Neural Engineering

Framework (NEF) [55], and was shown to perform routing as predicted by the mathematical model (Chapter 5). It was demonstrated that control signals can be computed and represented with minimal control neurons in each column, in a manner that is consistent with the anatomical connectivity and neural density of the modelled cortical areas.

Having demonstrated in Chapter 5 that the proposed mechanisms for selective routing can be implemented in spiking neurons, Chapter 6 evaluated the model's principles by comparing it to the attentional effects observed in real brains. This evaluation was conducted by simulating three physiology studies of visual attention in macaque, where each study reported different forms of attentional modulation.

The first simulation (Section 6.1) was of an experiment by Womelsdorf et al. [272], which found that attentional shifts result in the receptive fields shifting toward the attentional target and shrinking in size. The simulation results from the ARC were shown to be quantitatively consistent with each of the observed effects, from which several predictions about the properties of attentional modulation in this experiment were generated, and are discussed below.

The second simulation (Section 6.2) modelled an experiment by Treue and Martinez-Trujillo [245], which found that shifting attention between a variable motion and non-preferred stimulus presented inside the receptive field produces an increase in receptive field gain when attending the variable motion stimulus and a reduction in gain when attending the non-preferred stimulus. Additionally, that experiment reported that with attentional shifts, there was no significant change in the tuning selectivity. Using the same model as in the Womelsdorf et al. [272] simulations, the ARC was able to reproduce the effects reported experimentally. These results also allow the model to be distinguished from previous proposals, as discussed in Section 7.2.

The third experiment that was simulated using the ARC (Section 6.3) was a study by Lee and Maunsell [112] that investigated the effects of attentional shifts on contrast-response functions. This study demonstrated that although the attentional effect can be partially explained as a contrast gain, by using a more detailed analysis that quantitatively assesses both contrast gain and response gain, the effect is significantly better explained by response gain. The results of simulating this experiment with the ARC were also shown to be consistent with the experimental results.

The combined results of these three simulations support the plausibility of the model's principles being found in cortex, as they demonstrate that the proposed mechanisms for selective routing are consistent with the attentional modulation observed in cortical neurons. Together, they also support the model's prediction of a single mechanism producing these different forms of attentional modulation. The simulations of the Womelsdorf et al. study were presented first, as they provide strong evidence of one of the model's central predictions, that of dynamic shifting and shrinking of receptive fields to perform selective

processing. The ARC was shown to be statistically indistinguishable from the experimental results, while using the same mechanisms for selective routing as in the large-scale mathematical model and in the simulation of the other experiments.

The mathematical model provides a context for interpreting the results of Womelsdorf et al., by suggesting a functional role for the receptive field shifts within the larger problem of selective routing. Specifically, the ARC proposes that the selective processing of attended stimuli is performed by dynamically modulating the portion of the receptive field from which information is selectively processed. This receptive field location is computed by cortical control neurons in each column that use global control signals to specify the centre of the Gaussian-shaped routing function.

The results of the Womelsdorf et al. simulations in turn provide a context for interpreting the effects of spatial attention on feature tuning and contrast-response functions reported by Treue and Martinez-Trujillo [245] and Lee and Maunsell [112]. The results of simulating the experiments by Treue and Martinez-Trujillo [245] and Lee and Maunsell [112] are consistent with the model's prediction that directing attention to a location imposes a multiplicative gain term on visual inputs that is based on their spatial location, but not on the signals encoded by neurons at each spatial position. Specifically, both of these studies report that shifts of spatial attention produce a multiplicative scaling of responses that similarly affects all motion directions and contrast values, without attenuating responses to non-preferred motion directions or stimuli with high contrasts.

7.1.1 Predictions

As stated at the outset of this thesis, a central strength of computational models is that they can be used to bridge the gap between theoretical proposals of brain function and experimental work by quantitatively testing whether the theoretical proposals are consistent with experimental observations. If the results of simulations with the model are not consistent with experimental findings, then the discrepancies between the model and data can inform where the model is wrong and how it may be revised. If however, the simulation results are consistent with experimental data, the model can be used to generate predictions that may then be tested experimentally. This iterative process of theorizing, simulating, and testing experimentally is crucial to advancing our understanding of brain function.

The ARC's consistency with the experimental data lend support to the model's principles. And, as a result, the model provides several predictions regarding other aspects of attentional processing that can be tested to further evaluate whether the model's principles are biophysically consistent. Details of these predictions were discussed in Chapters 4 and 6, and are now summarized.

1. The attentional modulation of feature selectivity, contrast-response functions and receptive field profiles with spatial shifts of attention result from a single mechanism for selectively processing attended stimuli. Specifically, attending to a particular feature dimension imposes a Gaussian shaped multiplicative gain term defined over the dimension being attended, and centred on the target feature.
2. Visual cortical neurons operate in a default routing state prior to the target being known and when they are not encoding information from the attentional target. The ARC predicts four properties of neurons in their default routing state:
 - (a) Neurons operating in their default routing state will have larger receptive fields on average than when their receptive field contains an attentional target.
 - (b) In the Womelsdorf task [272], neurons with receptive fields in the periphery will on average respond more strongly to probe stimuli presented on the side of their receptive field that is more distant from the fovea.
 - (c) When two stimuli are placed inside a peripheral neuron’s receptive field, the receptive field profile will on average, shift by a larger amount when attention is directed to the receptive field target that is located closer to the fovea than when attending the target on the opposite side of the receptive field.
 - (d) An increasing proportion of neurons at each higher cortical level will show attentional modulation on average, as a result of them switching from their default routing state to a selective routing state, and the size of the shift will also be larger.
3. Independent of the target’s size, the effect of spatial shifts of attention can be explained by a multiplicative scaling of responses across different feature dimensions including contrast, motion direction, orientation, and colour.
4. In the Womelsdorf task [272], as the two target stimuli are placed closer to the receptive field edge, there will be a systematic increase in the receptive field shift and decrease in amplitude on average.

7.1.2 Assumptions

Although the combined results of computational, physiological and psychophysical research have greatly advanced our knowledge of neural processing and attention, a considerable amount remains presently unknown. Given this uncertainty, in defining the model, several assumptions were made based on the currently available information. It is possible that future studies will provide evidence that conflicts with the ARC’s assumptions, and the remainder of this section discusses the impact on the model of these assumptions being incorrect.

Pulvinar Providing a Global Control Signal

The ARC assumes that pulvinar provides global control signals that are used to perform selective routing in visual cortex. This assumption is shared by several models of attention [2, 7, 162, 169], and there is a considerable amount of experimental support for this hypothesis. Single-cell recording studies have shown stronger activity in pulvinar neurons when their receptive field contains an attentional target [16, 179, 180, 192], imaging studies demonstrating stronger activation in pulvinar in attentional tasks [98, 110, 148, 260, 275] and pulvinar deactivation and lesion studies have reported deficits in engaging attention [45, 185, 262, 263].

However, myriad other cortical and subcortical areas are also known to be involved in guiding saccades and attentional allocation, including superior colliculus, FEF and LIP [146, 184, 215]. It is probable that the pulvinar is not the sole locus of attentional control, but rather serves as a hub for distributing high level control signals to visual cortex.

The model's principles for selective routing are independent of how the high level control signals are generated and in which brain region(s), as the ARC addresses the issue of selective routing once the target has been determined. If the high level control signals are computed in FEF, LIP, pulvinar, or distributed across these areas, it is suspected that the size and position of the attentional target will still be encoded in the signals from these areas. The basis for this proposal is that many neurons in FEF and LIP respond more strongly when their receptive fields contain an attentional target or target of an upcoming saccade [146, 184, 215] which, when combined across the populations of neurons in these areas, forms a map from which the location and spatial extent of the target can be inferred.

Global Control Signals Are Only Projected to the Topmost Level of the Hierarchy

The ARC proposes that global control signals for visuospatial attention terminate in layer-I of PIT, although it is possible that future studies may find that this information is also projected to lower areas such as V4, V2 and V1. This may be the case if the global control signals are encoded in FEF or LIP, which both project to V4 and IT [59, 86, 146]. This could affect routing in the ARC in two ways.

1. One possibility is that the global control signals A_{len} and A_{pos} are not exclusively projected to PIT, but are *simultaneously* projected to each ventral stream area (Figure 7.1(a)). In this case, the global control signals would still carry a coarse and low-dimensional signal indicating the size and position of the target, and each ventral stream area would simultaneously compute local control signals using the same mechanisms as in the current model. In this arrangement, the predictions from the

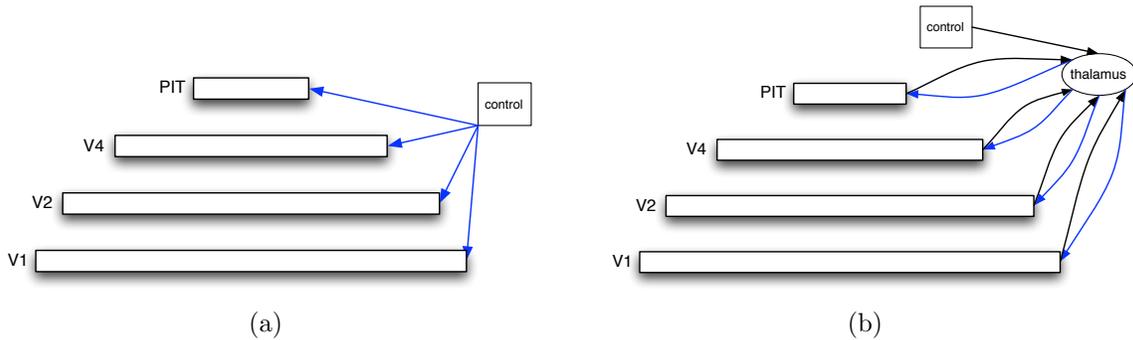


Figure 7.1: Alternative methods for distributing high level attentional control signals. (a) Global control signals are projected simultaneously to each ventral stream area from some unspecified area(s), such as FEF, LIP, pulvinar, or some combination therein. (b) Global control signals are projected from the area(s) in which the target is selected to thalamus. From thalamus, the global control signals are first projected to PIT, which relays the signals back to thalamus, and this cycle continues for areas V4, V2, and V1.

model still hold, as do the routing mechanisms, although the role of feedback projections would need to be reconsidered, as it is unclear how the differential timing of modulation in different levels would arise.

2. A second alternative is that global control signals are distributed to the hierarchy in a top-down manner, but are relayed through the thalamus between each cortical level (Figure 7.1(b)). This arrangement involves the global control signals first being projected from the brain region(s) encoding the signals, be it pulvinar, areas in the frontal or parietal cortex, or some combination therein, to the top level of the hierarchy via thalamus, where local control signals are computed as in the current model. However, instead of the global signals being fed back from PIT to V4, they would be relayed to the thalamus, and then from thalamus to V4. This communication could be realized using the corticothalamic projections from neurons layer-V of visual cortex [240] and the thalamocortical projections terminating in layer-I and layer-IV [97]. In this arrangement, the methods for computing local control signals do not need to be altered from the present model, and may also allow thalamus to restrict the distribution of the global control signals depending on the task demands. For example, in simple tasks that do not require fine spatial details for discrimination, thalamus could prevent the global control signals from being projected to V2 or V1. This arrangement requires the global control signals to traverse through several additional synapses, and thus additional time would be added to the delay at which attentional modulation is observable in each level. The timing of attentional modu-

lation in each area varies considerably between studies [26, 32, 138, 196], and only recently have experiments been conducted in which the timing of modulation has been measured in different ventral stream areas of the same monkey. As additional studies are conducted that can describe the timing of modulation in more detail, their results will allow the suitability of these different configurations to be better evaluated.

Dendritic Nonlinearities in Layer-IV Pyramidal Cells

The ARC assumes that pyramidal cells with dendritic nonlinearities will be found in layer-IV of visual cortex. As was discussed in Section 2.4, this hypothesis is supported by the electrochemical and morphological similarities of layer-IV pyramidal neurons in cortex and the neurons with dendritic nonlinearities reported experimentally. If future experiments fail to locate any such cells in visual cortex, this would require alterations to the laminar mapping of the ARC, but not to the model’s principles. Specifically, the processing performed using the dendritic nonlinearities in the present model could be combined with the computations performed using the layer-IV neuron’s second nonlinearity. In order to handle the additional computational requirements of these neurons, a similar optimization of the encoding and decoding vectors to that described in Section 5.2 would likely need to be performed, although given the number and density of layer-IV neurons, there seem to be sufficient neural resources to perform these computations, although it remains to be explicitly tested.

Object-Centred Reference Frame

The ARC proposes that information from an attended stimulus is routed to an object-centred reference frame in which the representation of the object is largely invariant to its retinal size and position. Although there is strong experimental evidence that supports this hypothesis (Section 4.6), this hypothesis is not agreed upon by all theorists [252].

The present model defines routing through cortical areas V1 to PIT and does not include areas above PIT, as beyond this level, much less is understood about the spatial distribution of an object’s representation across the cortical sheet. For example, the retinotopic organization and moderate receptive field sizes in V1 to V4 significantly constrain which columns could encode information from the target. Beyond PIT however, the columnar organization of each area becomes much more complex. In CIT and AIT, receptive fields can cover 50° or more of the visual field [41, 200], and the columnar organization seems to be defined according to the high dimensional feature space that neurons in each column are selective for (e.g. circles, corners, textured oriented ellipses, and common facial features) [62, 233]. It has been proposed that the complex properties of AIT columns are

assembled through the specific columnar projections from columns in PIT that are selective to more simple features [60, 203].

The general proposal of an object-centred reference frame is that the representation of an attended stimulus is largely invariant to the object’s retinal size and position. Further, in the cortical area where the object-centred reference frame is formed, whether it is PIT, CIT, AIT, or an even higher area, all columns in that area attempt to selectively process information from the attended stimulus. Due to the increasingly specific columnar tuning in these areas, some columns will respond weakly or not at all when the target is non-preferred, although any column that is sensitive to stimuli similar to the target will be selectively processing its information. This would be seen as “hotspots” of activity in response to a particular attended stimulus, and these hotspots may be separated by columns showing little activity. This is consistent with reports of neurons in IT that are selective for a particular stimulus and give weak responses when attention is directed to a non-preferred stimulus, despite the preferred stimulus still being within their receptive field [147].

Such activity patterns have been observed in several studies of macaque area TE, where spatially separated columnar clusters of cells, each covering $\sim 400\mu m$, show strong activation when a preferred stimulus is displayed [202, 234, 254, 274]. Koch [104] estimates that the entire area of TE could be parcelled into more than a thousand such “hotspots,” which is consistent with the estimate of 30×30 sampling units proposed by the shifter circuit [168, 169, 257].

In higher cortical areas beyond PIT, selective routing could still be performed using the principles of the ARC, although it is unclear if each column would selectively process a portion of the target as seen in some V4 neurons [23, 177], or whether the entire attended object would be routed to each column. Given the complexity of tuning in these areas, which requires pooled information from a larger spatial areas to form high dimensional feature conjunctions, it seems plausible that in areas beyond PIT, columns are selectively processing information from an increasingly large portion of the target, and eventually, the entire target.

7.2 Summary of Comparison to Past Models

In Section 3.6, a set of limitations that are common amongst existing models of attention was given. A central goal of the ARC is to overcome these limitations in a single mechanistic model, by proposing a model which is:

1. Implemented in spiking neurons, within biophysical constraints;
2. Defined for an entire cortical area, rather than for a single cell;

Table 7.1: Summary of common limitations in existing models of attention. Column headings indicate whether each model satisfies the following criteria: Spiking - implementation using spiking neurons. Entire cortical area - describes the relative effects of attentional modulation throughout an entire cortical area. Multiple cortical areas - describes the relationship of attentional effects between multiple cortical areas. Matching to physiology - the model is shown to account for specific physiological data. Multiple forms of modulation - the model is shown to account for multiple forms of attentional modulation within a single model.

Model	Spiking	Entire cortical area	Multiple cortical areas	Matching to physiology	Multiple forms of modulation
ARC	yes	yes	yes	yes	yes
Salinas and Abbott [204]	yes	no	no	no	no
Womelsdorf et al. [272]	no	no	no	yes	no
Shifter Circuit [168]	no	yes	yes	no	yes
SAIM [82]	yes	no	no	no	no
Biased Competition [188]	no	no	no	yes	yes
Normalization Model [189]	no	yes	no	yes	yes
Synchrony and Oscillation [163, 162]	yes	yes	no	yes	yes

3. Able to account for the relationship of attentional modulation between multiple cortical areas, and the interaction of attentional routing between these areas;
4. Well-matched to physiological data; and
5. Able to account for different forms of attentional modulation using a common mechanism.

Table 7.1 summarizes the limitations of each of the models discussed in Chapter 3. In the ARC, the first point was addressed in Chapter 5, in which the spiking neuron implementation was first presented and the model’s detailed laminar mapping was specified. The second and third point were addressed in Chapter 4 where the methods for performing selective routing in each hierarchical level were defined. The final two points were then addressed in Chapter 6, where the ARC was shown to account for three seemingly different forms of attentional modulation using the same mechanisms for each simulation as defined in the models presented in Chapters 4 and 5.

The ARC is most closely related to the shifter circuit [168, 169], which provides a strong starting point for modelling visual attention, as it frames the problem of visual attention as a problem of dynamically routing attended information through the visual hierarchy to an object-centred reference frame, provides a well-defined functional account of selective routing, and is largely biologically consistent. However, the shifter circuit is a purely mathematical implementation that lacks detail of how the model may be implemented in neurons. Additionally, it has several physiological inconsistencies.

Several aspects of the shifter circuit are incorporated in the ARC, including the proposal of dynamically routing an attended stimulus to an object-centred reference frame, and performing this routing by dynamically modulating the information being processed in each column. The shifter circuit performs selective processing in each column by changing the connection weights of inputs to each column, while the ARC demonstrated that a functionally similar effect can be produced in a biologically plausible manner through the multiplicative interaction of feedforward visual signals and a Gaussian-shaped routing function.

A section limitation of the shifter circuit is its proposal that all of the control signals for routing are computed in pulvinar, from where they are projected to specific connections in each ventral stream area. When the model is scaled to approximate the size of the human brain, the number of unique control signals required approaches the total number of neurons in pulvinar, and since pulvinar neurons are also required to compute these control signals, it suggests that pulvinar may not be sufficient for performing routing in such a model.

The ARC overcomes this limitation by proposing that pulvinar only coarsely encodes the size and position of the attentional target, with all detailed local control signals being computed by cortical control neurons. By implementing the ARC in spiking neurons, the neural requirements for computing the local control signals can be evaluated, and it was demonstrated that the control signals can be efficiently and accurately computed with minimal neural requirements.

The normalization model [189] is the other model with which the ARC can be most directly compared, as several aspects of that model are consistent with the ARC, and they both aim to explain similar attentional phenomena (see Section 4.7). However, the normalization model has several limitations that remain to be addressed. First, it is a mathematical implementation, which makes it unclear how biologically plausible its computations are, and limits the model's explanatory and predictive capacity.

Although the model can qualitatively account for several different forms of attention in a single model, several of its predictions are inconsistent with experimental results. Specifically, the model predicts that in the experiment by Treue and Martinez-Trujillo [245], the receptive field gain will increase when attending the variable motion stimulus and the tuning curve will become sharper (i.e. increasingly selective for the preferred stimulus). The first prediction that directing attention to a preferred receptive field stimulus produces an increase in gain has been demonstrated by numerous studies [118, 147, 186, 188, 246, 247], and the simulation result from normalization model was shown to be consistent with this prediction. The second prediction, that neurons will exhibit a sharpening of their tuning curves, is inconsistent with the results of the study that was being modelled. Rather, the study by Treue and Martinez-Trujillo [245] demonstrated that spatial shifts of attention do not alter the tuning selectivity of neurons in MT. The normalization model

suggests that the amount of tuning curve sharpening is related to the size of the attention field, although using their model code to simulate that experiment, the effects of increased gain without changes in selectivity could not be reproduced with larger or smaller attention field sizes. The ARC was also used to simulate the experiment by Treue and Martinez-Trujillo, and was able to quantitatively match the experimental data in a spiking neuron implementation.

The normalization model also predicts that the size of an attentional target will differently affect the attentional modulation for low and medium, although as was discussed in Section 4.7, the results of several studies have shown that stimulus size has little or no effect on contrast-response functions. The ARC however, predicts that the stimulus size will not cause different effects in the contrast-response function. With two stimuli inside the receptive field, the ARC predicts that the effect of spatial attention is a multiplicative scaling of responses, and that the amount of scaling is constant for all feature values at each position in the receptive field, independent of their size, contrast, or direction of motion.

7.3 Future Work

There are numerous directions in which future work with the ARC may proceed. In general, these extensions are related to increasing the scale of the model to larger sizes and including neurons that encode additional feature dimensions.

Multi-level hierarchy in spiking neurons

Due to computational constraints, the spiking neuron implementations in this thesis performs selective routing between two cortical areas. However, recent advancements in GPU-accelerated processing allow larger networks to be simulated, and the introduction of specialized computational hardware called Neuromorphic Computational Fabrics [216] can be used to emulate spiking neural networks for models containing up to a million neurons, running in real-time. Both of these advancements can be used to conduct significantly larger simulations with the ARC, and when the spiking neuron model is extended to include multiple hierarchical levels, there are three particularly promising directions for investigation.

- (a) Quantitatively measuring the timing of attentional modulation in each cortical area.
- (b) Quantitatively measuring the proportion of cells in each area and cortical layer that exhibit attentional modulation.
- (c) Quantitatively measuring the strength of modulation in each area and cortical layer.

Higher dimensional feature tuning of neurons in each level

It is also well established that the ventral stream areas modelled in the ARC play a crucial role in processing visual information for object recognition. This hierarchical processing, involving neurons in higher areas pooling spatial and temporal information from lower areas in order to construct more complex representations, presumably occurs concurrently with the routing mechanisms employed by the ARC. For simplicity, the model ignores the increasingly complexity of neural tuning at higher levels of the ventral stream, although it is assumed that incorporating neurons tuned to more complex stimuli will not affect the proposed mechanisms.

By incorporating neurons tuned to higher dimensional stimuli, there are two main directions for future work:

Incorporating the ARC Into a Hierarchical Vision System

A recent model by Tang and Eliasmith [235] that uses Deep Belief Networks for performing visual recognition has been implemented in spiking neurons using the NEF. Their results show that using a simple attention-esque top-down feed back signal significantly improved the model's recognition performance on a dataset of handwritten digits. The ARC is well suited for integration with hierarchical vision models, and suggests a promising application for the model.

Verifying the model's prediction of spatial attention not affecting tuning sensitivity with higher dimensional tuning

The simulations with the ARC presented in this thesis involve neurons that are tuned for one or two dimensional visual stimuli, and future work may explicitly test the model's prediction that spatial attention does not affect tuning sensitivity when using neurons tuned to higher dimensional stimuli. When using neurons sensitive to multiple feature dimensions, it may be that spatial attention causes an interaction between other feature dimensions that was not detectable with the low dimensional stimuli used in this thesis. Although the model predicts that there will not be any interactions, it remains to be explicitly tested.

Feature-based attention

That attentional modulation seems to have a Gaussian profile which interacts multiplicatively in the dimension being attended was shown here for spatial attention, although

several experimental studies show a similar effect for feature-based attention for orientation in V4 [131, 222] and motion direction in MT [126]. If we consider that space is just another stimulus dimension to which neurons are sensitive, then the selective routing in terms of orientation and motion may well be performed using a similar mechanism to that defined here.

Investigation of the temporal structure of spike trains

The simulations presented in this thesis were all conducted using a spiking LIF neuron model [103], which was selected because it provides a suitable trade-off between biological realism and computational efficiency. However, the regularity of activity in LIF neurons in response to a constant stimulus drive limits the ability for simulation data from the ARC to be investigated for synchronous or oscillatory properties. The NEF allows other neuron models to be used instead of LIF neurons, and future work may also investigate the effects of attention on the temporal structure of spike trains across populations in the ARC.

Concluding Remarks

The research presented in this thesis has proposed and demonstrated a mechanism by which the selective routing of attended stimuli may be performed in cortex. This proposed ARC model provides a platform for unifying several seemingly disparate forms of attentional modulation, and suggests that they each may be characterized as resulting from a common process for performing selective routing. However, the model contains several speculative proposals, and as more data becomes available regarding neural function, anatomy and physiology, parts of the model may need to be reconsidered. However, the general principles of the model are expected to be relatively unchanged, even with a variety of possible future findings.

As it stands, the ARC model provides several testable predictions, which make the model and its underlying assumptions clearly empirically testable. While the ARC is but a stepping stone to an improved future theory, it has clear advantages over several past theories of attention. Its unique combination of proposing both a large-scale architecture and a biologically plausible implementation, put it in a good position to continue to unify behavioral, physiological, and anatomical data regarding how attentional processing is actually performed in the brain.

Bibliography

- [1] M.M. Adams, P.R. Hof, R. Gattass, M.J. Webster, and L.G. Ungerleider. Visual cortical projections and chemoarchitecture of macaque monkey pulvinar. *Journal of Comparative Neurology*, 419(3), 2000. 20
- [2] S. Ahmad. VISIT: A neural model of covert visual attention. In *Advances in Neural Information Processing Systems*, pages 420–427, 1991. 3, 21, 151
- [3] D.G. Albrecht and D.B. Hamilton. Striate cortex of monkey and cat: contrast response function. *Journal of Neurophysiology*, 48(1):217–237, 1982. 132, 136, 142
- [4] G.A. Alvarez and A. Oliva. The representation of simple ensemble visual features outside the focus of attention. *Psychological Science*, 19(4):392–398, 2008. 70
- [5] R.A. Andersen, R.M. Bracewell, S. Barash, J.W. Gnadt, and L. Fogass. Eye position effects on visual, memory, and saccade-related activity in areas LIP and 7a of macaque. *Journal of Neuroscience*, 10(4):1176–1196, 1990. 50, 72
- [6] R.A. Andersen, G.K. Essick, and R.M. Siegel. Encoding of spatial location by posterior parietal neurons. *Science*, 230(4724):456–458, 1985. 50, 72
- [7] C.H. Anderson and D.C. Van Essen. Shifter circuits: A computational strategy for dynamic aspects of visual processing. *Proceedings of the National Academy of Sciences, USA*, 84:6297–6301, 1987. 21, 37, 151
- [8] C.H. Anderson, H. Nover, and G.C. DeAngelis. Modeling the velocity tuning of macaque MT neurons. *Journal of Vision*, 3(9):404, 2003. 123
- [9] J.C. Anderson and K.A.C. Martin. The synaptic connections between cortical areas V1 and V2 in macaque monkey. *Journal of Neuroscience*, 29(36):11283–11293, 2009. 14
- [10] S. Ardid, X.J. Wang, and A. Compte. An integrated microcircuit model of attentional processing in the neocortex. *Journal of Neuroscience*, 27(32):8486–8495, 2007. 49

- [11] S. Ardid, X.J. Wang, D. Gomez-Cabrero, and A. Compte. Reconciling coherent oscillation with modulation of irregular spiking activity in selective attention: gamma-range synchronization between sensory and executive cortical areas. *Journal of Neuroscience*, 30(8):2856–2870, 2010. 49
- [12] J.S. Baizer, L.G. Ungerleider, and R. Desimone. Organization of visual inputs to the inferior temporal and posterior parietal cortex in macaques. *Journal of Neuroscience*, 11(1):168–190, 1991. 14
- [13] I. Ballesteros-Yanez, R. Benavides-Piccione, G.N. Elston, R. Yuste, and J. DeFelipe. Density and morphology of dendritic spines in mouse neocortex. *Neuroscience*, 138(2):403–9, 2006. 66
- [14] M. Behrmann and M. Moscovitch. Object-centered neglect in patients with unilateral neglect: Effects of left-right coordinates of objects. *Journal of Cognitive Neuroscience*, 6(1):1–16, 1994. 73
- [15] S. Ben Hamed, J.R. Duhamel, F. Bremmer, and W. Graf. Representation of the visual field in the lateral intraparietal area of macaque monkeys: a quantitative receptive field analysis. *Experimental Brain Research*, 140(2):127–144, 2001. 19
- [16] D.B. Bender and M. Youakim. Effect of attentive fixation in macaque thalamus and cortex. *Journal of Neurophysiology*, 85(1):219–234, 2001. 22, 151
- [17] T. Binzegger, R.J. Douglas, and K.A.C. Martin. A quantitative map of the circuit of cat primary visual cortex. *Journal of Neuroscience*, 24(39):8441–53, 2004. 101, 102
- [18] G.J. Blatt, R.A. Andersen, and G.R. Stoner. Visual receptive field organization and cortico-cortical connections of the lateral intraparietal area (area LIP) in the macaque. *Journal of Comparative Neurology*, 299(4):421–445, 1990. 19
- [19] C. Borgers, S. Epstein, and N. Kopell. Background gamma rhythmicity and attention in cortical local circuits: A computational study. *Proceedings of the National Academy of Sciences*, 105(19):7002–7009, 2005. 49
- [20] C. Borgers and N. Kopell. Gamma oscillations and stimulus selection. *Neural Computation*, 20(2):383–414, 2008. 49
- [21] E.C. Boudreau, T.H. Williford, and J.H. Maunsell. Effects of task difficulty and target likelihood in area V4 of macaque monkeys. *Journal of Neurophysiology*, 96(5):2377–2387, 2006. 21

- [22] F. Briggs and E.M. Callaway. Layer-specific input to distinct cell types in layer 6 of monkey primary visual cortex. *Journal of Neuroscience*, 21:3600–3608, 2001. 105, 106
- [23] S.L. Brincat and C.E. Connor. Underlying principles of visual shape selectivity in posterior inferotemporal cortex. *Nature Neuroscience*, 7(8):880–886, 2004. 73, 154
- [24] D.E. Broadbent. *Perception and communication*. Pergamon Press, London, 1958. 5
- [25] A. Buehlmann and G. Deco. The neuronal basis of attention: rate versus synchronization modulation. *Journal of Neuroscience*, 28(30):7679–7686, 2008. 49
- [26] E.A. Buffalo, P. Fries, R. Landman, H. Liang, and R. Desimone. A backward progression of attentional effects in the ventral stream. *Proceedings of the National Academy of Sciences*, 107(1):361–365, 2010. 8, 49, 51, 64, 70, 153
- [27] J. Bullier. Integrated model of visual processing. *Brain Research Reviews*, 36(2-3):96–107, 2001. 15, 67, 69
- [28] D.P. Buxhoeveden and M.F. Casanova. The minicolumn hypothesis in neuroscience. *Brain*, 125:935–951, 2002. 25
- [29] E.L. Cameron, J.C. Tai, and M. Carrasco. Covert attention affects the psychometric function of contrast sensitivity. *Vision Research*, 42:949–967, 2002. 6
- [30] K.R. Cave. Modulation of oscillatory neuronal synchronization by selective visual attention. *Psychological Research*, 62(182–194), 1999. 3
- [31] L. Chelazzi and M. Corbetta. Cortical mechanisms of visuospatial attention in the primate brain. In M.S. Gazzaniga and E. Bizzi, editors, *The New Cognitive Neurosciences*, pages 667–685. The MIT Press, Cambridge, MA, 2 edition, 2000. 20
- [32] L. Chelazzi, E.K. Miller, J. Duncan, and R. Desimone. A neural basis for visual search in inferior temporal cortex. *Nature*, 363:345–348, 1993. 8, 64, 153
- [33] D.B. Chklovskii, T. Schikorski, and C.F. Stevens. Wiring optimization in cortical circuits. *Neuron*, 34(3):341–347, 2002. 102
- [34] A. Compte and X.J. Wang. Tuning curve shift by attention modulation in cortical neurons: a computational study of its mechanisms. *Cerebral Cortex*, 16(6):761–778, 2006. 35
- [35] C.E. Connor, J.T. Gallant, D.C. Preddie, and D. C. Van Essen. Responses in area V4 depend on the spatial relationship between stimulus and attention. *Journal of Neurophysiology*, 75(3):1306–1308, 1996. 35, 50, 56, 72, 108

- [36] C.E. Connor, D.C. Preddie, J.T. Gallant, and D.C. Van Essen. Spatial attention effects in macaque area V4. *Journal of Neuroscience*, 17(9):3201–3214, 1997. 8, 35, 50, 56, 69, 72, 108, 120
- [37] E.P. Cook and J.H. Maunsell. Attentional modulation of motion integration of individual neurons in the middle temporal visual area. *Journal of Neuroscience*, 24(36):7964–7977, 2004. 10
- [38] M. Corbetta, FM Miezin, S. Dobmeyer, GL Shulman, and SE Petersen. Selective and divided attention during visual discriminations of shape, color, and speed: functional anatomy by positron emission tomography. *Journal of Neuroscience*, 11(8):2383–2402, 1991. 22
- [39] F. Crick and C. Koch. Some reflections on visual awareness. In *Cold Spring Harbor symposia on quantitative biology*, volume 55, page 953, 1990. 3, 47
- [40] M. Cynader and N. Berman. Receptive-field organization of monkey superior colliculus. *Journal of Neurophysiology*, 35(2):187–201, 1972. 24
- [41] G. Deco and E.T. Rolls. A neurodynamical cortical model of visual attention and invariant object recognition. *Vision Research*, 44(6):621 – 642, 2004. 153
- [42] R. Desimone, T.D. Albright, C. G. Gross, and C. Bruce. Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience*, 4(8):2051–2062, 1984. 71, 73
- [43] R. Desimone and J. Duncan. Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18(1):193–222, 1995. 9, 41, 42
- [44] R. Desimone and S.J. Schein. Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. *Journal of Neurophysiology*, 57(3):835–868, 1987. 17
- [45] R. Desimone, M. Wessinger, L. Thomas, and W. Schneider. Attentional control of visual perception: cortical and subcortical mechanisms. *Cold Spring Harbor Symposia on Quantitative Biology*, 55:963–971, 1990. 21, 22, 23, 33, 151
- [46] A.A. deSousa, C.C. Sherwood, A. Schleicher, K. Amunts, C.E. MacLeod, P.R. Hof, and K. Zilles. Comparative cytoarchitectural analyses of striate and extrastriate areas in hominoids. *Cerebral Cortex*, 20(4):966–981, 2010. 106
- [47] J.A. Deutsch and D. Deutsch. Attention: Some theoretical considerations. *Psychological Review*, 70:80–90, 1963. 6

- [48] M.C. Dorris, R.M. Klein, S. Everling, and D.P. Munoz. Contribution of the primate superior colliculus to inhibition of return. *Journal of Cognitive Neuroscience*, 14(8):1256–1263, 2002. 24, 25
- [49] R.F. Dougherty, V.M. Koch, A.A. Brewer, B. Fischer, J. Modersitzki, and B.A. Wandell. Visual field representations and locations of visual areas V1/2/3 in human visual cortex. *Journal of Vision*, 3(10):586–98, 2003. 55
- [50] R.J. Douglas and K.A.C. Martin. Neuronal circuits of the neocortex. *Annual Reviews of Neuroscience*, 27:419–51, 2004. 102, 103, 105, 106
- [51] J.R. Duhamel, C.L. Colby, and M.E. Goldberg. The updating of the representation of visual space in parietal cortex by intended eye movements. *Science*, 255(5040):90–92, 1992. 19
- [52] J. Duncan and G.W. Humphreys. Visual search and stimulus similarity. *Psychological Review*, 96(3):433–458, 1989. 6, 32
- [53] M.P. Eckstein, S.S. Shimozaki, and C.K. Abbey. The footprints of visual attention in the Posner cueing paradigm revealed by classification images. *Journal of Vision*, 2(1):25–45, 2002. 6
- [54] H.W. Egeth. Parallel versus serial processes in multidimensional stimulus discrimination. *Perception and Psychophysics*, 1(245–252), 1966. 6
- [55] C. Eliasmith and C.H. Anderson. *Neural Engineering: Computation, Representation, and Dynamics in Neurobiological Systems*. MIT Press, Cambridge, MA, USA, 2003. 11, 40, 51, 79, 80, 148
- [56] G.N. Elston, R. Benavides-Piccione, A. Elston, J. Defelipe, and P.R. Manger. Specialization in pyramidal cell structure in the sensory-motor cortex of the vervet monkey (*Cercopithecus Pygerythrus*). *Neuroscience*, 134(3):1057–1068, 2005. 26, 66
- [57] G.N. Elston and M.G. Rosa. Morphological variation of layer III pyramidal neurones in the occipitotemporal pathway of the macaque monkey visual cortex. *Cerebral Cortex*, 8:278–294, 1998. 55
- [58] G.N. Elston and M.G. Rosa. Pyramidal cells, patches, and cortical columns: a comparative study of infragranular neurons in TEO, TE, and the superior temporal polysensory area of the macaque monkey. *Journal of Neuroscience*, 20:1–5, 2000. 55
- [59] D.J. Felleman and D.C. Van Essen. Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, 1(1):1–47, 1991. 13, 14, 15, 19, 73, 151

- [60] D.J. Felleman, Y. Xiao, and E. McClendon. Modular organization of occipito-temporal pathways: Cortical connections between visual area 4 and visual area 2 and posterior inferotemporal ventral area in macaque monkeys. *Journal of Neuroscience*, 17(9):3185–3200, 1997. 55, 154
- [61] P. Fries, J.H. Reynolds, A.E. Rorie, and R. Desimone. Modulation of oscillatory neuronal synchronization by selective visual attention. *Science*, 291(1560–1563), 2001. 10
- [62] I. Fujita, K. Tanaka, M. Ito, and K. Cheng. Columns for visual features of objects in monkey inferotemporal cortex. *Nature*, 360:343–346, 1992. 18, 74, 153
- [63] K. Funahashi. On the approximate realization of continuous mappings by neural networks. *Neural Networks*, 2(3):183–192, 1989. 29, 59
- [64] P.L.A. Gabbott and P. Somogyi. Quantitative distribution of GABA-immunoreactive neurons in the visual cortex (area 17) of the cat. *Experimental Brain Research*, 61:323–331, 1986. 101
- [65] R. Galambos, S. Makeig, and P.J. Talmachoff. A 40- hz auditory potential recorded from the human scalp. *Proceedings of the National Academy of Sciences, USA*, 78:2643–2647, 1981. 47
- [66] R. Gattass and C.G. Gross. Visual topography of striate projection zone (MT) in posterior superior temporal sulcus of the macaque. *Journal of Neurophysiology*, 46(3):621–638, 1981. 19
- [67] R. Gattass, A. P. Sousa, and C. G. Gross. Visuotopic organization and extent of V3 and V4 of the macaque. *Journal of Neuroscience*, 8(6):1831–1845, June 1988. 13
- [68] K.R. Gegenfurtner, D.C. Kiper, and J.B. Levitt. Functional properties of neurons in macaque area V3. *Journal of Neurophysiology*, 77(4):1906, 1997. 19
- [69] D. George and J. Hawkins. Towards a mathematical theory of cortical micro-circuits. *PLoS computational biology*, 5(10):1–26, 2009. 25, 54, 102
- [70] C.D. Gilbert. Laminar differences in receptive field properties of cells in cat primary visual cortex. *Journal of Physiology*, 268:391–421, 1977. 86
- [71] C.D. Gilbert. Microcircuitry of the visual cortex. *Annual Review of Neuroscience*, 6:217–247, 1983. 20, 102, 103, 105, 106
- [72] V. Goffaux, J. Peters, J. Haubrechts, C. Schiltz, B. Jansma, and R. Goebel. From coarse to fine? spatial and temporal dynamics of cortical face processing. *Cerebral Cortex*, 21(2):467–476, 2011. 68, 73

- [73] M.E. Goldberg, J.W. Bisley, K.D. Powell, and J. Gottlieb. Saccades, salience and attention: the role of the lateral intraparietal area in visual behavior. *Progress in Brain Research*, 155:157, 2006. 19
- [74] J. Gottlieb, M. Kusunoki, and M.E. Goldberg. The representation of visual salience in monkey parietal cortex. *Nature*, 391:481–484, 1998. 33
- [75] C.M. Gray, P. König, A.K. Engel, and W. Singer. Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, 338(6213):334–337, 1989. 47
- [76] M.S.A. Graziano, R.A. Andersen, and R.J. Snowden. Tuning of MST neurons to spiral motions. *Journal of Neuroscience*, 14(1):54–67, 1994. 19
- [77] C. G. Gross, C.E. Rocha-Miranda, and D.B. Bender. Visual properties of neurons in inferotemporal cortex of the macaque. *Journal of Neurophysiology*, 35(1):96–111, 1972. 17, 71
- [78] C.G. Gross. Representation of visual stimuli in inferior temporal cortex. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 335:3–10, 1992. 17
- [79] R.W. Guillery and S.M. Sherman. Thalamic relay functions and their role in cortico-cortical communication: Generalizations from the visual system. *Neuron*, 33(2):163–175, 2002. 14
- [80] D.J. Heeger. Normalization of cell responses in cat striate cortex. *Visual Neuroscience*, 9:181–197, 1992. 43
- [81] D. Heinke and G.W. Humphreys. Attention, spatial representation, and visual neglect: simulating emergent attention and spatial memory in the selective attention for identification model (SAIM). *Psychological Review*, 110(29–87), 2003. 25, 40, 54
- [82] D. Heinke, E. Mavritsaki, A. Backhaus, and M. Kreyling. The selective attention for identification model (SAIM): A framework for closing the gap between the behavioural and neurological levels. In D. Heinke and E. Mavritsaki, editors, *Computational Modelling in Behavioural Neuroscience*, chapter 5, pages 80–106. Psychology Press, New York, 1st edition, 2009. 3, 11, 34, 41, 146, 155
- [83] J. Heinzle, K. Hepp, and K.A.C. Martin. A microcircuit model of the frontal eye fields. *Journal of Neuroscience*, 27(35):9341–9353, 2007. 33
- [84] A.E. Hendrickson, J.R. Wilson, and M.P. Ogren. The neuroanatomical organization of pathways between the dorsal lateral geniculate nucleus and visual cortex in old world and new world primates. *Journal of Comparative Neurology*, 182:123–136, 1978. 103

- [85] J.L. Herrero, M.J. Roberts, L.S. Delicato, M.A. Gieselmann, P. Dayan, and A. Thiele. Acetylcholine contributes through muscarinic receptors to attentional modulation in V1. *Nature*, 454(28):1110–1114, 2008. 3
- [86] T.M. Herrington and J.A. Assad. Temporal sequence of attentional modulation in the lateral intraparietal area and middle temporal area during rapid covert shifts of attention. *Journal of Neuroscience*, 30(9):3287–3296, 2010. 8, 64, 151
- [87] G.E. Hinton. A parallel computation that assigns canonical object-based frames of reference. In *the Seventh International Joint Conference on Artificial Intelligence*, Vancouver, B.C. Canada, 1981. 71
- [88] J.C. Horton and D.L. Adams. The cortical column: a structure without a function. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 360(1456):837–862, Apr 2005. 25
- [89] D.H. Hubel and T.N. Wiesel. Anatomical demonstration of columns in the monkey striate cortex. *Nature*, 221(5182):747–750, 1969. 16
- [90] D.H. Hubel and T.N. Wiesel. Laminar and columnar distribution of geniculocortical fibers in the macaque monkey. *Journal of Comparative Neurology*, 146:421–450, 1972. 103
- [91] D.H. Hubel and T.N. Wiesel. Sequence regularity and geometry of orientation columns in the monkey striate cortex. *Journal of Computational Neuroscience*, 158:267–294, 1974. 25, 26
- [92] D.H. Hubel and T.N. Wiesel. Ferrier lecture: Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 198(1130):1, 1977. 16
- [93] M. Ito, H. Tamura, I. Fujita, and K. Tanaka. Size and position invariance of neuronal responses in monkey inferotemporal cortex. *Journal of Neurophysiology*, 73(1):218–226, 1995. 72
- [94] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 20(11):1254–1259, 1998. 32, 60, 67, 69
- [95] William James. *The Principles of Psychology*, volume 1. Henry Holt, New York, 1890. 1
- [96] E.G. Jones. *The Thalamus*. Plenn Press, New York, 1985. 20, 103

- [97] E.R. Kandel, J.H. Schwartz, and T.M. Jessell. *Principles of neural science*. McGraw-Hill Health Professions Division, New York, 4th edition, 2000. 16, 102, 103, 136, 152
- [98] S. Kastner, D.H. O'Connor, M.M. Fukui, H.M. Fehd, U. Herwig, and M.A. Pinsk. Functional imaging of the human lateral geniculate nucleus and pulvinar. *Journal of Neurophysiology*, 91(1):438–448, 2004. 22, 151
- [99] R. Kimchi. Primacy of wholistic processing and global/local paradigm: a critical review. *Psychological Bulletin*, 112(1):24–38, 1992. 67
- [100] H. Kirchner, E.J. Barbeau, S.J. Thorpe, J. Regis, and C. Liegeois-Chauvel. Ultra-rapid sensory responses in the human frontal eye field region. *Journal of Neuroscience*, 29(23):7599–7606, 2009. 19
- [101] E. Kobatake and K. Tanaka. Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *Journal of Neurophysiology*, 71(3):856, 1994. 18
- [102] C. Koch. Computation and the single neuron. *Nature*, 385:207–210, 1997. 2, 27
- [103] C. Koch. *Biophysics of Computation*. Oxford University Press, Oxford, 1999. 80, 159
- [104] C. Koch. *The quest for consciousness: A neurobiological approach*. Roberts & Company Publishers, Englewood, Colorado, 2004. 14, 60, 74, 102, 103, 154
- [105] C. Koch. Selective visual attention and computational models. <http://www.klab.caltech.edu/cns186/PS/attention-koch.pdf>, March 2004. 34
- [106] C. Koch, T. Poggio, and V. Torre. Retinal ganglion cells: A functional interpretation of dendritic morphology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 298:227–264, 1982. 66
- [107] C. Koch and S. Ullman. Shifts in selective visual attention: towards the underlying neural circuitry. *Human neurobiology*, 4(4):219–227, 1985. 32, 33, 37, 48
- [108] M. Kusunoki, J. Gottlieb, and M.E. Goldberg. the lateral intraparietal area as a salience map: the representation of abrupt onset, stimulus motion, and task relevance. *Vision Research*, 40(10–12):1459–1468, 2000. 33
- [109] D. LaBerge. *Attentional processing: the brain's art of mindfulness*. Harvard University Press, Cambridge, MA, 1995. 14

- [110] D. LaBerge and M.S. Buchsbaum. Positron emission tomographic measurements of pulvinar activity during an attention task. *Journal of Neuroscience*, 10(2):613–619, 1990. 21, 22, 151
- [111] J. Lee and J.H. Maunsell. Attentional modulation of MT neurons with single or multiple stimuli in their receptive fields. *Journal of Neuroscience*, 30(8):3058–3066, 2010. 76
- [112] J. Lee and J.H. Maunsell. The effect of attention on neuronal responses to high and low contrast stimuli. *Journal of Neurophysiology*, 104(2):960–971, 2010. 9, 10, 21, 76, 107, 132, 133, 135, 136, 137, 138, 139, 141, 142, 144, 145, 148, 149
- [113] J.B. Levitt, D.C. Kiper, and J.A. Movshon. Receptive fields and functional architecture of macaque V2. *Journal of Neurophysiology*, 71(6):2517–2542, 1994. 17
- [114] J.W. Lewis and D.C. Van Essen. Mapping of architectonic subdivisions in the macaque monkey, with emphasis on parieto-occipital cortex. *Journal of Comparative Neurology*, 428:79–111, 2000. 73
- [115] X. Li and M.A. Basso. Preparing to move increases the sensitivity of superior colliculus neurons. *Journal of Neuroscience*, 28(17):4561–4577, 2008. 10, 141
- [116] Z. Li. A saliency map in primary visual cortex. *Trends in Cognitive Sciences*, 6:9–16, 2002. 33
- [117] A. Losonczy and J.C. Magee. Integrative properties of radial oblique dendrites in hippocampal CA1 pyramidal neurons. *Neuron*, 50(2):291–307, 2006. 59, 66, 67
- [118] S.J. Luck, L. Chelazzi, S.A. Hillyard, and R. Desimone. Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *Journal of Neurophysiology*, 77(1):24, 1997. 7, 8, 10, 21, 49, 66, 156
- [119] J. Lund. Anatomical organization of macaque monkey striate visual cortex. *Annual Review of Neuroscience*, 11:253–288, 1988. 103, 105, 106
- [120] D.C. Lyon and J.H. Kaas. Evidence for a modified V3 with dorsal and ventral halves in macaque monkeys. *Neuron*, 33(3):453 – 461, 2002. 18
- [121] D.C. Lyon, X. Xu, V.A. Casagrande, J.D. Stefansic, D. Shima, and J.H. Kaas. Optical imaging reveals retinotopic organization of dorsal V3 in New World owl monkeys. *Proceedings of the National Academy of Sciences*, 99(24):15735–15742, 2002. 18
- [122] M. Margulis and C.M. Tang. Temporal integration can readily switch between sub-linear and supralinear. *Journal of Neurophysiology*, 79:2809–2813, 1998. 59, 66

- [123] D. Marr. *Vision. A Computational Investigation into the Human Representation and Processing of Visual Information*. W.H. Freeman and Company, San Francisco, 1982. 67, 69
- [124] D. Marr and H.K. Nishihara. Representation and recognition of the spatial organization of three dimensional structure. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 200:269–294, 1978. 71
- [125] J.C. Martinez-Trujillo and S. Treue. Attentional modulation strength in cortical area MT depends on stimulus contrast. *Neuron*, 35:1–20, 2002. 9, 10, 141, 143
- [126] J.C. Martinez-Trujillo and S. Treue. Feature-based attention increases the selectivity of population responses in primate visual cortex. *Current Biology*, 14:744–751, 2004. 47, 159
- [127] J.H. Maunsell and D.C. Van Essen. Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *Journal of Neurophysiology*, 49(5):1127–1147, 1983. 19
- [128] J.H. Maunsell and D.C. Van Essen. Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. *Journal of Neurophysiology*, 49(5):1148–1167, 1983. 19
- [129] J.H. Maunsell and D.C. Van Essen. The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *Journal of Neuroscience*, 3(12):2563–2586, 1983. 19, 106
- [130] P.J. May. The mammalian superior colliculus: laminar structure and connections. *Progress in Brain Research*, 151:321–378, 2006. 24
- [131] C.J. McAdams and J.H. Maunsell. Tuning functions of single neurons in macaque cortical area V4. *Journal of Neuroscience*, 19(1):431–441, 1999. 8, 9, 35, 66, 159
- [132] K. McAlonan, J. Cavanaugh, and R.H. Wurtz. Guarding the gateway to cortex with attention in visual thalamus. *Nature*, 456(7220):391–4, 2008. 8, 49
- [133] M. McCloskey and B. Rapp. Attention-referenced visual representations: evidence from impaired visual localization. *Journal of Experimental Psychology*, 26(3):917–933, 2000. 73
- [134] W.S. McCullough and W. Pitts. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biology*, 5(4):115–133, 1943. 27

- [135] P. McLeod, J. Driver, and J. Crisp. Visual search for a conjunction of movement and form is parallel. *Nature*, 332(6160):154–155, 1988. 32
- [136] W.P. Medendorp, H.C. Goltz, and T. Vilis. Remapping the remembered target location for anti-saccades in human posterior parietal cortex. *Journal of Neurophysiology*, 94(1):734–740, 2005. 19
- [137] M. Megías, Z. Emri, T.F. Freund, and I. Gulyás. Total number and distribution of inhibitory and excitatory synapses on hippocampal CA1 pyramidal cells. *Neuroscience*, 102(3):527–40, 2001. 29, 86
- [138] A.D. Mehta, I. Ulbert, and C.E. Schroeder. Intermodal selective attention in monkeys. I: distribution and timing of effects across visual areas. *Cerebral Cortex*, 10(4):343–358, 2000. 8, 49, 51, 64, 153
- [139] B.W. Mel. Synaptic integration in an excitable dendritic tree. *Journal of Neurophysiology*, 70(3):1086–1101, 1993. 66
- [140] B.W. Mel. Why Have Dendrites? A Computational Perspective. In G. Stuart, N. Spruston, and M. Häusser, editors, *Dendrites*, chapter 16, pages 421–440. Oxford University Press, Oxford, 2nd edition, 2006. 26, 27, 29
- [141] B.W. Mel and C. Koch. Sigma-pi learning: On radial basis functions and cortical associative learning. In D.S. Touretzky, editor, *Advances in Neural Information Processing Systems*. Morgan Kaufmann, San Mateo, CA, 2 edition, 1990. 29, 66
- [142] M. Migliore and G.M. Shepherd. Emerging rules for the distributions of active dendritic conductances. *Nature Reviews Neuroscience*, 3(5):362–370, 2002. 66
- [143] J. Mishra, J. Fellous, and T.J. Sejnowski. Selective attention through phase relationship of excitatory and inhibitory input synchrony in a model cortical neuron. *Neural Networks*, 19(9):1329–1346, 2006. 49
- [144] J.F. Mitchell, K.A. Sundberg, and J.H. Reynolds. Differential attention-dependent response modulation across cell classes in macaque visual area V4. *Neuron*, 55(1):131–41, 2007. 10, 26, 66
- [145] C. Möller, J. Lücke, J. Zhu, P.M. Faustmann, and C. von der Malsburg. Glial cells for information routing? *Cognitive Systems Research*, 8(1):28–35, 2007. 3
- [146] T. Moore, K.M. Armstrong, and M. Fallah. Visuomotor origins of covert spatial attention. *Neuron*, 40(4):671–683, 2003. 24, 151

- [147] J. Moran and R. Desimone. Selective attention gates visual processing in the extrastriate cortex. *Science*, 229(4715):782–784, 1985. 7, 8, 10, 17, 21, 43, 66, 71, 154, 156
- [148] J.S. Morris, K.J. Friston, and R.J. Dolan. Neural responses to salient visual stimuli. *Proceedings: Biological Sciences*, 264(1382):769–775, 1997. 22, 151
- [149] B. Motter. Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *Journal of Neurophysiology*, 70(3):909–919, 1993. 8, 49
- [150] V. Mountcastle. Modality and topographic properties of single neurons of cat’s somatic sensory cortex. *Journal of Neurophysiology*, 20:408–434, 1957. 25
- [151] V. Mountcastle. An organizing principle for cerebral function: the unit model and the distributed system. In G. Edelman and V. Mountcastle, editors, *The Mindful Brain*. MIT Press, Cambridge, Mass., 1978. 25, 54
- [152] V. Mountcastle. The columnar organization of the neocortex. *Brain*, 120:701–722, 1997. 25
- [153] J.A. Movshon and W.T. Newsome. Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. *Journal of Neuroscience*, 16(23):7733–7741, 1996. 106
- [154] J.R. Müller, M.G. Philiastides, and W.T. Newsome. Microstimulation of the superior colliculus focuses attention without moving the eyes. *Proceedings of the National Academy of Sciences*, 102(3):524, 2005. 24
- [155] D.P. Munoz and R.H. Wurtz. Saccade-related activity in monkey superior colliculus. II. Spread of activity during saccades. *Journal of Neurophysiology*, 73(6):2334, 1995. 25
- [156] K. Nakamura and C.L. Colby. Updating of the visual representation in monkey striate and extrastriate cortex during saccades. *Proceedings of the National Academy of Sciences*, 99(6):4026, 2002. 19
- [157] K. Nakayama and G.H. Silverman. Serial and parallel processing of visual feature conjunctions. *Nature*, 320(6059):264–265, 1986. 32
- [158] D. Navon. Forest before trees: the precedence of global features in visual perception. *Cognitive psychology*, 9:353–383, 1977. 67
- [159] U. Neisser. *Cognitive Psychology*. Appleton-Century-Crofts, New York, 1967. 67

- [160] J.S. Nettleton and W.J. Spain. Linear to supralinear summation of AMPA-mediated EPSPs in neocortical pyramidal neurons. *Journal of Neurophysiology*, 83:3310–3322, 2000. 59, 66
- [161] E. Niebur. Temporal tagging of attended objects. *Proceedings of the National Academy of Sciences, USA*, 106(8):2479–2480, 2009. 47
- [162] E. Niebur and C. Koch. A model for the neuronal implementation of selective visual attention based on temporal correlation among neurons. *Journal of Computational Neuroscience*, 1(1):141–158, 1994. 11, 20, 21, 25, 34, 48, 76, 146, 151, 155
- [163] E. Niebur, C. Koch, and C. Rosin. An oscillation-based model for the neuronal basis of attention. *Vision Research*, 33(18):2789–2802, 1993. 3, 11, 25, 47, 76, 146, 155
- [164] A. Oliva and P. G. Schyns. Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognitive psychology*, 34:72–107, 1997. 67
- [165] A. Oliva and P. G. Schyns. Diagnostic colors mediate scene recognition. *Cognitive psychology*, 41:176–210, 2000. 67
- [166] A. Oliva and A. Torralba. Building the gist of a scene: The role of global image features in recognition. *Progress in Brain Research*, 155:23–36, 2006. 67
- [167] B.A. Olshausen. *Neural Routing Circuits for Forming Invariant Representations of Visual Objects*. PhD thesis, California Institute of Technology, 1994. 37, 38
- [168] B.A. Olshausen, C.H. Anderson, and D.C. Van Essen. A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. *Journal of Neuroscience*, 13(11):4700–4719, 1993. 3, 11, 25, 31, 32, 33, 34, 35, 37, 45, 49, 50, 54, 71, 72, 75, 146, 147, 154, 155
- [169] B.A. Olshausen, C.H. Anderson, and D.C. Van Essen. A multiscale dynamic routing circuit for forming size- and position-invariant object representations. *Journal of Computational Neuroscience*, 2(1):45–62, 1995. 21, 34, 35, 37, 40, 49, 71, 75, 147, 151, 154, 155
- [170] C.R. Olson. Brain representation of object-centered space in monkeys and humans. *Annual Review of Neuroscience*, 26:331–354, 2003. 71
- [171] H. Op De Beeck and R. Vogels. Spatial sensitivity of macaque inferior temporal neurons. *Journal of Comparative Neurology*, 426:505–518, 2000. 56, 108

- [172] S.E. Palmer. Hierarchical structure in perceptual representation. *Cognitive psychology*, 9:441–474, 1977. 67
- [173] S.E. Palmer. *Vision Science : Photons to Phenomenology*. MIT Press, 1999. 71
- [174] C. Parisien, C.H. Anderson, and C. Eliasmith. Solving the problem of negative synaptic weights in cortical models. *Neural Computation*, 20:1473–1494, 2008. 98
- [175] D.M. Parker, J.R. Lishman, and J. Hughes. Evidence for the view that temporospatial integration in vision is temporally anisotropic. *Perception*, 26(9):1169–1180, 1997. 68
- [176] T. Pasternak, J.W. Bisley, and D. Calkins. Visual processing in the primate brain. In M. Gallagher and Nelson R.J., editors, *Handbook Of Psychology - Vol 03 - Biological Psychology*, pages 139–185. John Wiley & Sons, Inc., Hoboken, New Jersey., 2003. 16, 19
- [177] A. Pasupathy and C.E. Connor. Responses to contour features in macaque area V4. *Journal of Neurophysiology*, 82(5):2490–2502, 1999. 17, 73, 154
- [178] D. Perrett, E.T. Rolls, and W. Caan. Visual neurones responsive to faces in the monkey temporal cortex. *Experimental Brain Research*, 47:329–342, 1982. 72
- [179] S.E. Petersen, D.L. Robinson, and W. Keys. Pulvinar nuclei of the behaving rhesus monkey: visual responses and their modulation. *Journal of Neurophysiology*, 54(4):867, 1985. 19, 20, 21, 39, 151
- [180] S.E. Petersen, D.L. Robinson, and J.D. Morris. Contributions of the pulvinar to visual spatial attention. *Neuropsychologia*, 25(1A):97 – 105, 1987. 20, 21, 22, 151
- [181] P. Poirazi, T. Brannon, and B.W. Mel. Pyramidal neuron as two-layer neural network. *Neuron*, 37(6):989–99, 2003. 27, 28, 29, 59, 67, 86
- [182] A. Polsky, B.W. Mel, and J. Schiller. Computational subunits in thin dendrites of pyramidal cells. *Nature Neuroscience*, 7(6):621–627, 2004. 27, 28, 29, 59, 66, 67
- [183] M.I. Posner. Orienting of attention. *The Quarterly Journal of Experimental Psychology*, 32(1):3–25, 1980. 3, 6
- [184] M.I. Posner and S.E. Petersen. The attention system of the human brain. *Annual Review of Neuroscience*, 13(1):25–42, 1990. 21, 24, 151
- [185] R.D. Rafal and M.I. Posner. Deficits in human visual spatial attention following thalamic lesions. *Proceedings of the National Academy of Sciences*, 84(20):7349–7353, 1987. 22, 151

- [186] G.H. Recanzone and R.H. Wurtz. Effects of attention on MT and MST neuronal activity during pursuit initiation. *Journal of Neurophysiology*, 83(2):777, 2000. 7, 156
- [187] D. Ress, S. Zughni, S. Katyal, B. Rokers, A. Huk, L. Cormack, and C. Greene. Topography of covert attention in human superior colliculus. *Neuroimage*, 47:64–64, 2009. 24
- [188] J.H. Reynolds, L. Chelazzi, and R. Desimone. Competitive mechanisms subserve attention in macaque areas V2 and V4. *Journal of Neuroscience*, 19(5):1736–1753, 1999. 3, 7, 11, 17, 31, 34, 41, 42, 43, 45, 50, 51, 76, 122, 131, 144, 146, 155, 156
- [189] J.H. Reynolds and D.J. Heeger. The normalization model of attention. *Neuron*, 61(2):168–185, 2009. 3, 7, 11, 35, 43, 44, 47, 50, 51, 76, 122, 131, 144, 146, 155, 156
- [190] J.H. Reynolds, T. Pasternak, and R. Desimone. Attention increases sensitivity of V4 neurons. *Neuron*, 26:703–714, 2000. 9, 10, 136, 141
- [191] D.L. Robinson and C. Kertzman. Covert orienting of attention in macaques. III. contributions of the superior colliculus. *Journal of Neurophysiology*, 74(2):713–721, 1995. 20
- [192] D.L. Robinson and S.E. Petersen. The pulvinar and visual salience. *Trends in Neuroscience*, 15:127–132, 1992. 20, 33, 151
- [193] K.S. Rockland and G.W. Drash. Collateralized divergent feedback connections that target multiple cortical areas. *Journal of Comparative Neurology*, 373:529–548, 1996. 103, 105
- [194] K.S. Rockland and D.N. Pandya. Laminar origins and terminations of cortical connections of the occipital lobe in the rhesus monkey. *Brain Research*, 179(1):3–20, Dec 1979. 103, 105, 106
- [195] A.W. Roe. Modular complexity of area V2 in the macaque monkey. In J.H. Kaas and C.E. Collins, editors, *The Primate Visual System*, pages 109–139. CRC Press LLC, Cambridge, MA, USA, 2004. 17
- [196] P.R. Roelfsema, V.A.F. Lamme, and H. Spekreijse. Object-based attention in the primary visual cortex of the macaque monkey. *Nature*, 395:376–381, 1998. 8, 49, 64, 153
- [197] E.T. Rolls, N.C. Aggelopoulos, and F. Zheng. The receptive fields of inferior temporal cortex neurons in natural scenes. *Journal of Neuroscience*, 23(1):339–348, 2003. 70

- [198] E.T. Rolls and G.C. Baylis. Size and contrast have only small effects on the responses to faces of neurons in the cortex of the superior temporal sulcus of the monkey. *Experimental Brain Research*, 65:38–48, 1986. 72
- [199] E.T. Rolls and G. Deco. *Computational neuroscience of vision*. Oxford University Press Oxford, 2002. 15, 17, 19, 42
- [200] E.T. Rolls and G. Deco. Attention in natural scenes: Neurophysiological and computational bases. *Neural Networks*, 19:1383–1394, 2006. 33, 153
- [201] L.M. Romanski, M. Giguere, J.F. Bates, and P.S. Goldman-Rakic. Topographic organization of medial pulvinar connections with the prefrontal cortex in the rhesus monkey. *Journal of Comparative Neurology*, 379(3), 1997. 21
- [202] K.S. Saleem, W. Suzuki, K. Tanaka, and T. Hashikawa. Connections between anterior inferotemporal cortex and superior temporal sulcus regions in the macaque monkey. *Journal of Neuroscience*, 20(13):5083–5101, 2000. 73, 74, 154
- [203] K.S. Saleem, K. Tanaka, and K.S. Rockland. Specific and columnar projection from area TEO to TE in the macaque inferotemporal cortex. *Cerebral Cortex*, 3:454–464, 1995. 74, 154
- [204] E. Salinas and L.F. Abbott. Invariant visual responses from attentional gain fields. *Journal of Neurophysiology*, 77(6):3267–3272, 1997. 3, 11, 35, 36, 43, 45, 51, 60, 74, 75, 76, 145, 155
- [205] E. Salinas and L.F. Abbott. Coordinate transformations in the visual system: How to generate gain fields and what to compute with them. *Progress in Brain Research*, 130:175–190, 2001. 49, 50
- [206] E. Salinas and P. Thier. Gain modulation: A major computational principle of the central nervous system. *Neuron*, 27:15–21, 2000. 35
- [207] S.J. Schein and R. Desimone. Spectral properties of V4 neurons in the macaque. *Journal of Neuroscience*, 10(10):3369–3389, 1990. 17
- [208] J. Schiller, G. Major, H.J. Koester, and Y. Schiller. NMDA spikes in basal dendrites of cortical pyramidal neurons. *Nature*, 404(6775):285–289, 2000. 27
- [209] M.T. Schmolesky, Y. Wang, D.P. Hanes, K.G. Thompson, S. Leutgeb, J.D. Schall, and A.G. Leventhal. Signal timing across the macaque visual system. *Journal of Neurophysiology*, 79(6):3272–3278, 1998. 19

- [210] P. G. Schyns and A. Oliva. From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science*, 5(2):195–200, 1994. 68, 69
- [211] L.D. Selemon, A. Begović, and P. Rakic. Selective reduction of neuron number and volume of the mediodorsal nucleus of the thalamus in macaques following irradiation at early gestational ages. *Journal of Comparative Neurology*, 515(4):454–64, 2009. 39
- [212] T. Serre, L. Wolf, S. Bileschi, M. Riesenhuber, and T. Poggio. Robust object recognition with cortex-like mechanisms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(3):411–426, 2007. 60
- [213] M.N. Shadlen and W.T. Newsome. Noise, neural codes and cortical organization. *Current Opinion in Neurobiology*, 4:569–579, 1994. 49
- [214] S. Shipp. The functional logic of cortico-pulvinar connections. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 358(1438):1605–1624, 2003. 21, 105
- [215] S. Shipp. The brain circuitry of attention. *Trends in Cognitive Sciences*, 8(5):223–230, 2004. 151
- [216] R. Silver, K. Boahen, S. Grillner, N. Kopell, and K. L. Olsen. Neurotech for neuroscience: Unifying concepts, organizing principles, and emerging tools. *Journal of Neuroscience*, 27(44):807–819, 2007. 157
- [217] L.H. Snyder, K.L. Grieve, P. Brotchie, and R.A. Andersen. Separate body- and world-referenced representations of visual space in parietal cortex. *Nature*, 394(5448–5465), 1998. 50, 72
- [218] A. Soltani and C. Koch. Visual saliency computations: mechanisms, constraints, and the effect of feedback. *Journal of Neuroscience*, 30(38):12831–12843, 2010. 3
- [219] D.L. Sparks and R. Hartwich-Young. The deep layers of the superior colliculus. *Reviews of oculomotor research*, 3:213–255, 1989. 25
- [220] D.L. Sparks, C. Lee, and W.H. Rohrer. Population coding of the direction, amplitude, and velocity of saccadic eye movements by neurons in the superior colliculus. In *Cold Spring Harbor symposia on quantitative biology*, volume 55, pages 805–811. Cold Spring Harbor Laboratory Press, 1990. 25

- [221] G. Sperling, M. Landy, Y. Cohen, and M. Pavel. Intelligible encoding of ASL image sequences at extremely low information rates. *Computer Vision, Graphics and Image Processing*, 31:335–391, 1985. 72
- [222] H. Spitzer, R. Desimone, and J. Moran. Increased attention enhances both behavioral and neuronal performance. *Science*, 240(4850):338, 1988. 10, 21, 159
- [223] N. Spruston. Pyramidal neurons: dendritic structure and synaptic integration. *Nature*, 9(March):206–221, 2008. 66, 67
- [224] N. Spruston and W.L. Kath. Dendritic arithmetic. *Nature Neuroscience*, 7(6):567–9, 2004. 67
- [225] L.R. Squire, J.L. Roberts, N.C. Spitzer, M.J. Zigmond, S.K. McConnell, and F.E. Bloom. *Fundamental neuroscience*. Academic Press, 3 edition, 2008. 18
- [226] A.P. Sripati and C.R. Olson. Representing the forest before the trees: a global advantage effect in monkey inferotemporal cortex. *Journal of Neuroscience*, 29(24):7788–7796, 2009. 68, 73
- [227] I. Stepniewska. The pulvinar complex. In J.H. Kaas and C.E. Collins, editors, *The Primate Visual System*, pages 53–80. CRC Press LLC, Cambridge, MA, USA, 2004. 20, 21
- [228] T.C. Stewart. *A methodology for computational cognitive modelling*. PhD thesis, Carleton University Institute of Cognitive Science, 2007. 4
- [229] T.C. Stewart and R.L. West. Testing for equivalence: A methodology for computational cognitive modelling. *Journal of Artificial General Intelligence*, 2(2):69–87, 2010. 129
- [230] B. Stricanne, RA Andersen, and P. Mazzoni. Eye-centered, head-centered, and intermediate coding of remembered sound locations in area LIP. *Journal of Neurophysiology*, 76(3):2071–2076, 1996. 19
- [231] G.J. Stuart and M. Häusser. Dendritic coincidence detection of EPSPs and action potentials. *Nature Neuroscience*, 4(1):63–71, 2001. 67
- [232] M. Sur, P.E. Garraghty, and C.J. Bruce. Somatosensory cortex in macaque monkeys: laminar differences in receptive field size in areas 3b and 1. *Brain Research*, 342(2):391–395, 1985. 86
- [233] K. Tanaka. Neuronal mechanisms of object recognition. *Science*, 262(5134):685–688, 1993. 74, 153

- [234] K. Tanaka. Columns for complex visual object features in the inferotemporal cortex: Clustering of cells with similar but slightly different stimulus selectivities. *Cerebral Cortex*, 13:90–99, 2003. 154
- [235] Y. Tang and C. Eliasmith. Deep networks for robust visual recognition. In J. Furnkranz and T. Joachims, editors, *Proceedings of the 28th international conference on machine learning*, 2010. 71, 158
- [236] H. Tanigawa, Q.X Wang, and I. Fujita. Organization of horizontal axons in the inferior temporal cortex and primary visual cortex of the macaque monkey. *Cerebral Cortex*, 15(12):1887–1899, 2005. 55
- [237] M.J. Tarr and S. Pinker. Mental rotation and orientation-dependence in shape recognition. *Cognitive psychology*, 21:233–282, 1989. 73
- [238] M.J. Tarr and S. Pinker. When does human object recognition use a viewer-centered reference frame? *Psychological Science*, 1(4):253–256, 1990. 73
- [239] K.G. Thompson, D.P. Hanes, N.P. Bichot, and J.D. Schall. Perceptual and motor processing stages identified in the activity of macaque frontal eye field neurons during visual search. *Journal of Neurophysiology*, 76(6):4040–4055, 1996. 19, 33
- [240] A.M. Thomson and A.P. Bannister. Interlaminar connections in the neocortex. *Cerebral Cortex*, 13:5–14, 2003. 102, 103, 105, 152
- [241] A.M. Thomson and C. Lamy. Functional maps of neocortical local circuitry. *Frontiers in Neuroscience*, 1(1):19–42, 2007. 102, 103, 105, 106
- [242] P.H. Tiesinga. Stimulus competition by inhibitory interference stimulus competition by inhibitory interference. *Neural Computation*, 17:2421–2453, 2005. 49
- [243] A. Toet, H.L. van Eekhout, M.P. and Simons, and J.J. Koenderink. Scale invariant features of differential spatial displacement discrimination. *Vision Research*, 27(3):441–451, 1987. 72
- [244] A.M. Treisman and G. Gelade. A feature-integration theory of attention. *Cognitive psychology*, 12(1):97–136, 1980. 31
- [245] S. Treue and J.C. Martinez-Trujillo. Feature-based attention influences motion processing gain in macaque visual cortex. *Nature*, 399:575–579, 1999. 3, 9, 35, 46, 47, 107, 121, 122, 123, 124, 125, 127, 128, 129, 130, 134, 144, 145, 146, 148, 149, 156
- [246] S. Treue and J.H. Maunsell. Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature*, 382:539–541, 1996. 7, 156

- [247] S. Treue and J.H. Maunsell. Effects of attention on the processing of motion in macaque middle temporal and medial superior temporal visual cortical areas. *Journal of Neuroscience*, 19(17):7591–7602, 1999. 7, 76, 156
- [248] A.M. Triesman. Selective attention in man. *British Medical Bulletin*, 20(1):12–16, 1964. 5
- [249] W. Tryon. Evaluating statistical difference, equivalence, and indeterminacy using inferential confidence intervals: An integrated alternative method of conducting null hypothesis statistical tests. *Psychological Methods*, 6(4):371–386, 2001. 129
- [250] D.Y. Tsao, W.A. Freiwald, T.A. Knutsen, J.B. Mandeville, and R.B.H. Tootell. Faces and objects in macaque cerebral cortex. *Nature Neuroscience*, 6(9):989–995, 2003. 18
- [251] D.Y. Tsao, W.A. Freiwald, R.B.H. Tootell, and M.S. Livingstone. A cortical region consisting entirely of face-selective cells. *Science*, 311(5761):670–674, 2006. 18, 73
- [252] J.K. Tsotsos. Toward a computational model of visual attention. In T.V. Pappas, C. Chubb, A. Gorea, and E. Kowler, editors, *Early vision and beyond*, pages 207–218. MIT Press, Cambridge, Mass., 1995. 153
- [253] J.K. Tsotsos, S.M. Culhane, W.Y. Kei Wai, Y. Lai, N. Davis, and F. Nufflo. Modeling visual attention via selective tuning. *Artificial intelligence*, 78:507–545, 1995. 32, 34, 42
- [254] K. Tsunoda, Y. Yamane, M. Nishizaki, and M. Tanifuji. Complex objects are represented in macaque inferotemporal cortex by the combination of feature columns. *Nature Neuroscience*, 4(8):832–838, 2001. 154
- [255] L.G. Ungerleider, T.W. Galkin, and M. Mishkin. Visuotopic organization of projections from striate cortex to inferior and lateral pulvinar in rhesus monkey. *Journal of Comparative Neurology*, 217(2):137–157, 1983. 20, 21
- [256] L.G. Ungerleider and T. Pasternak. Ventral and dorsal cortical processing streams. In L.M. Chalupa and J.S. Werner, editors, *The visual neurosciences*, pages 541–562. MIT Press, Cambridge, MA, 2004. 17, 18
- [257] D. Van Essen, B.A. Olshausen, C.H. Anderson, and J.T. Gallant. Pattern recognition, attention, and information bottlenecks in the primate visual system. In *Proc. SPIE Conf. on Visual Information Processing: From Neurons to Chips*, volume 1473, pages 17–28, 1991. 72, 154

- [258] D.C. Van Essen and C.H. Anderson. Information processing strategies and pathways in the primate visual system. In S.F. Zornetzer, J.L. Davis, and C. Lau, editors, *An introduction to neural and electronic networks*, pages 45–76. Academic Press, Orlando, FL, 1995. 102, 137, 142
- [259] D.C. Van Essen and S.M. Zeki. The topographic organization of rhesus monkey prestriate cortex. *Journal of Physiology*, 277(1):193–226, 1978. 18
- [260] R. Vandenberghe, J. Duncan, P. Dupont, R. Ward, J.B. Poline, G. Bormans, J. Michiels, L. Mortelmans, and G.A. Orban. Attention to one or two features in left or right visual field: a positron emission tomography study. *Journal of Neuroscience*, 17(10):3739–3750, 1997. 22, 151
- [261] D. Walther and C. Koch. Modeling attention to salient proto-objects. *Neural Networks*, 19(9):1395–1407, 2006. 33
- [262] R. Ward and S. Danziger. Selective attention and response control following damage to the human pulvinar. In G.W. Humphreys and M.J. Riddoch, editors, *Attention in Action: Advances from Cognitive Neuroscience*, volume 1, pages 325–350. Psychology Press, New York, 2004. 23, 151
- [263] R. Ward, S. Danziger, V. Owen, and R. Rafal. Deficits in spatial coding and feature binding following damage to spatiotopic maps in the human pulvinar. *Nature Neuroscience*, 5(2):99–100, 2002. 23, 151
- [264] H. Wässle, U. Grünert, J. Röhrenbeck, and B.B. Boycott. Cortical magnification factor and the ganglion cell density of the primate retina. *Nature*, 341:643–646, 1989. 14
- [265] M. Watanabe, H. Tanaka, T. Uka, and I. Fujita. Disparity-selective neurons in area V4 of macaque monkeys. *Journal of Neurophysiology*, 87(4):1960–1973, 2002. 17
- [266] T. Williford and J.H. Maunsell. Effects of spatial attention on contrast response functions in macaque area V4. *Journal of Neurophysiology*, 96:40–54, 2006. 10, 136, 137, 141, 142
- [267] H.R. Wilson. Non-Fourier cortical processes in texture, form, and motion perception. In P.S. Ulinski and E.G. Jones, editors, *Cerebral Cortex*, pages 445–478. Kluwer Academic/Plenum Publishers, New York, 1999. 5
- [268] A.K. Wiser and E.M. Callaway. Contributions of individual layer 6 pyramidal neurons to local circuitry in macaque primary visual cortex. *Journal of Neuroscience*, 16:2724–39, 1996. 105, 106

- [269] J.M. Wolfe. Visual search. In H. Pashler, editor, *Attention*. University College London Press, London, 1998. 1, 6
- [270] J.M. Wolfe and T.S Horowitz. What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, 5:1–7, 2004. 6
- [271] P. Wolfrum and C. Von Der Malsburg. What is the optimal architecture for visual information routing? *Neural Computation*, 19:3293–3309, 2007. 55
- [272] T. Womelsdorf, K. Anton-Erxleben, and S. Treue. Receptive field shift and shrinkage in macaque middle temporal area through attentional gain modulation. *Journal of Neuroscience*, 28(36):8934–8944, 2008. 8, 11, 36, 37, 50, 51, 69, 72, 74, 75, 107, 108, 109, 110, 111, 112, 113, 115, 117, 119, 120, 122, 125, 134, 144, 148, 150, 155
- [273] R. H. Wurtz, Basso M. A., Paré M., and M. A. Sommer. The superior colliculus and the cognitive control of movement. In M.S. Gazzaniga, editor, *The New Cognitive Neurosciences*, pages 573–587. The MIT Press, Cambridge, MA, 2000. 24, 25
- [274] Y. Yamane, K. Tsunoda, M. Matsumoto, A.N. Phillips, and M. Tanifuji. Representation of the spatial relationship among object parts by neurons in macaque inferotemporal cortex. *Journal of Neurophysiology*, 96(6):3147–3156, 2006. 154
- [275] S. Yantis, J. Schwarzbach, J.T. Serences, R.L. Carlson, M.A. Steinmetz, J.J. Pekar, and S.M. Courtney. Transient neural activity in human parietal cortex during spatial attention shifts. *Nature Neuroscience*, 5(10):995–1002, 2002. 22, 151
- [276] Y. Yeshurun and M. Carrasco. Spatial attention improves performance in spatial resolution tasks. *Vision Research*, 39:293–306, 1999. 6
- [277] Y. Yeshurun, B. Montagna, and M. Carrasco. On the flexibility of sustained attention and its effects on a texture segmentation task. *Vision Research*, 48(1):80–95, 2008. 2