Neural Theory and Model of Selective Visual Attention and 2D Shape Recognition in Visual Clutter

by

Peter Lozo, B.Sc.

A thesis submitted in fulfilment of the requirement for the degree of

Doctor of Philosophy

The University of Adelaide
Faculty of Engineering
Department of Electrical and Electronic Engineering

June 1996

Dedication

I dedicate this thesis to my wife Isabel, with whose immense patience and understanding I was able to pursue and complete this research; and to our children, Troy and Carlia, who have eagerly waited till the end of my student days so that we can again spend leisure time together.

PL

Contents

Abstract

Extensive psychophysical and the more recent neurophysiological data from single cellular recordings suggest that selective attention and memory guided processing are some of the key properties of the primate visual brain that endows it with cognitive visual abilities that have not yet been matched by traditional artificial intelligence nor by the current artificial neural network models of learning and pattern recognition. Most neural network models of object and pattern recognition either ignore the mechanism of selective attention or are based on feedforward processes that ignore the role of the feedback pathways and the established memory, which therefore limits their application to simple visual scenarios.

This thesis proposes a neural theory, Selective Attention Adaptive Resonance Theory, and a neuro-engineered solution to selective visual attention, memory guided processing and illumination invariant recognition of complete (unoccluded) but distorted 2D shapes of 3D objects in cluttered visual images. We propose a family of modulated competitive neural layers and neuroengineering design principles for the design of multi-layered competitive 2-D neural circuits whose stability and success depends on feedforward-feedback interactions. The proposed feedback pathways and the top-down modulatory processes simultaneously supervise and stabilise the circuit dynamics, selectively retune the signal transmission gains and the filtering characteristics of the lower layers to enable unsupervised learning and recognition of 2D shapes obtained from unoccluded 3D objects in cluttered images. We propose neural circuits and networks that are capable of self-regulated attentional learning, selective attention and memory guided processing, autonomous detection of novelty/familiarity, distortion and illumination invariant recognition of familiar 2D shapes of real objects in cluttered images.

We conclude that flexible design principles that are based on feedforward-feedback interactions in a closed-loop real-time competitive neural circuit whose modulatory

mechanisms can dynamically retune the signal transmission gains and the cellular receptive field profiles at various stages of processing overcomes some of the problems and limitations that are faced by the rigid architecture of the current artificial neural networks. The neuro-engineeering design principles, mechanisms and circuits as proposed in the thesis provide a new and robust method for solving some of the most difficult problems in visual object recognition that are currently not well handled by the state-of-the-art artificial neural networks and the more conventional computer vision systems. These design principles also open new avenues for further research into more advanced modelling of cognitive and perceptual real-time artificial neural systems that use selective information processing.

DECLARATION

This thesis contains no material which has been accepted for the award of any other degree or diploma in any University or other tertiary institution, and to the best of the author's knowledge and belief contains no material previously published or written by another person, except where due reference has been made in the text.

Should the thesis be accepted for the award of the Degree, the author hereby consents to this copy, when deposited in the University Library, being made available for loan and photocopying.

Date 10/6/96

Acknowledgements

The neural theory and the neuro-engineering design principles, mechanisms, neural circuits and network architectures presented in this thesis have their major mathematical and scientific foundation in the work of Dr. Stephen Grossberg, who over the past three and a half decades has produced the most remarkable theories and mathematical tools to model brain functions. I was fortunate to have attended the course on Real-time Neural Networks at Wang Institute of Boston University, in May 1991 and acquainted myself with these tools and theories. I was truly fascinated by the Adaptive Resonance Theory (ART), although at the time I had no notion that some two years later it was going to have a huge impact on my way of thinking. I thank the staff of Centre for Adaptive Systems, Boston University, for presenting an excellent course and for sharing their theories and models.

The second crucial factor that has influenced me happened soon after at a conference on Neural Networks for Vision and Image Processing when Dr. Robert Desimone presented a paper describing the effects of selective attention on the receptive field profiles of cells in visual area V4 and inferior temporal cortex of a monkey. It was almost two years later when I read about other experimental results on selective attention, did I realise that Adaptive Resonance Theory can be reformulated to explain these experimental results. Hence my proposal of Selective Attention Adaptive Resonance Theory (SAART). SAART does not challenge the existing ART principles but elevates ART to a new paradigm that offers a more detailed account of neurobiological mechanisms, circuit dynamics, massive feedback pathways in the brain and self-organised real-time learning in a complex visual sensory environment. The success of the new theory as physical and explanatory theory of certain brain functions is thus largely due to the major paradigm shift that Dr. Grossberg has caused by introducing the Adaptive Resonance Theory.

An opportunity to develop Selective Attention Adaptive Resonance Theory and the associated neuro-engineering design principles presented itself when I was granted an approval by my employer, Defence Science and Technology Organisation (DSTO), to investigate neural mechanisms for robust Automatic Target Recognition systems as part of a Ph.D. research project at the Department of Electrical and Electronic Engineering, University of Adelaide. I thank my work colleagues and supervisors, Mr. Ray Johnson and Dr. Nanda Nandagopal for sharing their enthusiasm in the field of Artificial Neural Networks and for supporting my initial proposal; Dr. Roger Lough (former Chief Guided Weapons Division) who approved my project; Dr. Nandagopal for being a valuable co-supervisor and for arranging Dr. Cheng-Chew Lim of The University of Adelaide to be my principle supervisor.

Dr. Lim and I have had numerous interesting and valuable discussions which varied from Artificial Neural Networks, to Psychology, Neuropsychology, Neurobiology, visual illusions, consciousness, hypnosis, meditation, etc. What has surprised both of was the fact that most of our discussions were on Psychology and Neurobiology (neither one of us has had any previous training in either of the fields, Dr. Lim is a control engineer, whereas my graduate training was in Physics). Both of us now appreciate that a serious attempt to study Adaptive Resonance Theory, in particular its relationship to cognitive information processing, does require the study of the other fields. I owe my sincere gratitude to Cheng-Chew who always had time to listen to me and allowed me to wonder into the unknown. Despite the mammoth task that lay ahead of me, I was armed with lots of experience in image processing and a firm conviction that Adaptive Resonance Theory would lead me in the right direction.

I am indebted to the Cooperative Research Centre for Sensor Signal and Information Processing (CSSIP) for providing conducive research atmosphere and in particular to Prof. Doug Gray and Dr. Nanda Nandagopal for arranging a room for me during my two year affiliation with CSSIP.

P.L.

List of Publications

The following is a list of publications which are relevant to the research reported in this thesis.

Lozo, P. (1996a). Dynamic tuning in the Feedforward Excitation-Feedback Presynaptic Facilitation (FFE-FBPF) neural circuit. *Proceedings of the Seventh Australian Conference on Neural Networks (ACNN'96)*, Canberra, ACT, April 10-12, pp. 79-84.

Lozo, P. (1996b). Neural circuit for match/mismatch and familiarity/novelty detection in SAART neural networks. (To appear in *Fourth International Symposium on Signal Processing and its Applications (ISSPA'96)*, Gold Coast, Australia, August 1996).

Lozo, P. (1996c). Neural circuit for self-regulated attentional learning in Selective Attention Adaptive Resonance Theory (SAART) neural networks. (To appear in *Fourth International Symposium on Signal Processing and its Applications (ISSPA'96)*, Gold Coast, Australia, August 1996).

Lozo, P. (1996d). Neural circuit for the recognition of distorted 2-D shapes in cluttered background by SAART neural networks. (*Submitted May 1996*).

Lozo, P. and Lim, C-C. (1996). Neural circuit for object recognition in complex and cluttered visual images. (*Submitted May 1996*).

Lozo, P. and Nandagopal, D.N. (1996). Selective transfer of spatial patterns by presynaptic facilitation in a shunting competitive neural layer. (*Submitted May 1996*).

Lozo, P. (1995). Selective Attention Adaptive Resonance Theory (SAART) neural network for neuro-engineering of robust ATR systems. *IEEE International Conference on Neural Networks,* Perth 27 Nov.- 1 Dec., pp. 2461-2466.

Lozo, P., Lim, C.C., and Nandagopal, D.N. (1995). Translation invariant pattern recognition: Neural network architecture based on biological visual spatial attention. *Australian Journal of Intelligent Information Processing Systems*, **2 (1)**, 1-10.

Lozo, P., Lim, C.C. and Nandagopal, D. (1993). Towards Dynamic Neuro-computational Vision Part 1: Covert and exogenous visual spatial attention neural network with translation invariance. *IEEE Visual Signal Processing and Communications* - Workshop Proceedings, Melbourne 20-22 Sept., pp. 145-149.

Lozo, P., Nandagopal, D. and Lim, C.C. (1993). A self-organising neural network architecture for real-time and stable learning and recognition of translation variant patterns. *Proceedings of the First Australian and New Zealand Conference on Intelligent Information Systems.* Perth, 1-3 December, 10-14.

Lozo, P., Lim, C. and Nandagopal, D. (1994). A neural network model of 2-D visual spatial attention and translation invariant pattern recognition. *Proceedings of the Fifth Australian Conference on Neural Networks*, Brisbane, pp. 84-89.

List of Patents

The following is a Provisional patent that is based on the research in this thesis:

Selective Attention Adaptive Resonance Theory (Australian Provisional Patent Application No. PN4315 (24/7/95)).

List of Principal Symbols

A	passive decay rate of postsynaptic cellular activation;
$\overline{A}$	passive decay rate of inhibitory interneurons;
B	upper saturation level for postsynaptic cellular activation;
$\overline{B}$	charging rate of inhibitory interneurons;
C	lower saturation level for postsynaptic cellular activation;
D	passive decay of postsynaptic potential
F_i	facilitatory signal;
G	postsynaptic gain;
$\overline{G}$	lateral feedback inhibition gain;
H_y	gain of synaptic facilitation;
$\overline{H}_y$	gain of synaptic inhibition;
I_i	inhibitory synaptic input;
J_i	excitatory synaptic input;
K_u	gain of the postsynaptic feedback induced depletion of the stored transmitter;
K_v	gain of the postsynaptic feedback induced postsynaptic potential;
K_y	gain of the postsynaptic feedback induced depletion of the mobilized transmitter;
L	gain of lateral presynaptic facilitation and lateral presynaptic transmitter release;
Y	threshold for transmitter release;
f(.), g(.)	linear functions above a threshold;
h(.)	faster than linear function above a threshold;

c_{ij}	coefficient defining the coupling neighbourhood;
r_{ijk}	postsynaptic receptor level;
u_i	stored transmitter;
$\hat{u}_i$	stored transmitter at steady state;
v_i, v_i^+	excitatory postsynaptic potential (EPSP)
$\hat{v}_i$	steady state excitatory postsynaptic potential
v_i^-	inhibitory postsynaptic potential (IPSP)
$\overline{v}_i$	activation level of inhibitory neurons;
$\hat{\overline{v}}_i$	steady state activation level of inhibitory neurons;
x_i	postsynaptic cellular activation level;
$\hat{x}_i$	steady state postsynaptic cellular activation level;
w_i	released transmitter;
y_i	mobilized transmitter;
$\hat{y}_i$	steady state mobilized transmitter;
z_i	transmitter production rate;
α_u	passive storage of the produced transmitter;
β_u	passive decay and mobilization of the stored transmitter;
β_y	passive transmitter mobilization
γ_y	passive decay of the mobilized transmitter
Θ	threshold level for postsynaptic activation;
$\overline{\Theta}$	threshold in SAART's top-down memory pathway

List of Figures

Chapter 1

Introduction

"We find ourselves in a bewildering world. We want to make sense of what we see around us and to ask: What is the nature of the universe? What is our place in it and where did it and we come from? Why is it the way it is?.

S. Hawking (1988)

1.1 Background and Motivation

A typical visual scene is cluttered with many objects that may be embedded in complex background and yet our visual brain is extremely efficient at recognizing the individual objects as well as the whole scene. What neural mechanisms enable our visual brain to quickly recognize familiar objects in complex and cluttered visual scenes where current artificial neural network models and conventional computer vision systems fail? How do stable visual memories get established and what effect do they have on the attentional processes and the success of the brain in complex and novel scenes? What is selective attention, how is it implemented and what does it contribute to? What characterises the actual moments of visual perception and awareness? How do these brain states arise and how are they sensed in a dynamic and potentially chaotic brain? These and related cognitive aspects of the human brain, currently barely understood, need to be solved if we are to unravel the fundamental architectural design principles of biological neural networks in general and the primate visual brain in particular so that we can design more intelligent and robust artificial neural systems for technological applications.

Most known cognitive properties of the human visual brain have been scientifically revealed through psychophysical experiments using brief visual displays and through the study of neuropsychological patients. These studies indicate that cognitive vision and visual memory depend on interacting pre-attentive and attentive processes, without which the visual system is virtually useless. Most vision researchers, cognitive theoreticians and neural network modellers have, however, largely ignored the importance of feedforward-feedback interactions between the bottom-up or pre-attentive visual processes and the top-down attentive and memory driven processes. Current literature therefore abounds with various artificial neural (and non-neural) models of pre-attentive vision (filtering, edge and contrast detection, figure ground separation, texture segmentation, etc.) and object feature extraction and recognition that are basically of the feedforward type. Top-down processes, although largely ignored in computer vision and neural network models, are very important, both in central and peripheral human vision, and often override the bottom-up pre-attentive processing through cognitive control. Typical examples of visual problems where top-down attentive processes may be required is when recognizing overlapping shapes, occluding figures, objects embedded in cluttered background of low contrast and uneven illumination. Top-down attentional process may also be driven and influenced by prior memory, such as when searching for a particular item in a cluttered input field.

The basic architectural organization and the functioning of low-level (pre-attentive) biological vision, having been extensively studied at the experimental level and more recently at the neuro-theoretic level (Grossberg and Mingolla, 1985a,b), are far better understood than any other part of the brain. Although the detailed knowledge of early visual processes is still to be worked out, it is now well established that visual information from the retina proceeds through several stages of retinotopically organized and massively parallel feature detecting cells in the lateral geniticulate nucleus (LGN), the primary visual cortex (area V1) and the extrastriate cortex (areas V2 and V4), (Hubel and Wiesel, 1977; Van Essen *et al.*, 1990; Zeki, 1993). However, very little is known about the organizational principles and the functioning of the higher levels of attentive and cognitive primate vision. For examples, it is not yet known how the subsequent cortical layers (inferior temporal (IT) cortex, parietal cortex, prefrontal cortex, etc.) process the information from early visual layers to give rise to coherent visual perception, memory storage and ability to recognise an object in a complex and cluttered visual world despite numerous transformations that the sensory stimulation from an object may project onto the retina. The neuronal mechanism that links vision, perception and memory is thus largely unknown, although it is established that the neural circuits within inferior temporal cortex are depositories for long term visual memories (Gross *et al.*, 1972; Gross 1973a,b; Desimone and Gross, 1979; Perret *et al.*, 1982; Richmond *et al.*, 1983; Miyashita, 1993).

It is the goal of this thesis to propose neurobiologically inspired and plausible neural design principles, mechanisms and circuits of high level primate vision, with a particular emphasis on selective visual attention, memory and object recognition in complex and cluttered visual scenes.

1.2 The Need for Selective Visual Attention

The most basic cognitive requirement of a real-time adaptive neural sensory information processing system (be it specialised for visual, auditory or another modality) is the ability to extract and recognise a familiar or a previously learned sensory stimulus when that stimulus is embedded in a novel, cluttered or noisy sensory environment that may often be severely degraded. A more general cognitive neural system should also be able to recognise the same stimulus when it is partly occluded by other visual stimuli or when its projection to the early sensory neurons has been transformed or warped, such as when a viewed object undergos complex 3-D motion (leading to relative changes in projected retinal size, position, orientation, etc). The visual brain of primates is an example of a biological cognitive visual system that can learn to recognise familiar objects and figures in complex visual scenes, even when portions of the object are obstructed or appear in unfamiliar contexts.

A cognitive neural system that interacts with its environment must also be able to adapt in real-time to the demands placed on it by a continuously changing environment. Real-time learning is not the only prerequisite for an animal's survival since the organism needs to be able to extract the stationary components of its environment even when the contexts may change, otherwise frustration and death may follow. A typical example of such a problem arises if the animal is unable to recognise its food whenever it appears in different contexts, or is unable to recognise its predator in unfamiliar contexts. Thus, organisms need to be able to extract the relevant and relatively stable portions of the sensory input and be able to learn stable internal memory representation of the extracted information. Otherwise, previously learned information will be useless if it can be continuously eroded by recent experiences. At the same time, sensory stimulation that has immediate physiological or psychological consequence (ie. are important to immediate survival) needs to be processed and evaluated, even if it is at the expense of other processing, so that an appropriate response can be made (such as pain avoidance, escape from danger or satisfaction of internal desires such as hunger relief). In order to process the relevant stimulus in a continuum of simultaneous sensory stimuli (and to ascertain its meaning or consequence), the neural systems controlling the animal's behaviour and response need to be able to selectively attend to (or tune into) the important (or desired) stimulus while at the same

time temporarily ignore the rest of the stimulation. If the selected stimulus has no direct consequence to the animal, then the neural system has to be able to shift attention to other sensory events within.

Although the early layers of the primate visual system are massively parallel, the shear volume of real-world visual information that may simultaneously barrage the primate visual system suggests that there ought to be a neural mechanism that can filter most of the information, thus allowing organisms to respond to the relevant portion of the overall stimulation. On the basis of psychophysical (Sperling, 1960; Eriksen, 1988; Eriksen and Schultz, 1979; Eriksen and Yeh, 1985; Eriksen and Murphy, 1987; Posner, 1980; Shulman *et al.,* 1979; Treisman and Gelade, 1980; Tsal, 1983), neuropsychological (Posner *et al.,* 1984, 1987; Rafal and Posner, 1987) and neurophysiological experiments (Lynch *et al.,* 1977; Mountcastle 1978; Moran and Desimone, 1985; Mountcastle *et al.,* 1987; Petersen *et al.,* 1987), consensus has been reached that selective visual attention enables only the relevant portion of the sensory stimulation to reach the higher cortical layers. Neurophysiological data has shown that various neural centres may be involved (either directly or indirectly) in selective visual attention. The main cortical areas that have been found to be involved in selective visual attention, or are affected by it, include the parietal cortex (Mountcastle, 1978; Desimone *et al.,* 1990; Robertson *et al.,* 1988; Posner *et al.,* 1987), the medial pulvinar (Robinson, 1993; Chalupa, 1991; Bender and Butter, 1987; Petersen *et al.,* 1987), visual area V4 (Moran and Desimone, 1985; Haenny *et al.,* 1988; Haenny and Schiller, 1988; Motter, 1993; Braun, 1994), the inferior temporal cortex (Moran and Desimone, 1985; Sato, 1988) and the anterior cingulate gyrus (Posner and Rothbart, 1992, 1994). It is highly likely that other areas, if not directly involved are in some way affected by attentional factors, such as the hippocampus and other memory pathways. Recent experimental data indicates that even the earliest visual cortical areas (V1 and V2) are influenced by attentional factors (Motter, 1993), with the presence or absence of competing stimuli being the determining factor for the differential selectivity of the cell.

Extensive psychophysical experiments conducted over the past three decades (for a good summary refer to Van der Heijden, 1992) provide evidence that selective visual attention has spatial and non-spatial components which have automated rudimentary selectivity to the incoming visual information so as to orient the organism (human) towards salient and novel features of the input. The automatic component of visual attention facilitates what is often referred to by vision researchers as the "pre-attentive vision", while the voluntary component contributes to the so called "attentive vision". These two highly coupled components of human vision enable us to selectively attend to (or tune into) a portion of input for further processing, the purpose of which may be to verify or nullify a hypothesis that we generate during our exploratory behaviour and interaction with environment.

If we did not posses the ability to ignore the irrelevant inputs, we would find it very difficult to comprehend anything in a complex visual scene. The fact that we do indeed posses an attentional filtering system is exemplified by our limited ability to simultaneously recognise more than one object. For example, we cannot simultaneously recognize with high degree of certainty more than one word in this sentence nor can we simultaneously ascertain whether two successive words in the sentence are spelled correctly. We typically "pay attention" to one object at a time and then move attention to another. How much processing time we allocate to a particular object depends on our current goals, needs, relevance and the familiarity of the attended object. However, we would be at a disadvantage if we were always limited to perceiving a scene in terms of single isolated objects (which may be true for a scene containing novel objects). Once familiar with the contents of a given scene, that is, when we have constructed an internal representation of the scene, we should be able to attend to the whole scene rather than to its constituent elements. In such cases, our attentional system should not only alert us about a potential discrepancy between our stored representations of a given scene and its real contents (which may be caused by the dynamic changes in the scene), but it should also orient us to the approximate location of the mismatch to allow us to attend to and learn the novelty within the scene.

It is also recognised that selective attention may also be subject to voluntary and adaptive influence from higher visual centres. Experimental psychological data shows that there is a relationship between selective visual attention and visual memory (Sperling, 1960; Rock and Gutman, 1981). For example, the experimental data on the recognition memory of a briefly presented set of overlapping figures shows that the attended stimuli tend to be remembered reasonably well, whereas memory of unattended stimuli is at chance level (Rock and Gutman, 1981). In addition to selective transmission of visual information, it has been proposed that visual attention provides the key to forming invariant object representations (Palmer, 1983) and, may also be involved in the encoding of depth information (Epstein and Lovitts, 1985).

Evolutionary processes have thus endowed visually intelligent life-forms with a neural mechanism to selectively pass the behaviourally relevant visual information through to brain centres where learning and recognition takes place. Although visual attention is not the only form of attention in biological systems, it is by far the most studied at the psychological and the cellular level. Whether such a mechanism is needed in technological applications of artificial neural networks remains to be seen.

Selective attention and memory guided processing thus appear to be some of the key mechanisms of biological neural systems that provide them with cognitive abilities that have not yet been matched by traditional artificial intelligence approaches nor by current artificial neural network models of learning and pattern recognition. It

is therefore both of theoretical and practical interest to understand the neural mechanism of selective attention and how it may contribute to intelligent behaviour in a self-organising artificial neural network model of sensory information processing and visual perception. Of particular interest are the interactions between attention and memory, that is how attention affects memory and how the established memories affect the attentional processes.

Very little is known about the neurobiological interactions and circuits that mediate selective attention, although it has been recently suggested that the massive feedback pathways in the brain may play a significant role (Mumford, 1994; Posner and Rothbart, 1994; Churchland *et al.,* 1994). Feedback pathways in the brain are perhaps one of the more important but the least understood and appreciated design principles of the brain. With the advent of real-time competitive and self-organising neural network models of learning during the past two decades, particularly the invention of the Adaptive Resonance Theory (ART) by S. Grossberg (1976, 1980) and its subsequent mathematical implementation in various neural network architectures (Carpenter and Grossberg, 1987a, b, 1990), coupled with the advances and knowledge being gained from psychology and experimental neuroscience and neurophysiology, it should now be possible to make significant inroads into the basic understanding of fundamental neural design principles of biological neural networks.

1.3 Artificial Neural Networks and Computer Vision

Since the re-emergence of Artificial Neural Networks (ANNs) in the mid 1980's, through the influential paper by Hopfield (1982), ANNs have become very popular, particularly in pattern recognition applications. One of the main reasons for their re-emergence and popularity since the early work by McCulloch and Pitts (1943) are the new architectures and learning laws that enable ANNs to be adopted to various problems domains. Because ANNs learn to solve a problem by being trained on examples that are chosen from the problem domain, they re claimed to be able to generalise to new, previously "unseen" examples within the same problem domain. It is also generally regarded that the application of ANNs reduces the time needed to generate a solution to a particular problem since the burden of providing a formal mathematical or engineering solution (which cannot be always easily done) is replaced by training the chosen network on the relevant data.

Even though ANNs have proven to be extremely good as pattern recognizers and in some cases have exceeded the performance of conventional classifiers, yet after some 10 years of their applications to various problems, particularly to visual object recognition, we still do not have robust object segmentation and recognition systems. This appears to be due to our failure to appreciate the intricacy and the complexity of

mechanisms that enable biological vision systems to succeed so effortlessly on visual problems to which ANNs are applied only to fail miserably. Perhaps we fail to appreciate the fact that real-world visual sensory, and perhaps other sensory domains, are too complex to be considered as the problems of simple pattern recognition.

The image of a visual scene is a two-dimensional mosaic of luminances that are reflected from the objects and the background within the scene. In order to recognize a particular object in a cluttered background, the object recognition system first needs to separate the various luminances from the other objects and the background so that it can segregate the luminances that belong to one object. That is, the system must be able to determine which of the various luminances in the retinal image belong to the same object. Segregating the various luminances that belong to the same object is in general a very difficult problem in computer vision. Problems arise when the scene contains several occluding objects or when the luminance from an object is non-uniform and merges with that of other objects or the background. While such scenes do not pose much of a problem to biological visual systems, they are particularly difficult to solve by machine vision systems which would fail to properly segment an object, thus leading to recognition failure (recognition of an improperly segmented object would most certainly fail).

Computer vision and object recognition researchers generally apply various pre-processing and feature extraction algorithms to visual images in order to reduce the image data to a computationally manageable level while generating a simplified description of its contents. Typical pre-processing algorithms for image filtering and object segmentation include median filtering, edge detection, thresholding and boundary following (Pratt, 1978; Ballard and Brown, 1982; Rosenfeld and Kak, 1982; Marr, 1982; Peli and Malah, 1982), etc. These low level pre-processing steps are then followed by a feature extraction process that transforms the shape of the segmented object into a small number of shape descriptors that are representative and unique to the object. These object descriptors are then used as the input to the classification (recognition) module that may or may not be based on a neural network.

For example, in a typical application of ANNs to visual or infrared (IR) object recognition (Lozo *et al.,* 1991), the problems of object segmentation (separation of the object from its background), feature extraction and classification are done by independent modules, all working in a feedforward manner. Neural networks, although not restricted to, are generally applied in the final stages of processing (i.e., feature learning and classification). This traditional approach to object segmentation and recognition, schematized in Figure 1.1, where each processing stage depends on the success of the

previous stage, works quite well on well contrasted objects, such as those shown in Figure 1.2, where feedforward computations can reliably separate the target object from its background.

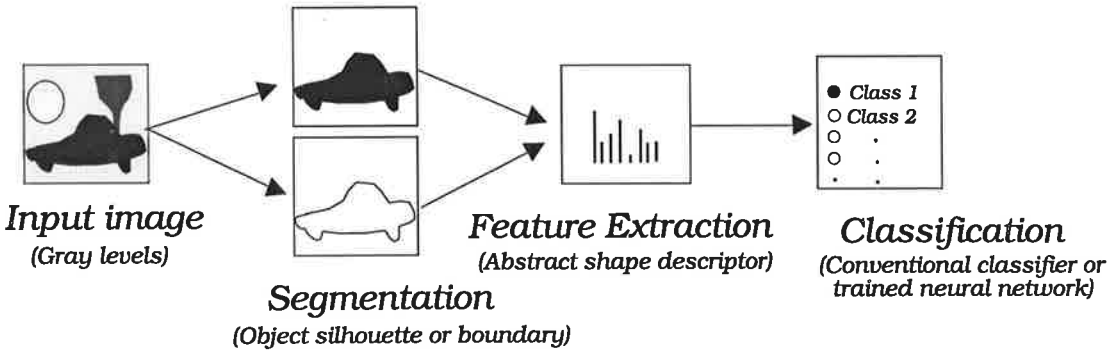
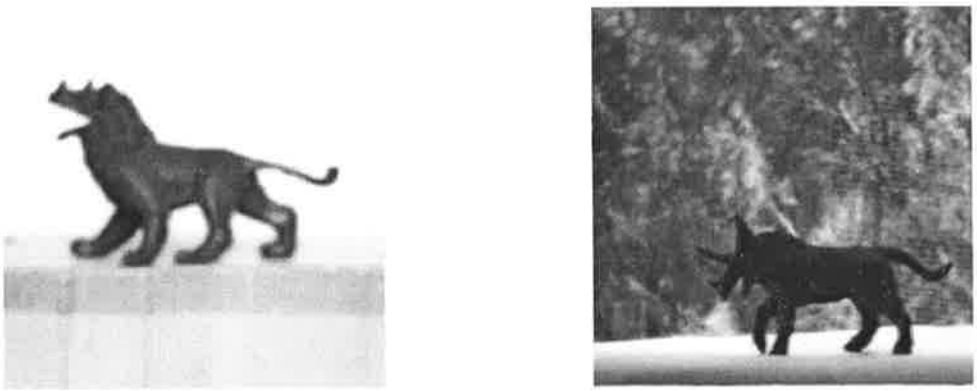


FIGURE 1.1.Traditional feedforward processing steps in object segmentation and recognition.



(a) Scene 1 (b) Scene 2

FIGURE 1.2. Examples of well contrasted objects in visual images where current feedforward object segmentation and recognition systems are likely to succeed.

However, these approaches generally fail on scenarios that typify the real visual world of noise, clutter, complexity, occlusions and overlapping figures, etc. When the segmentation stage fails to cleanly separate an object from its background, the whole system is doomed to fail even though the higher levels of the network architecture may contain the memory (network weights) of the desired object. Why bother training an Artificial Neural Network if its acquired memory weights do not help it one iota in circumstances when it is really needed, i.e., when the preceding stages fail? In fact, if the pre-processing stages do succeed, there is really no need for a neural network at the recognition stage since there are already many well established traditional

approaches that work quite well. It seems intuitive to suggest that perhaps the memory of a trained neural network should therefore play a greater role in the overall process than is currently appreciated.

(a) Scene 3 (b) Scene 4 (c) Scene 5

FIGURE 1.3. Examples of complex and cluttered visual images where current feedforward object segmentation and recognition systems are likely to fail.

To expose some reasons why current artificial neural networks and more conventional image processing systems are very likely to fail on the above images, below we show the result of processing the above images with a typical edge detector. We have used a simple 3x3 Sobel operator, although any other edge detector could have been applied for the current purpose.

(a) Scene 3 edges (b) Scene 4 edges (b) Scene 5 edges

FIGURE 1.4. Sobel edge processed images of complex visual scenes that demonstrate some of the difficulties faced by feedforward segmentation and recognition systems.

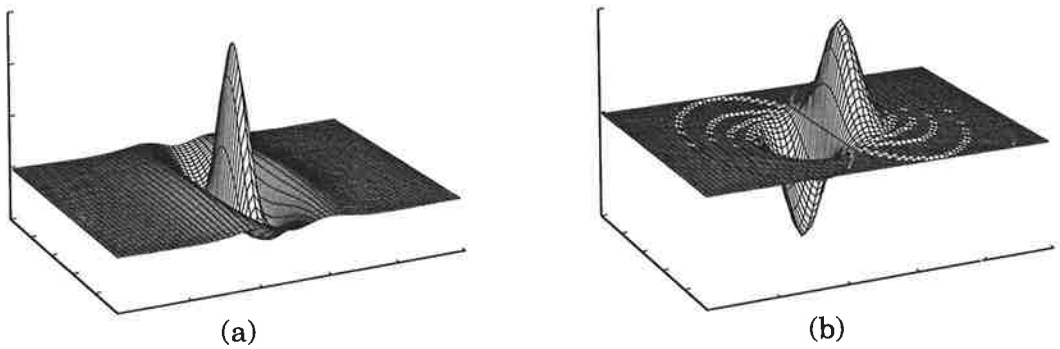
Let us now examine these edge maps in the context of some of the most commonly used algorithms that use image edges to segment individual objects. The reason for considering the edges (rather than the grey level images) is primarily because most

computer vision and object recognition systems typically include edge detection at the very earliest stages of processing (as is also experimentally found to be the case in the early stages of biological visual systems).

As can be seen in the three image edges of Figure 1.4, the boundary of the target object is not well defined and in some places has a much weaker edge strength than the non-relevant object and the background. If the edge maps are to be thresholded, as is typically done, then parts of the target boundary would be discarded while significant portion of non-relevant edges in close proximity to the target object would be retained. If on the other hand we were to use a boundary following technique that not only uses edge magnitudes but includes edge directions as well, the boundary follower would still fail to trace the proper boundary of the target object. For example, the target object in Figure 1.4(a) has strong edge magnitudes along most of its boundary, with the exception of a small portion below the neck of the toy animal. A boundary follower will thus be driven off the target object. Hence any recognition system whose segmentation is based on edges (and edge directions) is unlikely to succeed in segmenting the target object in above images. The problem is that it is not always possible to decide at the local level which of all the possible edges form part of the single object.

2-D Gabor functions in image segmentation and feature extraction

Gabor functions, which were initially introduced as one dimensional functions in the context of uncertainty theory of information (Gabor, 1946) have been widely used in image segmentation and feature extraction since their extension to two dimensions by Daugman (1985). Figure 1.5 shows an example of two elongated and spatially oriented 2-D Gabor functions.



(a) (b)

FIGURE 1.5. Example of oriented symmetric and asymmetric 2-D Gabor functions: (a) symmetric; (b) asymmetric filter.

The reason for their recent popularity is two fold: (i) Gabor functions provide an image filtering scheme whereby an arbitrary image area can be represented by the response of symmetrical and asymmetrical 2-D filters that are simultaneously maximally localized both in space and in spatial frequency; and (ii) 2-D Gabor filters resemble the receptive field profiles of the simple cells in the visual cortex. The early experimental neurophysiological data from single cell neuronal recordings (Hubel and Wiesel, 1962) has provided evidence that simple cells in the primary visual cortex (area V1) consist of elongated excitatory and inhibitory zones, thus producing a receptive field that responds maximally to oriented lines or edges.

Gabor functions, or convolution kernels, are typically represented by a product of a Gaussian and a sinusoid. The two most commonly used Gabor filters can be written mathematically as

$$G_S(x,y) = g(x,y)\cos(\Phi) \tag{1.1}$$

$$G_A(x,y) = g(x,y)\sin(\Phi) \tag{1.2}$$

where $g(x,y) = \exp\left(-\frac{1}{2}\left(\left[\frac{x}{\sigma_x}\right]^2 + \left[\frac{y}{\sigma_y}\right]^2\right)\right)$ is a 2-D Gaussian with σ_x and σ_y defining its spatial extent (and elongicity), while $\cos(\Phi)$ and $\sin(\Phi)$ are the two sinusoidal functions that determine the filter's spatial frequency and whether the filter is symmetric or asymmetric. The asymmetric and elongated filter, $G_A(x,y)$, is typically applied to the detection of edges and contrast differences in various directions and at various spatial scales. As filters, 2-D Gabor functions are considered to represent the lowest level of processing in the visual systems and have been applied to various image processing problems.

Applications of Gabor filters include image compression (Daugman, 1988), texture segmentation and discrimination (Du Buf, 1990; Fogel and Sagi, 1989; Clark *et al.,* 1987; Turner, 1986), computation of surface orientation from textured images (Gopal *et al.,* 1990) and combined image sampling and object feature extraction (Flaton and Toborg, 1989). In typical applications, a larger number of such kernels, each differing in spatial resolution and orientation is applied to an image. As the spatial extent of the filter is increased, its resolution is decreased. The features thus extracted are then used in the recognition process. For example, Flaton and Toborg (1989) have devised a scheme whereby the feature vectors obtained from Gabor convolutions with an image at different orientations and spatial scales are conveniently arranged into suitable data structures to enable efficient object recognition in infrared images by a simple similarity measure.

While the application of 2-D Gabor function to image segmentation and object feature extraction provides an alternate approach to many existing image segmentation and feature extraction techniques, there is no evidence that it provides a more robust solution to the existing problems when the image clutter and complexity increases.

Thus while there are many alternate approaches that have been proposed in the field of computer vision and object recognition over the past three decades (not all being inspired by neurobiology), it may be concluded that when a particular solution works well on a restricted set of images, one cannot guarantee its success on images that depart even slightly from the ones that it was aimed at. The problem in general is too difficult to solve. Current approaches to computer vision and object recognition, whether they are purely based on conventional image processing algorithms, state-of-the-art ANNs or a hybrid, are susceptible to failure as soon as the complexity and the clutter in the image increases, thus placing a heavy burden on the object segmentation processes. The overall system is likely to fail simply because the segmentation process may fail to properly segment an object. In the process of segmenting an object from its background and obtaining its simplifying descriptors or patterns, often the important information may also be discarded. This may lead to object descriptors or patterns that bear no resemblance to the descriptors of the complete object (even if only a small portion of the object information is discarded). Since the transformed simplifying abstract feature patterns obtained from an improperly segmented object are unlikely to bear any relation to the original pattern from a cleanly segmented object, the object recognition systems will likewise fail.

Image segmentation *'the Achilles heel'* of object recognition

Many researchers have realised that object segmentation in a realistic visual image is perhaps one of the most difficult problems in computer vision. For example, in their investigation of a neural network architecture for 3-D object representation and recognition by a mobile robot, Seibert and Waxman (1992), have noted that "*segmenting a 3-D object cleanly from a complex scene is a very difficult problem in general because of interference from noise, occluding objects, background, illumination and spatial sampling effects*". Although these authors have used a hybrid of image processing and neural networks in their work on 3-D object recognition (see section 1.3.1), it does not necessarily mean that a trained feedforward neural network is required in the segmentation stages nor does it guarantee the success of such a network if it were to be used.

Attempts to circumvent the failure of the segmentation process at higher levels is prone to errors not only because the important information may have been discarded by the lower level processing modules, but the higher levels may make a wrong classification or recognition response on the partial data that is available. This is unlike

the case of human vision where for example we are able to recognize an occluded object after seeing it first in its entirety. Even if an artificial neural network is able to recall the complete memory from a partial description, memory recall does not necessarily imply recognition unless the recalled memory is used effectively in the comparison process. But the recall of the complete memory from a partial input will not match the input. We are thus left with a paradox where on one hand we need a complete memory recall from a partial input while on the other hand we need the input to match the recalled memory. An interesting and relevant idea that was recently suggested is that "*it is much more significant for a part of the input pattern to match closely the stored memory, than for all of it to match slightly*", (Mumford, 1994).

What appears to be needed is a neural architecture where the bidirectional interactions between pre-processing and recognition stages enable the established memory of an object to influence and assist the pre-processing stages when required. The basic notion of bidirectional interactions or 'interactive vision' (Churchland *et al.*, 1994), is that higher levels of vision should provide top-down feedback to lower visual layers, such that 'partial segmentation helps partial recognition and partial recognition helps segmentation'. Bidirectional interactions should be easy to implement in 2-D neural circuits compared to conventional computer algorithms. However, we currently do not have feasible neural networks with such interactive properties, nor do we know what type of feedback should be used. Engineers are quite accustomed to electronic circuits where feedback is heavily used, particularly in the design of feedback control systems. Unfortunately, engineered feedback controls systems are typically characterized by positive or negative feedback. Positive feedback is not very popular amongst engineers (they loathe the positive feedback) because of potential instabilities that it may cause. Positive feedback is used in both the Boundary Contour System (BCS) of Grossberg and Mingolla (1985a, 1985b) and in the ART based neural networks (Carpenter and Grossberg, 1987a, 1987b, 1990) to provide new computational capabilities and yet these neural networks are stable! However, there is at least one type of electronic circuit, Automatic Gain Control (AGC), where a feedback signal is used in a completely different way, i.e., it is not strictly positive nor is it strictly negative, but modulatory (positive modulatory or negative modulatory). The feedback signal in the electronic AGC circuit biases the gate of a Field Effect Transistor (FET) such that the amplitude of the AC coupled input signal is modulated to produce a constant amplitude output signal (Millman, 1979, pg. 406).

Although the Boundary Contour System (BCS) does have feedforward-feedback interactions (of the excitatory type) and can explain and reproduce numerous visual illusions, it can still fail to properly segment objects in visual images such as those shown in Figure 1.3, where the human visual systems succeeds readily. Even when BCS is combined with an ART based neural network (such as the ART-3 neural network

of Carpenter and Grossberg, 1990), the whole system will still fail to segment and recognise a previously learned object when that object is subsequently embedded in cluttered background.

Thus, even with the advent of Artificial Neural Networks, we have not yet solved the problem of machine vision and object recognition, although we seem to be getting closer to understanding the reasons why. In an article on competitive learning Grossberg (1987) makes a following observation of the neural network field, particularly the inadequacy of the current neural network models to deal with complex sensory environments:

> *"The architecture of many popular learning and information processing models are often inadequate because they have not been constrained by the use of design principles whereby they could stably self-organize. Many models are actually incompatible with such constraints and some models utilize physically unrealizable formal mechanisms. Learning models which cannot adaptively cope with unpredictable changes in a complex environment have an unpromising future as models of mind and brain, and provide little hope of solving the outstanding cognitive problems which are not already well-handled by traditional methods of artificial intelligence and engineering."*

So what new neural mechanisms and interactions are required? It we refer back to the above images, it is interesting to note that when humans first view the object shown in Figure 1.2(a) and then the view the three images shown in Figure 1.3, their short term memory of the seen object enables them to quickly segment and recognize the same object in the cluttered backgrounds of Figure 1.3. This simple example seems to support the view that stored visual memories are able to influence (prime) the bottom-up neural signal processing, enabling only the relevant portions of the input to be matched with memory at the higher levels. But for the memory to have an effect on lower visual processing, there must be a top-down feedback. However this feedback cannot be solely excitatory. The reason for this is that an excitatory feedback (i.e., additive feedback) from memory to early visual layers would add information (memory signals) to these layers, which translated to human vision would imply that we should see an object whenever we remember or imagine it. What other feedback enables the stored visual memories to influence how the early visual layers subsequently processes the information in cluttered scenes?

In Chapter 4 we will propose two new types of feedback mechanisms that are neither positive nor negative but modulatory. The brain evolution may have discovered

modulatory feedback as means of linking various neural layers into bidirectional interactions, thus enabling memory to influence lower visual layers and attentional mechanisms.

1.3.1 Related neural architectures for 2D and/or 3D object recognition

In this section we review several related neural network architectures for 2D shape and/or 3D object recognition. At the outset, it should be stressed that there is a considerable difference between the recognition of 2D shapes (which is the focus of this thesis) and the recognition and representation of 3D objects, even though a 2D shape may represent the boundary of a real 3D object when viewed from one perspective. The latter generally requires that several different views (from different perspectives) of an object be linked into a unified representation.

Adaptive 3D object recognition from different aspect views

Beginning in 1989, M. Seibert and A. Waxman (1989, 1991, 1992) have pioneered the development of a novel hierarchical real-time artificial neural architecture that self-organizes internal representations of 3D objects from 2D view sequences. The approach, which is based on the concepts of aspect graphs (Koenderink and Doorn, 1979), uses transitions from one aspect view to another to link nearby aspects into a sequence that would be normally obtained when the viewed object is being observed as one moves around the object. One of the attractions of this approach is that the recognition confidence increases as one gets more views of the same object. To achieve positional invariance, Seibert and Waxman have developed a neural network called NADEL (Seibert and Waxman, 1989) which computes the centroid of an object via a diffusion process with feedback to align the camera towards the object. After extracting the shape features (typically corners) they then find the centroid of these features and apply the log-polar transform with respect to this feature centroid. Scale and rotation invariant recognition of each 2D shape is achieved by a linear shift to the new centroid in the transformed domain. Tolerance to small deformations (such as forthshorthening) is achieved by sampling the log-polar features via 2D kernels. Each 2D shape (aspect view) is then learned on-line by an ART-2 neural network. Successive views are then linked by the aspect graph network which learns transitions from one view to the next (Seibert and Waxman, 1991; 1992). In order to enable the aspect network to learn transitions, Seibert and Waxman have proposed the concept of neuromodulation, where one aspect facilitates the learning of the next by gating the appropriate pathways.

Whenever an aspect transition occurs from aspect x_i to aspect x_j, the activity of the previous aspect decays while the activity of the new aspects builds. During the transient time interval when both aspect nodes are co-active, the relevant weights are modified according to the following equation:

$$\frac{dw_{ij}^k}{dt} = k_w w_{ij}^k (1 - w_{ij}^k) \{\Phi_w[(x_i + \varepsilon)(x_j + \varepsilon)] - \lambda_w\} \Theta_y(y) \Theta_z(z_k) \tag{1.3}$$

where k_w governs the rate of the evolution of the weights relative to the x-node dynamics, λ_w is the decay rate of the weights. Object nodes y_k accumulate evidence for object hypothesis over time, as specified by the following equation:

$$\frac{dy_k}{dt} = k_y \left\{ \left[\sum_i \sum_{j>i} \Phi_y[(x_i + \varepsilon) w_{ij}^k (x_j + \varepsilon)] \right] - \lambda_y y_k \right\} \tag{1.4}$$

The Hebbian-like term $(x_i + \varepsilon)(x_j + \varepsilon)$, where ε is a small positive constant, combined with the the gating term w_{ij}^k forms an axo-axo-dendritic synapse (Seibert and Waxman, 1990, 1991). The complete neural network architecture is embedded in a mobile robot and has been demonstrated in real-time, albeit on simple visual scenarios where the object of interest is embedded in a reasonably clean background. In addition to being able to recognize 2D shapes of 3D objects and linking these into a 3D representation, they have also demonstrated that the neural architecture can learn to recognize simple 3D visual landmarks in real-time, the results of which have been compared with the recordings from rat hippocampal cells (Bachelder and Waxman, 1994; Waxman, Seibert and Bachelder, 1995). While the above work has introduced several novel concepts (such as neuromodulation, as represented by above equations) and some new network architectures (such as aspect network) that were strongly motivated by living biological visual systems (macaque monkey in particular), its main weakness is the inability to deal with more realistic visual scenarios. The use of the log-polar transform, while being based on known experimental findings from the early visual pathways (Schwartz, 1980) provides an interesting and a convenient solution to achieving size and rotation invariant recognition. However, the log-polar transform also limits the capability of the system. Not only is it necessary that the sensor (in their case a camera mounted on a mobile robot) be always pointed at the computed centroid of a 2D shape before the shape can be recognized. Not only that, the log-polar transform does not necessarily provide a biologically plausible method for achieving perceptual constancy since it cannot differentiate between two shapes that may represent quite a distinct entity in the real visual world, such as when one 2D shape may be transformed into the other by a simple geometrical transformation (see Appendix D for a critique on the log-polar transform).

VIEWNET

The work by Seibert & Waxman, and that of Bradski et al., (1992) has inspired another neural network model for 3D object recognition from accumulated views of 2D aspects. The network called VIEWNET (Bradsky and Grossberg, 1995) also learns to code 2D views but stores these categories in a working memory called STORE (Bradsky et al., 1992). The VIEWNET architecture basically consists of the following three parts: (i) an image preprocessor called CORTEX-2 (Carpenter et al., 1989; Grossberg and Wyse, 1991, 1992) which is a simplified version of the BCS network; (ii) a self-organizing pattern recognition network based on the Fuzzy ARTMAP (Carpenter et al., 1992); and (iii) a working memory network, STORE, to accumulate evidence over multiple views. As in the neural system of Seibert & Waxman, the VIEWNET architecture also uses the log-polar transform to achieve size and orientation invariant recognition of 2D aspects of 3D objects. However, both of the approaches suffer from the fact that neither can successfully recognize a 2D aspect of a 3D object when that object is embedded in a realistic sensory environment of clutter and noise, primarily because the image preprocessing components cannot handle these types of images nor do they benefit from a prior memory that the system may have previously learnt.

Neocognitron

Neocognitron (Fukushima, 1980, 1987, 1988; Fukushima and Miyake, 1984) is a multilayered neural network model for visual pattern recognition. The network embeds the mechanism of selective attention (i.e., it models feedforward-feedback pathways), is tolerant to small deformations of the input pattern and is also capable of translation invariant 2-D pattern recognition. Since the degree of translational invariance is controlled by the number of layers, the network is found to suffer in its recognition accuracy (Menon and Heinemann, 1988; Barnard and Casasent, 1990). Even though Neocognitron is capable of selective attention and shape recognition in cluttered backgrounds, the network as originally proposed was limited to binary inputs. A subsequent extension of the network (Ting and Chuang, 1993) has shown that it is possible to extend the network to analog images provided that two compensation parameters (m_0 and m_1, being representative "averages" of the background and the target shapes respectively) are first derived from the statistical profile of the image. One of the interesting features of this modification is that the network can be trained on binary patterns and used for recognizing analog input images.

1.4 Neuro-Engineering with Non-linear Neural Networks

The primate visual brain is too complex to be understood in its entirety and too large to be modelled on present day computers. It is therefore important to take

into consideration the knowledge that has been gained from the experimental study of simple biological neural circuits (such as those of invertebrates) and to integrate it with the data from primate neurophysiology while limiting ourselves to the consideration of those problems that could expose the key neurobiological design principles and mechanisms of primate vision, memory and attention.

On the assumption that the visual brain can be considered as a neural circuit (albeit complex) that is composed of dynamic elements, the problem then becomes that of *Neuro-Engineering*.

Definition: *Neuro-Engineering is a discipline that applies the engineering design methodologies to the design of parallel and adaptive artificial neural processing circuits and systems while incorporating the domain specific knowledge from cognitive science, psychology and psychophysics, neuroscience and neurophysiology, artificial neural networks, electronic engineering and mathematics.*

In addition to providing a better understanding of biological neural design principles and cognition, Neuro-Engineering also has a major role to play in modern computer technology and machine intelligence by providing us with new and robust design principles.

Neuro-engineering a dynamic artificial neural circuit that is not only capable of solving the outstanding technological problems but that is also capable of emulating and explaining some capabilities of biological vision requires that careful attention be paid to psychophysical and neurophysiological data, as well as to good engineering design principles and design logic. To be biologically plausible, as well as being of use in solving difficult technological problems, particularly in the fields of machine vision and automatic target recognition, artificial neural circuits should address the problem of vision and object recognition in complex visual scenes. The problem of how a self-organising neurobiological circuit of primate visual system can succeed to learn and perceive objects in a complex visual sensory environment remains unsolved. If the problem can be solved, at least theoretically, by a neuro-engineered artificial neural circuit that is constrained by the key psychophysical and neurophysiological data, then the neuro-engineered system forms a biologically plausible model of the brain function that it addresses.

Unfortunately, due to the lack of a rich variety of fundamental design blocks and guiding design principles and design logic that characterizes other engineering fields, such as that of digital systems design, the field of artificial neural networks has not yet evolved into a neuro-engineering discipline. It is thus appropriate to first develop a suite of robust and useful neuro-engineering design principles, mechanisms and

design logic that should enable us to design complex artificial neural circuits. While the designed neural circuits should be capable of reproducing and explaining the key neurobiological data and mechanisms of primate vision, their dynamics and behaviour for a given input should also be amenable to introspection by following their design logic. Just like an experienced digital systems design engineer can design complex digital circuits and analyse their behaviour by following the design logic, similarly, once we have mastered the fundamental neuro-engineering design principles through extensive applications and computer simulations, we should eventually reach a stage where we can design complex artificial neural systems and predict their behaviour without having to constantly carry out extensive computer simulation or mathematical analysis (which may not always be possible).

Before we can even contemplate the design principles of visual neurobiology and the brain functions in general, we need to be first aware of the problems faced by biological visual systems in the real world.

(a) What mechanisms are needed?

(b) How are they implemented?

(c) Does the massive psychological and neurophysiological data reveal any logic in the brain design?

(d) Are there any useful design principles and mechanisms embedded in the currently available Artificial Neural Networks (ANNs)?

(e) How does a designer of an artificial neural system know whether he or she is on the right track and whether their model is correct and, perhaps biologically plausible?

(f) Is there a universal solution to vision or are there many possible solutions?

(g) Are there some common neural mechanisms in the brains of rats, cats, monkeys and humans?

(h) What should a theorist and a neuro-engineer of brain functions hope to achieve?

(i) To what detail should one model visual neurobiology?

(j) Should the emphasis be on mathematics or design principles and mechanisms?

(k) What is the appropriate computational unit, a single cell or an array of competing cells?

(l) What is the most relevant mathematics to model brain functions?

In a text on neurophilosophy, Churchland (1986) offers the following glimpse into what a theorist of brain functions should hope for and how one may decide whether a theory is on the right track:

> *"The hope of any theorist is that if the basic principles governing how nervous systems operate are discovered, then other operations can be understood as evolution's articulation and refinement of these basic principles. Simplifications, idealizations, and approximations, therefore, are unavoidable as part of the first stage of getting a theory off the ground, and the trick is to find the simplification that is the Rosetta stone, so to speak, for the rest... What gives a theory the right "feel" in its early stages is whether or not it is capable of making sense of the phenomena, whether it shows itself capable of fitting in with established theory elsewhere in science (biology, genetics, physics, chemistry), whether it can unify and explain, and perhaps whether it is simple and elegant. There are, alas, no formal procedures for telling whether a theory is on the right track, let alone whether it is elegant, and perhaps the best indicator here is whether a theory captures the imagination of other scientists who, in making it their own, get it to reveal satisfying answers to important questions and solutions to long standing mysteries".*

Since the brain is dynamic, can learn in real time without formal external supervision (one exposure to a sensory stimulus can form a lasting memory of that stimulus) and has to process information in complex sensory environments, it seems obvious that plausible neural network models of brain functions must be capable of self-organised learning in complex inputs and be modelled by dynamic mechanisms. Most of the current ANNs are not capable of self-organised learning, while those whose design principles address self-organisation cannot learn nor recognize familiar visual stimuli in complex and cluttered inputs. It thus seems that one of the primary goals of an initial neural theory of biological vision should be to provide general neural design principles and mechanism for vision and object recognition in cluttered scenes.

We are now left to ask whether there exists a mathematical tool and a simplifying model of brain functions that could be considered as "the Rosetta stone, so to speak, for the rest". Amongst the numerous models of learning that are embedded in various artificial neural network architectures, the only ones that are based on a scientific theory of self-organised learning are the real-time neural network models of Carpenter and Grossberg (1987a, b, 1990). These neural networks are based on the Adaptive Resonance Theory (ART) of S. Grossberg (1976, 1980). ART is the first psychophysiological theory of self-organised learning that embeds the cognitive concepts of attention, vigilance, top-down priming and bidirectional learning in real-time neural systems.

Real-time neural systems are characterized by dynamic mechanisms (i.e., they are modelled by differential equations) and are able to learn and store memories without a need for external supervision. Biological learning systems, such as the human visual brain, can store the memories of new objects quickly and without the formal external supervision. ART suggests that this type of unsupervised learning may be achieved by a system that stores memories in the bottom-up and the top-down pathways, as shown in Figure 1.6, thus enabling the recalled top-down memory to influence the learning process.

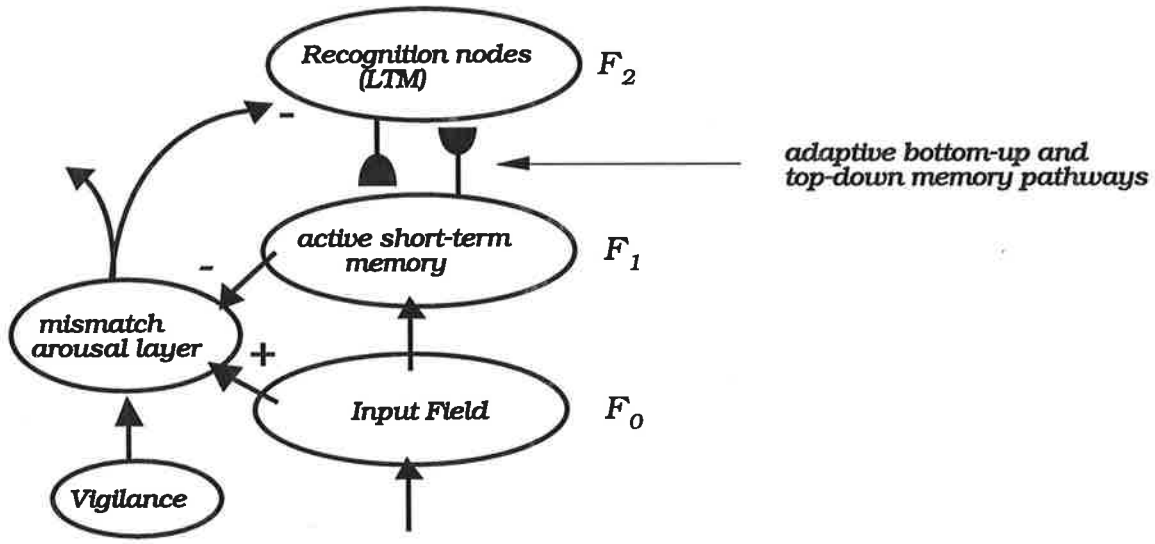


FIGURE 1.6. Basic ART concepts of bottom-up and top-down learning in a system capable of detecting the match/mismatch between the input and the recalled memory. Learning occurs when the spatial patterns across Fields F0 and F1 are matched to within a specified tolerance (vigilance) level.

The established top-down memory, when activated or recalled by the bottom-up input, serves as the internal supervisor against which all the inputs are compared. Since it is possible to store many different visual memories, this in effect provides for a large number of supervising memories. The main concept introduced in ART (reviewed in more detail in Chapter 3) is that stable learning in real-time neural systems is achieved by attentional mechanisms in which a reverberation or an *adaptive resonance* (a state during which a memory is modified by a standing wave or a resonating spatial pattern of neural activity that circulates across several interacting neural layers) is set up between the internal representations (the recalled supervisory memory) and the externally driven sensory stimuli. The attentional mechanism are actively engaged in comparing the internal representations with external events and enable learning only when the external events match their internal representations. Resonance occurs between layers that encode the bottom-up inputs and the top-down learned expectations.

Since ART derives its power from the interaction of the top-down memory and the bottom-up inputs, it provides a strong theoretical foundation for a more advanced neural theory of cognitive vision. A general neural theory of biological visual cognition should provide a plausible rationalisation for the existence and the purpose of the massive feedback pathways in the brain.

In addition to the cognitive concepts of attention, top-down priming and vigilance that are introduced in ART, the non-linear mathematics of shunting cooperative-competitive neural systems, introduced by Grossberg (1973,1988) offers a suitable mathematical tool for modelling the dynamics of complex neural systems. Non-linear neural networks have been proven to be successful in the implementation of various ART based neural networks (Carpenter and Grossberg, 1987a, 1987b, 1990) and various other neural network models of brain functions, including models of early vision (Grossberg and Mingolla, 1985a, 1985b). Real-time non-linear neural networks are characterized by non-linear differential equations, such as the one shown below. This is a version of the shunting cooperative-competitive feedback equation with the on-centre off-surround anatomy that was introduced by S. Grossberg (1973, 1988) and which represents the dynamics of a cell in a layer of competitive cells, as schematized in Figure 1.7.

$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)[J_i^+ + f(x_i)] - (C + x_i)\left[J_i^- + \sum_{j=1}^{n} F_{ji} g(x_j)\right] \tag{1.5}$$

where J_i^+ is the total excitatory input into the layer at i^{th} cell position; J_i^- is the total inhibitory input to the i^{th} cell (for more detailed description see Chapter 3 and Appendix A).

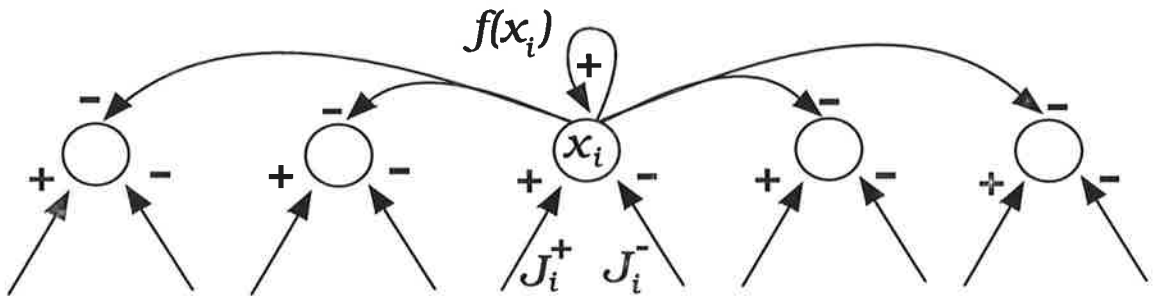


FIGURE 1.7. Layer of shunting competitive neurons.

This thesis therefore uses the non-linear mathematics of competitive neural layers and the concepts introduced in Adaptive Resonance Theory to propose a suite of neuro-engineering design principles, mechanisms and design logic for complex multi-layered 2-D artificial neural circuits with feedforward-feedback interactions. The neuro-engineered circuits described in the thesis are capable of self-regulated attentional learning, selective attention and memory guided processing, autonomous detection of novelty and familiarity and the recognition of familiar objects in cluttered visual images.

1.5 Outline of the Thesis

This thesis is principally concerned with the development of a neural theory of high level primate vision, with a particular emphasis on the interaction between attention, memory, self-regulation and object learning and recognition in complex and cluttered visual scenes. The thesis therefore attempts to unify the cognitive data from several experimental disciplines, most notably the psychological and neurophysiological data on selective attention in order to propose a generalisation of Adaptive Resonance Theory within the domain of attentive vision in realistic visual sensory inputs.

We thus provide a plausible account for the role of the massive feedback pathways in the brain and propose neuro-engineering design principles, mechanisms and design logic for two dimensional competitive neural circuits that are capable of object recognition in complex and cluttered visual scenes. Since the chief aim of the thesis is to provide a large-scale theoretical neural model of selective visual attention and related phenomena of cognitive biological vision, we are currently not concerned with the problems of implementation of the proposed 2-D neural circuits for real-time technological applications. All computer simulations are carried out on a general purpose workstation (486 PC) and are therefore limited to small scales. However, since the property of the proposed neural layers is independent of the lateral extent of the layer, the generated neural circuits are equally applicable to more realistic sizes of primate vision. Numerous computer simulations of 2-D neural circuits are presented in various chapters of the thesis to demonstrate the mechanisms and the dynamics of the generated circuits. The capability of the generated neural circuits and mechanisms is also demonstrated on grey level images. Simulations on grey level images demonstrate some of the key properties of the proposed neural circuits on the recognition of objects in cluttered images as well as the effect of dynamic attentional tuning of early vision feature detecting cells or filters.

In Chapter 2 we review the relevant psychophysical and neurophysiological data as well as the data from invertebrate learning. The neuroscientific data from invertebrates, particularly that of associative learning in the simple neural circuit controlling the gill withdrawal reflex of the sea snail *Aplysia*, is of particular theoretical interest and relevance to the central theme of the thesis because it provides the simplest form of evidence that modulatory mechanisms play a crucial role in biological neural systems. Although largely ignored in current artificial neural network models of learning, the synaptic modulatory mechanisms that have been found to be necessary for strengthening certain synapses in the *Aplysia* circuit, supports the proposal that learning is gated by attentional arousal. While the invertebrate data comes from extremely simple neural circuits whose interconnectivity is very low and whose neurons have an extremely small number of synaptic inputs (in some cases one or two input synapse), the data from more complex biological systems, such as that of monkey's visual cortex, also indicates the existence of modulatory mechanisms. For example, some of the recent data from visual areas V1, V2 and V4 indicates that the feature detecting cells of the early visual pathways do not have fixed receptive fields (as is currently modelled by 2-D Gabor functions) but that they are subject to modulation by the attentional factors (Moran and Desimone, 1985; Haenny *et al.,* 1988; Haenny and Schiller, 1988; Motter, 1993). We also review the latest neuroscientific experiments that demonstrate that top-down feedback from memory can directly influence the attentional processes in a cognitive visual task.

Chapter 3 provides a review of the major mathematical tools, theories and models of vision and learning. Since the thesis has its mathematical and scientific foundations in various non-linear neural networks of S. Grossberg and his colleagues (for a general summary of this field refer to Grossberg, 1988), we will primarily review their most relevant models. In particular we provide qualitative analysis of ART in simple and complex and cluttered visual images. This analysis, supported by computer simulations of ART-3's processing fields, will reveal that while ART is a powerful theory, it has not solved some of the major problems. We identify the weakness of ART to be in its attentional subsystem, which, because of its rigidity, does not allow memory to be effectively used to guide attentional processes. We then derive the theoretical concept of Selective Attention Adaptive Resonance Theory (SAART) and propose a new interaction, which not only overcomes the limitation of the ART model but also predicts the existence of modulatory feedback pathways in the primate visual brain.

In Chapter 4 we lay foundations for the remainder of the thesis by proposing a suite of fundamental neuro-engineering design principles, synaptic mechanisms and a family of new competitive neural layers, collectively called Presynaptically Modulated Shunting Competitive Neural Layers (PM-SCNLs). We thus combine the experimental neuroscientific data of synaptic signal transmission in simple invertebrates, particu-

larly the data from the gill withdrawal reflex of the sea snail *Aplysia* (Kandel and Schwartz, 1982; Gingrich and Byrne, 1985; Hochner *et al.*, 1986; Carew, 1987), with the most recent data from primates (Moran and Desimone, 1985; Desimone *et al.*, 1990; Chelazzi *et al.*, 1993; Motter, 1993) to extend the previously published models of chemical synapses (Grossberg, 1968, 1969; Carpenter and Grossberg, 1981) and the models of shunting competitive neural layers (Grossberg, 1973, 1988). More specifically, the chapter addresses the mechanisms of selective information transfer and processing; the mechanisms of synchronization in pulsating and non-pulsating neural layers; the mechanisms of top-down memory guided selective attention and recognition of familiar stimuli in cluttered background; the mechanisms for the modulation of cellular receptive field profiles, and a neural circuit for the recognition of distorted 2-D shapes.

In Chapter 5 we present a mathematical procedure for the parameter design of modulated neural layers to ensure stability and wide dynamic range.

In Chapter 6 we extend the neuro-engineering design principles and provide two dimensional neural circuits of increasing complexity. In particular, we propose neural circuits for more cognitive type of processing in order to address the various cognitive data that was reviewed in Chapters 2. The emphasis is on the dynamics of short term memory, awareness of familiarity/novelty, self-regulated attentional modulation, attentional selection, translation invariant 2-D pattern recognition and memory guided search in cluttered visual inputs. The proposed design logic and neural circuits form the building blocks for neuro-engineering of complex real-time cognitive and perceptual visual neural networks of subsequent chapters and in a general sense, they also form a neuro-engineering foundation for the design of a cognitive artificial visual neurocomputer. The Chapter concludes by presenting a minimum 2-D neural circuit that can simulate and explain a recent neurobiological experiment on the neural basis of memory guided visual search in the anterior temporal cortex of a monkey (Chelazzi *et al.*, 1993).

In Chapter 7 we propose a new self-organising real-time neural network called Selective Attention Adaptive Resonance Theory (SAART). The most significant property of the SAART neural network that is not shared by its predecessors, such as the ART-3 neural network (Carpenter and Grossberg, 1990), is that the SAART network is capable of selective attention to familiar inputs. That is, the SAART neural network can recognize previously learned 2-D shapes when they subsequently appear complete but in a cluttered input. One of the new interactions in the network, *top-down presynaptic facilitation*, enables the SAART neural network to use its established memory to selectively filter the desired object shape from the cluttered input. Although the SAART

model is more general than ART, it is a long way from being a general neural model of cognitive vision. Its major deficiency is an inability to deal with the various spatial transformations of the sensory input.

In Chapter 8 we present the Advanced Selective Attention Adaptive Resonance Theory (ASAART). ASAART extends SAART by addressing the complex issues of perceptual constancy and visual cognition. It is proposed that perceptual constancy is achieved by bidirectional (bottom-up and top-down) transformations of exogenously and endogenously generated signals via multiple and competing pathways whose selection may be regulated by pre-attentive as well as attentive and memory driven processes. Multiple and simultaneous spatial transformations in the bottom-up direction are required to activate stored visual memories whose spatial extent, orientation or location may not initially match the direct bottom-up input pathways. Multiple and simultaneous spatial transformations in the top-down direction (which are equated to imagined visual transformations) are required to enable the recalled memory to match and regulate the bottom-up signals.

In Chapter 9 we conclude the thesis by summarising the main concepts that were introduced. We discuss the main contribution that the thesis offers to the understanding of visual neurobiology and the solutions that it offers to some outstanding technological problems in visual information processing and object recognition.

1.6 Major Contributions of the Thesis

The principal contributions of the thesis are as follows:

- A critical review of the current algorithmic and artificial neural networks approaches to object recognition in cluttered images.
- A critical review of the Adaptive Resonance Theory in simple and complex visual inputs.
- Extension and generalisation of Adaptive Resonance Theory to cluttered visual inputs.
- The development of a new family of shunting competitive neural layers (Presynaptically Modulated Shunting Competitive Neural Layers) and a mathematical procedure for their parameter design.
- Proposal of a suite of biologically plausible neuro-engineering design principles, mechanisms and design logic for complex (multi-layered) two dimensional self-regulating attentional neural circuits.

- Proposal of a new self-organising real-time artificial neural network called Selective Attention Adaptive Resonance Theory (SAART) for cluttered inputs.
- Demonstrated applicability of the designed neural circuits and networks to the recognition of familiar 2D shapes of 3D objects in cluttered visual background.
- Proposal of the Advanced Selective Attention Adaptive Resonance Theory (ASAART) concepts for the generalisation of the SAART neural network to translation, size and orientation invariant object recognition.

Chapter 2

Neurophysiological and Psychological Background

"The cerebral cortex in general and the visual system in particular are too complex to understand from an exclusively intuitive analysis. Accordingly, there is a pressing need for specific models which are motivated by biological issues and which capture as much of the underlying structural and physiological data as possible, yet which can account for specific tasks or aspects of system performance. It is also critical that such models be robust and efficient in handling the types of noisy data presented by natural images and other real-world sensory input."

D.C. Van Essen, D.J. Felleman, E.A. DeYoe, J. Olavarria and J. Knierim (1990)

2.1 Introduction and Overview

Up to 1950's neurobiology was largely descriptive and it was not until the 1960's that the systematic experimental study begun to reveal the modes of neuronal communication, the types of cells and the architecture of biological neural systems (Whitfield, 1984). Modern anatomical data beyond the primary visual cortex (Van Essen *et al.,* 1990) has revealed that the connectivity among many visual areas is far more complex and intricate than the simple serial (feedforward) processing that characterizes most of the current artificial neural network models of pattern recognition. Although the primate visual system is too complex to be understood from a purely intuitive and mathematical analysis, the impressive body of experimental data about cortical structure and the properties of cells obtained from single cellular recordings does indicate that the visual system is highly organized, both in its structure and the cellular properties. From a neuro-engineering perspective, highly organized structures

(whether they be at the local level of an isolated neural circuit or at the level of complex neural systems) imply the existence of some fundamental design principles that are used repeatedly. Since the organization of the primate visual system probably did not arise instantaneously but may have evolved through numerous refinements of older more primitive and simpler structures, unravelling its architectural design principles requires that careful attention be paid not only to the experimental data from primate neurophysiology, but also to the data from the simplest studied organisms. It is therefore important that plausible neural models of primate vision, even if they are initially very crude, be founded on flexible design principles and neural circuits that could be easily extended to capture much of the structural, physiological and cognitive data as possible. Knowledge of the relevant experimental data (while itself being a major feat for a novice) thus forms a cornerstone around which one can begin to propose useful design principles.

In this chapter we review what we believe to be some of the most relevant neurobiological and psychological data from which a neuro-engineer can begin to formulate 'neurobiologically inspired' design principles. The most interesting property of complex neural systems that has been revealed through experimental work is the existence of highly specialized local neural circuits and reciprocal interconnectivity between many of them. Since the data from complex neural systems, such as the visual cortex of a monkey, is often hard to interpret in terms of the underlying mechanisms and circuit dynamics, it is even harder to discover the design principles upon which such circuits are based. The experimental data from invertebrates on the other hand, even if not related to primate vision, may be more useful in discovering some of the most basic neurobiological design principles because of the simplicity and the predictability of the underlying neural circuits. We will therefore review some of the key experimental data from invertebrate learning, particularly the mechanisms of associative learning in the simple neural circuit controlling the gill withdrawal reflex of the sea snail *Aplysia* (Kandel and Schwartz, 1982; Hawkins *et al.*, 1983; Carew, 1987). The experimental data obtained from *Aplysia* over the past two decades is of immense importance, not only because it reveals some of the most basic dynamics of modulatory mechanisms in biological neural systems but because the circuit embeds what is possibly the simplest form of selective information processing and selective learning.

In Appendix D we will discuss the psychophysical data on shape constancy in relation to the concepts introduced in the latter parts of the thesis.

2.2 General neuroscience

The fundamental computational unit of biological neural circuits is considered to be a cell or a neuron (Levitan and Kaczmarek, 1991), which in more complex systems (such as the visual cortex of a monkey) samples its inputs via a large number of input

pathways. When a neuron receives sufficient input to drive it above a threshold, it fires an electrical pulse (an action potential) which is then distributed to a large number of other cells via the output pathway called axon. Although it had been correctly predicted as early as 1905 that neurons often communicate with one another chemically, it had been generally regarded, until 1950's, that the central nervous system (CNS) did not use chemical neurotransmission. The reason for discarding the idea of chemical signal transmission was based on the grounds that only electrical processes would permit the necessary rapid transfer of signals between various brains regions (for an overview see Eccles, 1964). This view has since changed and it is now widely accepted that both electrical and chemical form of signal transmission is used throughout the brain. A neuron may thus communicate with other neurons either via electrical or chemical signals, the latter being much slower.

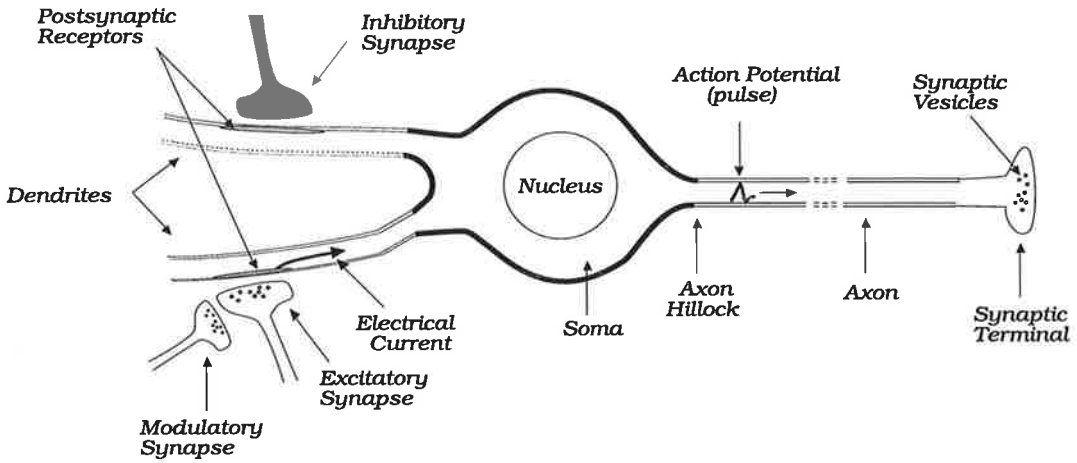


FIGURE 2.1. Schematic of a typical neuron with inhibitory and excitatory chemical synapses that may also be subject to modulation.

A typical neuron, shown in Figure 2.1, has a number of dendrites that form a dendritic tree. Dendrites sample their inputs from other neurons via chemical synapses (which may be either excitatory or inhibitory or both) that abut the dendritic branches. When a presynaptic neuron is active, it releases a certain amount of chemical neurotransmitter that slowly diffuses across the junction to be picked up by receptors on the postsynaptic site and converted into an electrical postsynaptic potential. Whether the postsynaptic cell fires an action potential depends on the magnitude of the combined excitatory and inhibitory postsynaptic potentials.

Chemical Synapses

Although largely ignored by the general artificial neural network community, chemical synapses, having a slower time scale than the electrical form of signal communication, have some very interesting properties that if properly modelled may

be very beneficial to a neuro-engineer. Experimental data gathered over the past four decades has revealed that the internal dynamics of chemical synapses depends on a number of inter-related mechanisms whose time scales may differ widely. For example, the three most important events in a chemical synapses which may be directly related to various forms of memory are transmitter production, storage and mobilization (Eccles, 1964). Transmitter production, being a relatively slow process, can thus be considered to represent the long term average memory trace of signals that have activated the synapse and led to its modification. Transmitter mobilization, on the other hand, being much faster can be considered to represent the most recent signal that has activated the synapse.

It has been suggested (Singer, 1987) that chemical synapses may be in two states or modes of operation: (i) a relay mode during which the synapse transmits the presynaptic signals to the postsynaptic cell without any long term modification of its internal dynamics; and (ii) a plastic mode during which the synapse gets modified, generally by an increased level of transmitter production.

Postsynaptic Feedback

The internal state of a chemical synapse may in general be influenced by feedback from postsynaptic cells. Neurobiological data suggests that the postsynaptic feedback signals, possibly mediated by nitric oxide, have a facilitatory increase on transmitter release (Bohme *et al.*, 1991; Schuman and Madison, 1991; Bredt *et al.*, 1990; Gally *et al.*, 1990; Haley *et al.*, 1992; O'Dell *et al.*, 1991).

Presynaptic and Postsynaptic Receptors

Postsynaptic receptors, which also exhibit time variant properties (i.e, they may be either active or inactive), convert the chemical signals into electrical pulses (Changeux, 1993). Forward transmitter diffusion causes the binding of the released transmitter to the postsynaptic receptors thus evoking an excitatory (or an inhibitory) postsynaptic potential. However, in addition to forward diffusion of the released transmitter onto the postsynaptic cell, some is returned back into the original synapses via the receptors of the chemical synapses, called *autoreceptors*. The effectiveness of some postsynaptic receptors, such as N-methyl-D-aspartate (NMDA) decreases with repeated use (Constantine-Paton, 1990; Bekkers and Stevens, 1990), i.e., they become desensitized. Because the NMDA receptor requires the binding of at least two distinct neurotransmitters (glutamate and glycine) for their activation (Johnson and Ascher, 1987; Bekkers and Stevens, 1990), it was suggested that NMDA receptors may in fact be performing a logical AND operation, such as in detecting temporal correlations in pre- and postsynaptic activity, (Collingridge, 1987; Friedlander *et al.*, 1993).

Neuromodulation

As shown in Figure 2.1, some chemical synapses, rather than being involved in neurotransmission, may be involved in modulation of other synapses, i.e., *neuromodulation*. Unlike neurotransmitters that basically transfer signals from the synaptic junction, neuromodulators may affect the synthesis of neurotransmitter, its release and interaction with postsynaptic receptors (Barchas *et al.*, 1978). A neuromodulator may in general act simultaneously on a large number of neurons with long lasting effects. In section 1.3.1 (Chapter 1) we have mentioned how the concept of neuromodulation via an axo-axo-dendritic synapse was modelled by Seibert and Waxman (1990, 1992) in the context of a neural system for 3D object recognition (association of viewed 2D aspects of a 3D object during aspect transitions). In the next section we discuss neuromodulation in the context of a simple biological neural circuit where it provides a facilitatory effect on the efficacy of neurotransmission and also seems to be necessary for the long term modification of the synapse.

2.3 Presynaptic Facilitation in Aplysia

In Figure 2.2 we reproduce a portion of a neural circuit that controls the gill withdrawal reflex in the sea snail *Aplysia* (Levitan and Kaczmarek, 1991). The figure shows the interaction between three neurons: sensory neuron (SN), the motor neuron (MN) and a facilitatory neuron (FN).

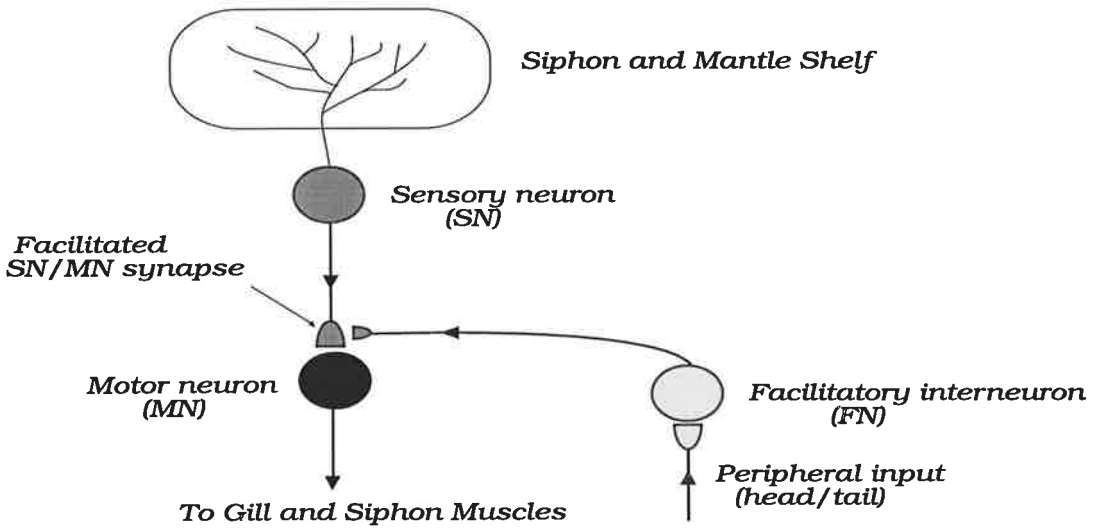


FIGURE 2.2. Facilitation of SN/MN chemical synapse in *Aplysia*. From Levitan and Kaczmarek (1991).

The facilitatory interneuron in *Aplysia* influences the sensory/motor neuron (SN/MN) synapse by increasing the amount of transmitter that is released by the presynaptic spike from the sensory neuron. This presynaptic facilitation enables a habituated sensory pathway to be quickly dishabituated when a noxious stimulus is applied to other parts of the animal (typically the head or the tail). Although the facilitatory interneuron releases its own neuromodulatory transmitter into the presynaptic terminal whose efficacy it modulates, it does not appear to be directly involved in the release of the transmitter of that pathway. That is, the presynaptic transmitter in the pathway from the sensory to the motor neuron synaptic junction (synaptic cleft) is not released unless it is acted upon by the presynaptic spike from the sensory neuron. The action of the facilitatory interneuron can thus be considered to be priming the SN/MN synapse. Experimental data on learning in *Aplysia* also provides evidence that in addition to the increased concentration of releaseable transmitter after the firing of the FN neuron alone, the SN/MN synaptic pathway is strengthened even further during the correlated firing of SN and FN neurons. The latter form of synaptic modification has a longer lasting effect on the synapse.

Presynaptic facilitation in sensory neurons has been attributed to transmitter mobilization (Gingrich and Bryne, 1985; Hochner *et al.*, 1986; Gingrich *et al.*, 1988; Bhara *et al.*, 1990). If presynaptic facilitation enhances transmitter mobilization then we can assume that the effect of presynaptic inhibition is to inhibit transmitter mobilization. There is no plausible reason why similar synaptic events, perhaps mediated by different neuromodulatory transmitters and by a distributed network of cells, should not occur at higher cortical layers of the primate visual system. Given the importance of presynaptic modulation as means of selectively amplifying the neural signal transmission in as simple organism as that of *Aplysia*, it is highly likely that neural systems of highly evolved organisms would also depend on similar (and perhaps more complex) mechanisms of neural signal modulation.

2.4 Primate Visual Neurophysiology

Since most of the experimental data on primate visual neurophysiology has been obtained from the monkey visual cortex, the data reviewed below will thus be primarily from monkeys, although, where relevant, the data from other animals (e.g., rats and cats) will be mentioned.

Transient and Sustained Channels

The early visual pathways in the primate retina have been found to consist of sustained ganglion cells or X-cells and transient or Y-cells (Enroth-Cugell and Robson, 1966). Sustained cells respond maximally to stationary stimuli within their receptive

field, have a long response latency and are affected by image blurring, whereas transient cells give transient responses to light onset or offset, are sensitive to rapid motion, have a short response latency and are unaffected by blurring. This early division between the sustained and transient cells reflects an early distinction between neural layers that process form and motion independently. Output of sustained cells may be the primary factor determining the pattern or form recognition, whereas output of transient cells may be important for the detection of flicker and motion. Because of the different temporal responses of the two channels, transient activation by the stimulus will become available prior to sustained activation. Physiological evidence (Singer and Bedworth, 1973) and the data from visual masking experiments (Breitmeyer and Ganz, 1976; Breitmeyer and Valberg, 1979; Breitmeyer, 1978, 1980) indicate that transient and sustained visual channels mutually inhibit each other.

Visual Pattern Recognition Pathway

The retinal information is routed to the striate cortex (area V1) via Lateral Geniculate Nucleus (LGN), where it is analysed by feature detecting cells that are tuned to various spatial frequencies and orientations. From there it is routed via areas V2 and V4 to the inferior temporal cortex (DeYoe and Van Essen, 1987; Desimone and Ungerleider, 1989).

In an influential study on the orientational selectivity in visual cortex, Hubel and Wiesel (1959, 1962, 1963, 1977) have found that neurons in the cat's striate cortex respond optimally to elongated bars that are oriented in a specific direction. As the experimental work (anatomical, lesion and cellular recordings) progressed over the past three decades, it has been revealed that there are something like 30 different areas in the monkey's brain that are related to vision and which constitute about 60% of the overall cortex (Van Essen and Maunsell, 1980). Most of these areas are organized as two dimensional maps, such that higher cortical areas, although having cells whose response is dependant on the larger areas of the stimulated retina, are in retinotopic register with cells at the lower layers. Primary visual cortex or area V1, being the first visual area where feature detecting cells are found, thus consists of an extremely large number of cell whose receptive fields are relatively small (about 0.5 degrees of visual angle) compared to the receptive fields of cells in area V4 (around 5 degrees). As one progresses to the inferior temporal (IT) cortex, cells are typically sensitive to very large areas of the retina and, unlike the lower visual areas, lack any discernible topography.

Inferior Temporal (IT) Cortex and Long-term Visual Memory

The visual information from the visual area V4 feeds into a large expanse of visual cortex called inferior temporal (IT) cortex. The IT cortex of the primate temporal lobe is considered to be the ultimate visual processing area (Gross, 1973a, 1973b;

Mishkin, 1982). Numerous lesion experiments have shown that monkeys whose IT cortex is removed fail to visually recognise objects and are deficient in learning and remembering visual discrimination habits (Chow, 1954; Mishkin, 1966; Humphrey and Weiskrantz, 1969; Gaffan and Harrison, 1986). However, their lower level visual functions remain intact, such as the threshold for the detection of a brief flash, backward masking functions and critical flicker frequency (Gross, 1973a). IT neurons, especially those in the anterior portion called TE, have large receptive fields that often include the fovea and respond to highly complex visual features (Gross *et al.,* 1972; Desimone and Gross, 1979; Perrett *et al.,* 1982; Richmond *et al.,* 1983; Miyashita, 1993). Experimental data also implicates IT neurons in short-term memory and selective attention tasks (Fuster and Jervey, 1981; Baylis and Rolls, 1987; Fuster, 1990; Miller *et al.,* 1991, 1993; Colombo and Gross, 1994); their activity is enhanced and modulated during visual discrimination and selective attention (Sato *et al.,* 1980; Richmond and Sato, 1987; Sato, 1988; Fuster, 1990); they are insensitive to pattern transformations such as size, orientation and translation (Desimone *et al.,* 1984; Gross and Mishkin, 1977; Schwartz *et al.,* 1983; Sary *et al.,* 1993). Single-unit recordings from IT cells (Perrett *et al.,* 1985; Hasselmo *et al.,* 1989) support the hypothesis of an "viewer-centered representation" of a viewed 3D object while "object-centred representation" may also be needed for 2D representations (i.e., the representation may depend on the centroid of the 2D shape). It has also been proposed that IT cortex is the convergence site of memory and perception (Miyashita, 1993).

IT cortex has been subdivided into two major areas, TEO and TE. On the basis of recent anatomical tracings, the classic IT cortex has further been subdivided into three major areas, PIT, CIT and AIT (Van Essen *et al.,* 1990). Each of these is also found to consist of two parts that have different anatomical connections.

Inferior temporal cortex thus plays an important role in shape recognition and the storage of long-term visual memories. It is now established that lesions of area TEO causes devastating impairment in the ability of monkeys to learn visual pattern discrimination that is even more severe than the impairment that is caused by the lesion of TE (Cowey and Gross, 1970; Kikuchi and Iwai, 1980; Weiskrantz and Saunders 1984; Gaffan and Harrison, 1986). Lesions of TEO makes it almost impossible for monkeys to discriminate between simple 2-D patterns differing in form, size, orientation, color or brightness (Iwai and Mishkin, 1969; Gross, 1973; Ungerleider and Mishkin, 1982). However, ablasions of the area TEO does not impair visual memory (Iwai and Mishkin, 1968; Delacour, 1977). It is now believed that area TEO is responsible for perceptual stability and constancy (i.e., ability to recognize a shape despite changes in its size, position and orientation). The receptive fields of TEO neurons are intermediate in size between those of V4 and TE (Boussaoud *et al.*, 1991). Through extensive probing of the IT pathway, (Tanaka, 1993; Tanaka et al., 1993), it was found that as one traverses

from posterior to more anterior portions of IT that the cells respond to more complex visual features, although it is not yet experimentally established how IT cortex establishes its specificity for object features and what type of a neural circuit may mediate such learning.

IT cortex interacts with many other cortical areas that are involved in memory retention tasks, such as the prefrontal cortex (Fuster *et al.,* 1985; Bauer and Fuster, 1976; Fuster and Alexander, 1970). In addition to providing further evidence that the visual information is split into the 'what' and the 'where' pathways (Ungerleider and Mishkin, 1982) one of the interpretations of the interactions between the TEO area and the prefrontal cortex is that TEO probably uses the prefrontal cortex for the temporary storage of visual information. Other experimental data from monkeys (Goldman-Rakic, 1987) provides evidence that the prefrontal cortex is involved in short term memory storage during delayed periods when a monkey is required to hold a stimulus in memory to compare it against the next.

Backprojection Pathways

Experimental data on large areas of the visual cortex in monkeys shows that the forward axon projections are typically equivalent to or are outnumbered by backprojections (Tigges *et al.*, 1973; Rockland and Pandya, 1979; Van Essen and Anderson, 1990; Van Essen *et al.*, 1990). Thus a projection form one area to another is matched by a projection in the reverse direction. However it soon became apparent that these reciprocal connections are usually asymmetric and are not merely reciprocating the feedforward connections (Rockland and Pandya, 1979; Friedman, 1983; Maunsell and Van Essen, 1983).

These feedforward-feedback pathways, especially at the earliest stages of visual processing where the operations are typically considered to be performing filtering and feature extraction (the loop between V1 and the LGN) is in marked contrast to the conventional feedforward engineering methods of filtering, noise removal and feature extraction. However, the function of the massive feedback pathways in the primate visual system is not yet understood. In a recent experimental study on the effect of the feedback from visual area V1 to LGN, Sillito *et al.* (1994), have found that the cortical feedback induces coherent oscillations in a pair of LGN cells that a driven by a stimulus that crosses their receptive fields. On the basis of their experimental results, Sillito *et al.* have concluded that "the feedback circuit searches for correlations that support the 'hypothesis' represented by a particular pattern of cortical activity and 'locks' the esemble onto the stimulus". Although these researchers have not suggested what type of feedback can test for the presence of the relevant pattern in the input cells, it seems obvious that it cannot be achieved by a mutual excitatory loop between LGN and V1. Whatever the nature of feedback, it must allow LGN cells to be sensitive to their inputs

from the retina and it must be able to synchronize those LGN cells that are activated by the retinal signals. Sillito *et al.* have suggested that the cortico-thalamic input is only strong enough to exert an effect on those LGN cells that are also depolarized by their retinal input. Could it be that the cortico-thalamic feedback in fact facilitates the input receptive field profiles of LGN cells?

In a recent extension of the Boundary Contour System and Feature Contour System (Gove et al., 1995) a top-down corticogeniculate feedback (top-down) was added to realize a matching process which enhances LGN cells that are consistent with those of active cortical cells while suppressing LGN activities that are not. In the implementation of the neural circuit, this feedback multiplies with the bottom-up retinal input to provide additional forward excitation of the LGN cells (see equation (8) -(10), Gove et al., 1995).

Selective Attention and Memory Guided Search

A number of neurophysiological experiments have demonstrated that the cellular receptive fields in the visual pathway from area V1 to IT cortex of an alert monkey are modulated by attentional factors. On the basis of their experimental results Moran and Desimone (1985) have concluded that spatial attention serves to remove irrelevant stimuli from the receptive fields and sharpen their selectivity to the attended stimulus. In a related experiment (Spitzer *et al.,* 1988; Desimone *et al.,* 1991) have obtained data which shows that when a monkey had to perform a difficult pattern recognition task that the response of V4 neurons was larger and more selective to the stimulus orientation.

Of all the neurophysiological experiments on selective attention that appear in literature, the one most relevant to the central theme of this thesis was performed by Chelazzi *et al.* (1993). These researchers have obtained experimental data from monkey's IT cortex that implicates this cortex in the control of the top-down memory guided selective attention and visual search in a cluttered visual display. Further experimental support for the influence of the higher order feedback systems (top-down signals) on the receptive fields of cortical neurons (V4) comes from the work by Haenny and Schiller (1988) and Haenny *et al.* (1988).

Response Decrement and Indication of Familiarity/Novelty

Several experiments on a working memory task (where an animal is required to hold a stimulus in memory) have shown that the response of IT neurons was attenuated if an incoming stimulus matched a stimulus held in memory (Miller *et al.,* 1991, 1993; Gross *et al.,* 1979; Baylis and Rolls, 1987; Eskandar *et al.,* 1991). Based on these results, several researchers have proposed that the IT cortex is comprised of two opposing classes of cells. For example, Miller *et al.* (1993) have proposed the existence

of an "adaptive mnemonic filter" class whose activity are compared with a "sensory" class. On the other hand, Eskandar *et al.* (1992a, 1992b) have proposed that IT cortex contributes primarily to the comparison process, but it does so on the result of temporal modulation of the visual input with the recalled memory. Miller *et al.* (1993) have also proposed that the comparison process works on just the stimuli presented within a trial.

Since the general decrease in cellular response has also been found in many other cortical areas (Riches *et al.,* 1991), it is unlikely that IT neurons indicate whether a stimulus is familiar or novel (although their response may be needed in the actual decision process). The most probable site of neurons that do the actual matching process seems to be located in the medial thalamus (Rolls *et al.,* 1982; Fahy *et al.,* 1993a,b).

2.5 Psychological Background on Selective Attention

Whereas the neurophysiological data on selective attention begun to appear about 20 years ago, the psychological data extends to almost a century with the early introspective work of William James (James, 1890) and Von Helmholtz (Helmholtz 1866/1925). However, the concept of selective attention to sensory stimuli has gained its foothold in the mainstream of psychology in 1960's through the influence of two theories, the 'filter theory of attention' (Broadbent, 1958) and the 'response selection theory of attention' (Deutsch and Deutsch, 1963). The common thread in these theories is the emphasis on the selective allocation of some limited processing resource, although there is no unified view on what that resource is. Below we provide a brief overview of the psychophysical data on selective attention that was obtained from humans.

Traditionally, selective attention has been likened to a filtering process that can either amplify the target information, attenuate the background information or a combination of both may take place (Broadbent, 1958; Deutsch & Deutsch, 1963; LaBerge & Brown, 1989). What all these theories have in common is the fact that there is a processing bottle-neck, although there is a disagreement on where in the neural system this bottle-neck occurs. In Broadbent's theory the bottle-neck occurs early in the processing, whereas Deutsch & Deutsch place the bottle-neck much nearer the response end of the system.

Spatial Allocation of Attention

Through an experimental paradigm that cues the locations in the visual field prior to target stimulus presentation, Eriksen *et al.* (1985, 1987, 1988) have obtained data which led them to conclude that the benefits of precueing (which are measured by the reduced reaction times) are attributed to the alignment of the attentional spotlight on the appropriate locations of the visual field, thus speeding the processing of the

stimulus. They have suggested that a variable power or *zoom lens* is a more apt analogy for the operation of visual attention rather than the *spotlight* metaphor of attention that had been proposed by Posner (1980). The zoom lens has the property of wide field of view and low resolution at low lens power and a smaller field of view but a higher resolution at high lens power.

The *gradient* model of attention, proposed by LaBerge and Brown (1989), assumes that visual attention is resource limited and that the resources can be directed over a space of varying size. The resources fall off continuously from the centre of the focus of attention as a function of spatial distance and the stimulus identification time varies with distance from the gradient peak. Accordingly, attending to a small region of space compared with a large region is equivalent to creating a more narrow peak in the resource gradient. Henderson (1991) conducted experiments where both the cued area size and cue-target spatial relationship was manipulated. The main finding was that a smaller cue led to greater facilitation at the cued location and greater inhibition at uncued locations. These findings seem to support the gradient model of attention.

Reaction Times and Response Competition

The experimental data from attentional cueing experiments shows that precueing target locations prior to target stimulus presentation has an effect on the reaction time (time to recognise or respond to the target stimulus). Subjects are faster in detecting and/or recognising a target stimulus when the spatial cue is valid than in the neutral (non cue) condition. However, when the cue is invalid, reaction times are slower that in the neutral condition. In addition, RT's for validly cued targets are typically faster as the interval between the cue and the target increases (Tsal, 1983; Shulman *et al.*, 1979). Pre-cueing target locations prior to stimulus presentation has a facilitating effect on the reaction time if the target is presented at the cued locations. Otherwise, the cue has an inhibitory effect with a corresponding increase in reaction time compared to a neutral (non cue) condition). This facilitory/inhibitory effect is dependant on the cue-target inter-stimulus interval (stimulus onset asynchrony or SOA) and on the distance of the cued location from the fixation point. Figure 2.3 shows a typical reaction-time data that is obtained from target cuing experiments. The reaction time decreases as the SOA is increased until it reaches a plateau (at SOA's of about 400 ms) after which the reaction time begins to increase. Results also show that RT's to validly cued targets increase with the distance of the cued position from the fixation centre.

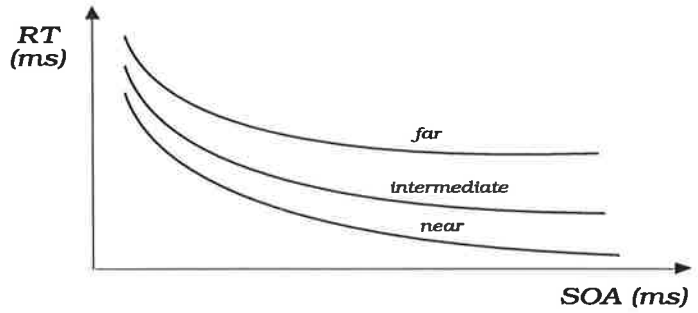


FIGURE 2.3. Typical reaction times in target cuing experiments as a function of the cued position from the fixation centre and at different cue-target intervals. From Eriksen (1988), pg. 14.

The reaction time to a target stimulus also increases when several irrelevant visual stimuli are simultaneously presented with the target. This increase is attributed to the response competition effect. It has been demonstrated that the response competition is quite robust and that the extent to which target incompatible stimuli affect the reaction time varies inversely with the distance between them, (Eriksen and Eriksen, 1974; Eriksen and Schultz, 1979; Miller, 1982). In addition to the distance related effect, it was also found that the interference between the target and the non-target stimuli has a component whose magnitude depends on feature similarity. Thus, if a target is flanked by stimuli with which it shares some features then the reaction time to the target is elevated. Response competition experiments have shown that while there is an irreducible minimum size of the attentional spotlight (of the order of about 3 degrees of visual angle) it can be expanded to cover large areas of the visual field (but with a corresponding reduction in the efficiency of stimulus processing).

Speed of Attentional Shifts

Although the psychophysical data from target cueing experiments demonstrates that there is a facilitory effect of pre-cueing target locations on the reaction time and/or recognition accuracy, these experiments do not provide conclusive data on the speed of attentional shifts and the distribution of attention across the visual field. Shulman *et al.*, (1979) have concluded that attention moved across the visual field in an analog manner at a constant rate and that all intervening stimuli are processed. On the other hand Tsal (1983), although confirming the analog movement of attention, rejected the notion that the intervening stimuli are processed. These conclusions are in conflict with that of Eriksen and Murphy (1987) whose experiments support the discrete shifts of attention (as opposed to a sweep).

Exogenous and Endogenous Control of Attention

In cueing experiments mentioned above, subjects are typically cued to a peripheral spatial location by a high contrast visual marker in advance of the target stimulus presentation, with the marker located at or near the impeding location of the target stimulus. Experimental data taken during brief presentations (brief enough to preclude overt eye movements) suggests that such a cue automatically captures attention. Attention may also be subject to voluntary control. For example, a centrally presented arrow may be another form of visual cue that indicates to the subject the peripheral location of an impeding stimulus (Posner *et al.*, 1978). Visually informative cues, such as an arrow, do not automatically cause attentional shifts away from the fixation point but rather require higher level processing to first interpret the cue (direction of the arrow) followed by a voluntary shift in the direction of the arrow. Voluntary shifts of attention may also be driven by other endogenously generated signals.

Two processes thus seemed to be involved in the control of visual attention: an automatic process that is driven by the property of the visual stimulus, and an endogenous process that is voluntary and allows subjects to have internal control over the spatial allocation of attention (Weichselgartner and Sperling, 1987; Posner, 1980). It is quite likely that the two attentional systems are coupled. This would be advantageous to an intelligent vision system in a natural scene when the exogenous component of visual attention may often be attracted to non-relevant stimuli, such as when the object of interest is camouflaged, partially occluded and surrounded by clutter and noise. In such cases the endogenous system, which is knowledge driven, would override the exogenous system either in spatial location or size of the attentional field, thus increasing the detectability of the object. When this fails, it may be necessary to shift fixation across several locations.

Some Deficits in Vision and Visual Attention

Compared to the psychophysical evidence mentioned above, the neuropsychological data was obtained from human subjects who have some sort of a defect in their visual ability. Here we will cite several well studied cases that expose the nature of vision and visual attention from a different perspective and, to some extent, reveal how deficit in one visual function affects another (and which functions are dissociated). For example, there are several striking examples that point to the dissociation between object identification function and those functions that are of a more spatial nature (Holmes 1919). One of the most commonly occurring and intriguing visual deficits is unilateral visual neglect (Bisiach and Luzzatti, 1978; Bisiach *et al.*, 1981). This deficit reflects a disturbance in the spatial distribution of attention and is manifested by patient's inability to notice objects presented in the affected visual field (contralateral to the site

of brain injury). Patients with left-sided neglect seem to ignore stimuli that fall to their left. This has been demonstrated on a number of visual tasks: crossing-out tasks, copying viewed drawings, drawing an object upon a verbal instruction, etc. Evidence from recent studies, Riddoch and Humphreys (1987), supports the view that this deficit is due to the malfunction of the visual attentional system. Often, patients with this visual defect can be made to attend to previously ignored stimuli if verbally prompted to do so. This seems to indicate that the bottom-up defective attentional visual pathways may be overridden by controlled top-down attentional pathways.

Closely related defect to unilateral neglect is visual extinction (De Renzi, 1982). Patients with this defect can identify a single stimulus when presented singularly in any part of the visual field and yet when two stimuli are presented simultaneously in the left and right visual field they do not seem to notice the stimulus in the affected visual field. An explanation for this deficit appears to be consistent with Posner *et al.*, (1984) findings that patients have a difficulty in disengaging attention once it is engaged.

Based on results of an experimental study of neuropsychological patients, Posner *et al.*, (1984) have proposed that the process of visual attention proceeds in the following three phases: (1) engaging attention to a target; (2) disengaging; and (3) shifting attention from one target to another. Failure of any one of these stages leads to visual neglect and to other related phenomena that have been observed in patients.

2.6 Conclusions

In this chapter we have provided a brief overview of the general neuroscience and the relevant neurophysiological and psychophysical background on primate vision. We have concentrated primarily on the experimental data of the primate visual memory (IT cortex) and selective attention. This data indicates that selective visual attention plays a crucial role in object recognition and memory. At the cellular level, selective attention seems to modulate the receptive field profiles, thus sharpening their selectivity and removing the irrelevant stimuli, while at the behavioural level it determines whether an animal will perceive and remember a visual stimulus. The reviewed data supports the view that selective attention is a selection (or a filtering) mechanism that has spatial and non-spatial components which: (i) influence memory; (ii) are influenced by memory; and (iii) are regulated by the familiarity/novelty of the stimulus or the difficulty of the pattern discrimination task. The general reduction in the level of cortical activation for familiar (or expected) stimuli suggests that memory circuits within the primate visual system rely on the results of a matching process which: (i) compares the contents of memory with the incoming stimulus; and (ii) regulates its activity, the degree of attention, the accuracy of the response and hence the rate of learning.

Since very little is known about the neural circuitry of the higher levels of vision, there is a need to develop a neural model of learning and visual object recognition that captures some of the experimental data on selective attention, response decrement to familiar stimuli, self-regulation, etc. While it may be concluded that the aforementioned experimental data is scattered across several experimental disciplines and is too complex to be incorporated into a single theoretical framework, we will show that the Adaptive Resonance Theory of S. Grossberg (1976, 1980) already captures some of the data and that it can be extended to explain most of the data. The main neurophysiological data that cannot be explained by the current neural network models of attention is how the feedback from higher cortical layers (including memory) can influence the orientational selectivity of lower cortical cells. Since ART embeds top-down signal flow (top-down memory) in a neural architecture that also relies on a matching process, we will show how additional top-down pathways (of modulatory nature) may be incorporated to influence the lower neural layers. The main purpose of the thesis, however, isn't to model and replicate the experimental data, but to use it in order to develop a neural theory of biological vision and neuro-engineering design principles and mechanisms for robust artificial neural vision systems.

Chapter 3

Mathematical and Theoretical Foundations

"How can a learning system be designed to remain plastic in response to significant new events, yet also remain stable in response to irrelevant events? How does the system know how to switch between its stable and its plastic modes in order to prevent the relentless degradation of its learned codes by the blooming buzzing confusion of irrelevant experience?."

S. Grossberg (1987)

3.1 Introduction and Overview

The theoretical work that is presented in subsequent chapters does not only rely on the experimental data that was reviewed in the previous chapter but also on numerous mathematical tools, models, concepts and mechanisms that have been proposed over the past three decades. The major theoretical and mathematical work of the thesis is, however, by far largely based on Grossberg's theoretical work of the period 1960 - 1980, particularly his Adaptive Resonance Theory (ART) and its subsequent implementation into a family of real-time artificial neural networks by Carpenter and Grossberg since mid 1980's.

This chapter will therefore review the relevant mathematical and theoretical background upon which the thesis is based, including some of the most important and recent models and theories of vision, learning and pattern recognition, particularly those that embed potentially useful mechanisms and design principles. We will concentrate primarily on non-linear artificial neural networks whose design principles

address self-organised learning (such as ART based networks, Carpenter and Grossberg, 1987a, 1987b, 1990), or are claimed to be biologically plausible and constrained by the key psychophysical and neurobiological data, such as models of visual spatial attention (Olshausen *et al.*, 1992, 1993; Niebur *et al.*, 1993). Although we have provided a brief review of several related neural networks for 2D shape and 3D object recognition in section 1.3.1 (Chapter 1), in this chapter we will provide a detailed review of Adaptive Resonance Theory (Grossberg, 1976; 1980) and ART-3 neural network (Carpenter & Grossberg, 1990). Other recently proposed theoretical concepts that address the role of the feedback pathways in the brain will be discussed in the light of the neurobiological data that is now available (and reviewed in Chapter 2).

At the end of the Chapter we propose the basic theoretical concept behind our Selective Attention Adaptive Resonance Theory (SAART), which is then fully developed in the subsequent chapters of the thesis.

3.2 Theories and Models of Vision

Marr's Theory of Vision

Marr's theory of vision (Marr, 1982), although not of much relevance to the central theme of this thesis, is mentioned here primarily because it is still considered to be the most influential theory of vision upon which most current computer vision models are based. Marr was influenced by several examples in which top-down information was not needed (e.g., fusing random-dot stereograms). This led him to propose a purely feedforward theory of vision that is based on linear mathematics and which does not address the role of the feedback pathways.

Boundary and Feature Contour Systems (BCS & FCS)

In a series of articles beginning in early 1980's (Cohen and Grossberg, 1984; Grossberg and Mingolla, 1985a,b) a real-time visual processing theory of 3-D form, color and brightness perception began to emerge. Largely founded on the perceptual data from visual psychophysics, the theory is composed of two complimentary and interacting systems, the Boundary Contour System (BCS) and the Feature Contour System (FCS). While addressing some aspects of early vision, such as grouping of textured regions via cooperative linking, as well as being able to explain and simulate many visual illusions, BCS and FCS still fall short of providing a robust solution to figure-ground segmentation in realistic visual scenes (such as those of Figure 1.3, Chapter 1). Nevertheless, BCS and FCS neural networks, although not yet widely accepted by vision researchers, have been evolving rapidly over the past decade and have the potential to eventually explain the key neurobiological design principles of early vision.

Interactive Vision

In a recent review of the current assumptions and methodologies of mammalian vision, which are largely based on Marr's theory of vision, Churchland *et al.* (1994) offer a critique of the current wisdom and methodology of "pure vision" and provide a theoretical outline for *interactive vision*. The basic tenet of interactive vision is that vision is exploratory and predictive. Vision interacts with other sensory modalities to improve motor control and facilitate the organism's success in feeding, fleeing, fighting and reproduction. Compared to the conventional concept of a chiefly unidirectional low-to-high feedforward processing hierarchy, interactive vision suggests that the brain's recognition success in the real-world case depends on richly recurrent networks. The notion of interactive vision suggests an important role for the interaction between memory and vision, whereby stored information (memory) from previous learning plays a role in what an animal literally sees (perceives).

Interactive vision thus recognizes the need for top-down feedback pathways from the higher to the lower cortical layers. In contrast to conventional engineered computer vision systems whose design is serialized into the segmentation problem followed by the recognition problem, interactive vision proposes interactive segmentation-recognition. The strategy of interactive segmentation-recognition may be extremely useful in real world cluttered scenes, where there are many objects that partially occlude one another. Churchland *et al.* (1994) thus note that figure-ground segmentation and recognition are more efficiently achieved in tandem than strictly sequential. They provide the following neurobiological explanation for the difficulty of strictly feedforward visual processing and how it may be overcome with the feedback pathways: "*The problem is essentially that global information is needed to make decisions at the local level concerning what goes with what. At lower levels of processing such as V1, however, the receptive fields are relatively small and it is not possible locally to decide which pieces of the image belong together. If lower levels can use information that is available at higher levels, such as representations of whole objects, then feedback connections could be used to help tune lower levels of processing*".

The proposition that one of the roles of the information flow in the feedback pathways may be to retune the characteristics of lower level cortical cells based upon the interpretation made in higher cortical layers has been made earlier by Tsotos (1991). Other researchers, for various reasons, have also realised that the top-down feedback pathways in the brain cannot be ignored.

Pattern Theory

"Pattern Theory", a term introduce by Grenander (1976-1981) and promulgated as a framework for a theory of vision by Mumford (1991, 1992, 1994) is similar in many respects to the interactive segmentation-recognition strategy. Pattern Theory proposes that in order to successfully reconstruct the world variables, it is necessary to first *synthesize* the coded signals that one observes and then compare the synthesized signals with the observed signals. Hence, the neural architecture implied by Pattern Theory is not purely feedforward, but depends on recursive computations and interactions between the bottom-up and the top-down processes. Pattern Theory thus recognizes that the top-down feedback processes are just as important as the feedforward processes. In summary, Pattern Theory presupposes an analysis-synthesis loop that combines feedforward feature extraction process with feedback loops that enable the system to duplicate the stimulus by combining and transforming its basic features. The purpose of the feedback pathways is to relay the interpretations of higher cortical areas to lower areas in order to verify the high-level interpretation of a scene.

Dynamic Routing Circuits and Visual Spatial Attention

In an attempt to understand the nature of visual spatial attention and translation invariant pattern recognition, Olshausen *et al.* (1992, 1993), Van Essen *et al.* (1994) have introduced a feedforward neural network model of visual spatial attention to achieve translation and size invariant recognition of 2-D patterns. The model uses the dynamic routing circuits (previously proposed by Anderson and Van Essen, 1987) with a set of control neurons that dynamically gate the alignment of the neural input to achieve translation invariant representation of the input stimulus. The model also uses the Hopfield network (Hopfield, 1984) for pattern recognition. These authors have proposed that visual spatial attention has evolved to subserve general purpose object recognition since it is too computationally demanding to have the requisite neural circuits replicated at each location in the visual field. Visual spatial attention also acts as an informational bottle-neck, whose purpose is to reduce the amount of visual data that is to be processed and learned by the higher neural layers. In addition to controlling what is to flow to the higher neural layers, visual spatial attention also preserves the information about spatial relationships within the window of attention. It has been estimated that less than 0.1% of the information from the optic nerve passes through the attentional bottleneck at any moment (Van Essen *et al.*, 1991).

Although the model presupposes the existence of top-down feedback pathways, the information processing role of these pathways has largely been ignored. Since the model does not incorporate the top-down control of attention it cannot explain how the

memory within the system may assist the feedforward attentional processes. Hence the model cannot explain the neural processes that enable the brain to recognize familiar stimuli in cluttered visual backgrounds.

Many other models of visual attention are also limited because they have ignored the role of memory on the attentional processes. For example, the model of attention by Koch and Ullman (1985) provides means for defining the saliency of the input based on relatively-low level cues such as pop-out due to motion, depth, texture, or color. While one cannot deny the crucial role that these low level feedforward processes play in our everyday vision, one certainly cannot ignore the role of memory, particularly in visual scenes where the visual stimulation from a scene does not always lend itself to easy determination of attentional cues by feedforward computations.

3.3 Mathematics of Non-linear Neural Networks

Continuous-nonlinear neural networks, summarised in Grossberg (1988) began to appear in 1960's. They either arose from a direct analysis of behavioural or neural data. The Hodgkin and Huxley model (1952), although focussing on individual cells rather than a network of cells, has provided a foundation for Grossberg's shunting competitive model (Grossberg, 1968b, 1973). The shunting competitive feedback equation, which was subsequently shown to encompass a wide variety of neural models (Grossberg, 1988), can be written in its simplest form as

$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)\{J^+ + f(x_i)\} - (C + x_i)\left\{J^- + \sum_{j \neq i} f(x_j)\right\} \tag{3.1}$$

where J^+ is the total excitatory input into the layer, J^- is the total inhibitory input onto the layer while $\sum_{j \neq i} f(x_j)$ is the inhibitory input that is due to the competitive interactions between the cells in the layer. Function $f(x)$ governs the nature of cellular interactions in a given layer and hence the property of the layer. If $f(x)$ is faster than linear above a threshold, then (3.1) represents a winner-take-all competitive network. Equation (3.1) has contrast enhancing properties whereby strong inputs are enhanced at the expense of the weaker inputs. Each node in a layer may also self-excite through positive feedback. This equation has two automatic gain control terms, $(B - x_i)$ and $(C + x_i)$, which ensure that the dynamics of each cell in a layer remains sensitive to input fluctuations and is bounded in range (-C, B).

In Appendix A, we provide a qualitative analysis of the equilibrium behaviour of (3.1), as well as computer simulations, which show that, from an engineering perspective, a competitive neural layer whose dynamics is described by the above equation has an undesirable characteristic whereby the inhibitory inputs do not effectively regulate the cellular dynamics. We then develop a more general model in Chapter 4, called Presynaptically Modulated Shunting Competitive Neural Layer, whose parameter design procedure and the relevant computer simulations are presented in Chapter 5.

Learning Laws

One of the most popular forms of learning widely adopted by neural network modelers was introduced by Hebb (1949). Hebb has proposed that whenever neuron A causes the firing of neuron B, then the synaptic strength between the two neurons increases. The neuroscientific data obtained from the gill withdrawal reflex in *Aplysia*, reviewed in Chapter 2, provides strong evidence against the biological plausibility of Hebb's proposal. The data clearly indicates that the magnitude of the excitatory post-synaptic potential (EPSP) on the motor neuron decreases with each successive pulse from the sensory neuron (provided that the facilitatory interneuron is silent). This decrease in the synaptic strength is considered to be the simplest form of learning (termed habituation) and is exactly opposite to what Hebb has proposed. The data does show, however, that the synaptic strength can be increased, but only when the facilitatory interneuron is active. Although the experimental data from *Aplysia* first became available in 1970's (well after Hebb published his book), it is quite remarkable that Hebb's rule is still widely used (albeit in slightly modified versions). One example is the following equation which also includes a passive decay term

$$\frac{dz_{ij}}{dt} = -F_{ij}z_{ij} + G_{ij}f_i(x_i)h_j(x_j) \tag{3.2}$$

where $f_i(x_i)h_j(x_j)$ is often called a Hebbian term. An alternate learning equation, called the gated decay long term memory (LTM) equation and proposed by Grossberg is

$$\frac{dz_{ij}}{dt} = h_j(x_j)\,[-F_{ij}z_{ij} + G_{ij}f_i(x_i)] \tag{3.3}$$

This gated decay LTM learning equation represents the form of learning that is implemented in all ART neural networks and also forms a foundation for the learning equations that we will use in subsequent chapters.

3.3.1 Model of Chemical Synapses

One of the earliest models of chemical synapses was by S. Grossberg (1968, 1969). This model is of extreme importance because it embeds three coupled variables that represent transmitter production, storage and mobilization, as well as modulated transmitter release by the synaptic input signal. A version of the model with transmitter mobilization is represented by the following two equations.

$$\frac{dZ_{ij}}{dt} = \lambda_{ij}^{+}(\delta_{ij} z_{ij} - Z_{ij}) - \lambda_{ij}^{-} F_{ij}\left[\tilde{Z}_{ij} - U_{ij}\right]^{+} \tag{3.4}$$

$$\frac{d\tilde{Z}_{ij}}{dt} = \omega^{+}\left(Z_{ij} - \tilde{Z}_{ij}\right) - \lambda_{ij}^{-} F_{ij}\left[\tilde{Z}_{ij} - U_{ij}\right]^{+} - \omega_{ij}^{-}\left[\tilde{Z}_{ij} - V_{ij}\right]^{+} \tag{3.5}$$

where $[w]^{+} = \max(w, 0)$ is a thresholding function; $Z_{ij}(t)$ is the total amount of the transmitter in the synaptic pathway at time t; $\tilde{Z}_{ij}(t)$ is the total amount of the *mobilized* transmitter at time t; $z_{ij}(t)$ is the total number of *active transmitter production sites* at time t (i.e., it represents the transmitter production rate); F_{ij} is the spiking frequency driving the synapse and the transmitter release process. By assuming a steady-state spiking frequency F and by ignoring the slow variations in $z_{ij}(t)$, Grossberg has shown that equations (3.4) and (3.5) can be solved explicitly for the transient response of the model. The solution for $\tilde{Z}_{ij}(t)$ is given by

$$\tilde{Z}_{ij}(t) = \left(\frac{w_{ij}^{+}\delta_{ij} z_{ij}(0)}{\omega_{ij}^{-} + \omega_{ij}^{+}}\right)\left(\exp(-(\lambda_{ij}^{+} + \lambda_{ij}^{-}F)t) + \frac{\lambda_{ij}^{+}}{\lambda_{ij}^{+} + \lambda_{ij}^{-}F}(1 - \exp(-(\lambda_{ij}^{+} + \lambda_{ij}^{-}F)t))\right) \tag{3.6}$$

which is a monotone decreasing function in F, with the steady state being given by

$$\tilde{Z}_{ij}(\infty) = \frac{w_{ij}^{+}\lambda_{ij}^{+}\delta_{ij} z_{ij}(0)}{(\omega_{ij}^{-} + \omega_{ij}^{+})(\lambda_{ij}^{+} + \lambda_{ij}^{-}F)} \tag{3.7}$$

It can be shown that the amount of the mobilized transmitter that is available for release per unit time (for the case of zero thresholds) jumps from a maximum of

$$\frac{\lambda_{ij}^{-}\omega_{ij}^{+}\delta_{ij}^{+} z_{ij}(0)F}{\omega_{ij}^{-} + \omega_{ij}^{+}} \tag{3.8}$$

at time zero and thereafter decays asymptotically to the value of

$$\frac{\lambda_{ij}^{-}\lambda_{ij}^{+}\omega_{ij}^{+}\delta_{ij}^{+}z_{ij}(0)F}{(\omega_{ij}^{-}+\omega_{ij}^{+})(\lambda_{ij}^{+}+\lambda_{ij}^{-}F)} \tag{3.9}$$

which is monotone increasing in F. Whereas the above model assumes that the rate of transmitter mobilization is independent of the spiking frequency, it can be shown that for the case where the mobilization rate does depend on the spiking frequency, the amount of the mobilized transmitter rises from zero and then decays to a positive asymptote. Thus, in either case, the model ends up in a stable state, which from an engineering point of view is crucial if the model is to be used in the design of complex neural circuits that depend on the model for signal transmission.

A slightly modified version of the above model, which forms the basis for the model used in this thesis, is represented by equations (3.10) and (3.11), from Carpenter and Grossberg (1981).

$$\frac{dw}{dt} = K(L-w) - (Mw - Nz) \tag{3.10}$$

$$\frac{dz}{dt} = (Mw - Nz) - Sz \tag{3.11}$$

where the term $K(L-w)$ in equation (3.10) says that $w(t)$ tries to maintain a level L via transmitter accumulation; the term $-(Mw-Nz)$ in (3.10) says that the storage transmitter is mobilized at a rate M and that the mobilized transmitter z is demobilized and restored at a rate N. The term $-Sz$ is (3.11) says that the mobilized transmitter is released (and depleted) by the input signal S.

A simplified version of the above models (which does not include the transmitter mobilization process) has recently been used in the implementation of the ART-3 neural network (reviewed in section 3.4.3), in modelling the dynamics of the fly visual system (Ögmen and Gagné, 1990) and in the implementation of our simple feedforward neural network model of visual spatial attention (Lozo *et al.*, 1993a, b) and translation invariant pattern recognition (Lozo *et al.*, 1994, 1995; see Chapter 6). Although the application of simple models of chemical synapses in a number of different artificial neural network architectures has provided these networks with new processing flexibility, there is no strong evidence that these networks have benefited much from the additional computational elements. For example, modelling chemical synapses in the ART-3 neural network replaces the need for a threshold upon the network reset. Similarly, in our application to a simple feedforward model of visual spatial attention (Chapter 6), the dynamics of chemical synapses has been chiefly used to implement the mechanism of attention shifting.

Self-regulated Transmitter Release

It has been suggested (Grossberg, 1984) that the reuptake of the released transmitter back into the synapse via autoreceptors has an inhibitory effect on further transmitter release, i.e., the reuptake process provides a self-regulated release of transmitter. Grossberg has also suggested that when lateral transmitter diffusion and reuptake occurs between several synapses, then the feedback inhibition from the total released transmitter provides a normalizing effect such that the maximum amount of transmitter that can be released is independent of the number of active synapses (assuming that all inputs are equal).

Equation (3.12) represents the inhibitory effect of the total released transmitter, for the case of a simple model of a synapse that does not include transmitter mobilization (Grossberg, 1988).

$$\frac{dz_{ij}}{dt} = \varepsilon S_i \left[-F z_{ij} + G x_j - H \sum_{k=1}^{n} S_k z_{kj} \right] \tag{3.12}$$

where z_{ij} is the transmitter level in the pathway from the ith to the jth cell; S_i is the signal emitted by the ith cell into this pathway; x_j is the activity of the jth cell; ε, F, G and H are constants. This equation says the reuptake via autoreceptors of a fraction of the released transmitter inhibits the growth of the corresponding LTM trace.

Presynaptic autoreceptors and the transmitter reuptake thereby mediate a type of self-regulated competition for the activation of the postsynaptic cells. This theoretical prediction (supported by the experimental results of Stricker and Zigmond, 1976) suggests that when some neurons are damaged and can no loner participate in transmitter release from their target synapses, then other neurons whose synapses are in close contact, compensate by producing more transmitter. We will make use of this concept to implement a normalization procedure in neural circuits whose neurons sample their input by more than one synapse.

3.4 Adaptive Resonance Theory and Implementation

ART was initially introduced by S. Grossberg (1976, 1980) as a physical theory of cognitive information processing in the human brain and has since been implemented in a family of artificial neural network architectures. These include: ART-1 for binary inputs (Carpenter & Grossberg, 1987a), ART-2 for binary and analog inputs (Carpenter & Grossberg, 1987b), ART-3 for hierarchical neural architectures (Carpenter & Grossberg, 1990), ART-MAP for supervised self-organisation of memory codes (Carpenter *et al.*, 1991), and various other versions.

ART arose from attempts to solve the *stability-plasticity dilemma* of real-time competitive learning systems. The dilemma is that real-time learning systems must be able to rapidly learn about significant novel events while at the same time being able to autonomously impose stability on their learning to prevent the erosion or corruption of past memories. The adaptive resonance concept suggests that only the resonant state in which the reverberation between feedforward (bottom-up) and feedback (top-down) computations within the system are consonant can lead to adaptive changes (Grossberg, 1980, 1982). ART models therefore embed bottom-up and top-down learning in a network that consists of two subsystem that regulate learning: (i) an attentional subsystem where top-down expectancies (recalled memories) interact with the bottom-up information; and (ii) an orienting (or a vigilance) subsystem that is sensitive to the mismatch between the two. Interactions between these two subsystems ensures that memory modification occurs under exceptional circumstances, i.e., memory can be modified when an approximate match has occurred. This state is called an *adaptive resonance*.

3.4.1 Strength of ART

Because the ART model departs from conventional neural network models by also introducing the top-down learned pathways and a mismatch arousal subsystem, the model has been able to predict and explain the existence and the nature of Event Related Potentials (ERP's), Banquet & Grossberg (1987). The strong correspondence between the neural mechanisms embedded in the ART model and that of cognitive information processing has enabled the model to be used as a physical and explanatory theory of various brain functions. For example, the theory has been used to explain the data on visual perception, speech perception, neural substrates of learning and memory, amnesias and a host of other cognitive data. It is quite remarkable that one theory alone, which conceptually is quite simple, can be used to explain such a diverse set of biological cognitive data. Below we qualitatively analyse the behaviour of the model during learning in simple and uncluttered inputs.

Let us assume that in the context of primate vision, ART represents the highest level of visual information processing, say at the level of the inferior temporal cortex (Carpenter & Grossberg, 1993). The theory then suggests how our visual memory (internal representation of visual stimuli) can be stored and evolved through further experience with the same (or almost the same) sensation, without eroding the internal representation of this memory when it is recalled or activated by another sensation that does not match it. Figure 3.1 shows how various ART neural mechanisms interact to ensure stable and real-time evolution of internal representations of external sensory stimuli.

To qualitatively analyse the behaviour of the ART model in the context of high level vision, let us consider how the long term memory traces are evolved through exposure to sensory stimulation.

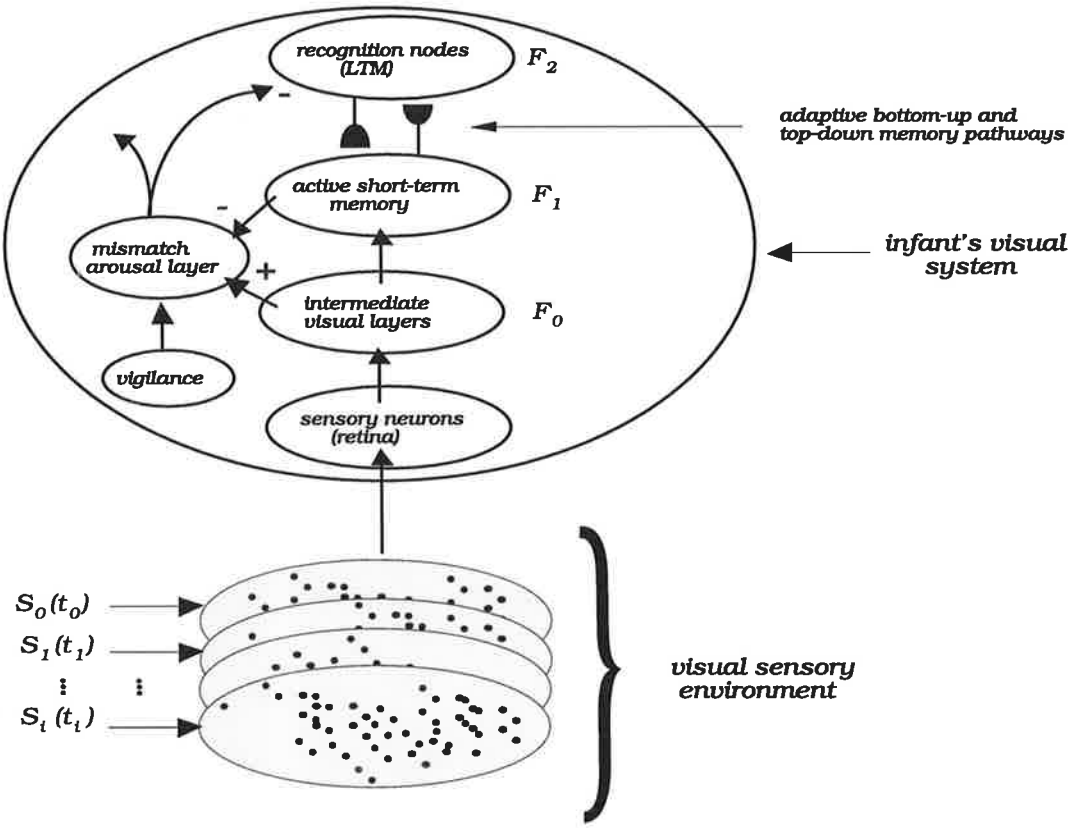


FIGURE 3.1. Very simplified representation of the visual brain sketched primarily to illustrate the main ART concepts (we are presently ignoring the numerous intermediate layers of the visual cortex).

The memory of the very first visual stimulation can be stored into LTM pathways since no prior top-down learned memory trace will be recalled to cause a mismatch between the input and the resultant pattern of activity across layer F1. Once the memory trace of this stimulus is stored (the strength of which depends on the learning rate that may itself depend on a host of other factors, such as the degree of arousal, etc.) it can then be evolved into a stable memory by further experience. For example, if at some later time this stored memory trace is recalled by another visual stimulus that does not differ very much from the recalled memory, then the memory trace is increased further. Once the top-down memory trace of this visual experience is strong enough to override the bottom-up activations, it can no longer be corrupted by sufficiently different visual stimuli that share some common features. This causes a mismatch arousal and a memory search for a better matching memory. If a memory search does get engaged by the new input stimulus but cannot activate or recall another stored memory, then

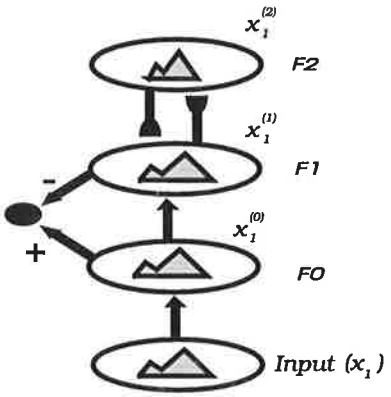
this sensory stimulation gets stored as a new visual memory. This example is a simplified treatment since the theory also suggests that when the memory traces are weak, they can be overwritten by non-matching inputs but should eventually stabilise through self-organisation.

The strength of Adaptive Resonance Theory therefore lies in the proposal of the top-down learned pathways whose signals mix with the bottom-up inputs to form an attentional subsystem, which by its interaction with the orienting (or vigilance) subsystem ensures that the memory modification of stored (or new) representations occurs under exceptional circumstances, i.e., learning is gated. In the next subsection we present a simple example to support our claim that ART's attentional subsystem is incomplete and is too rigid to be of much use in more realistic visual scenes.

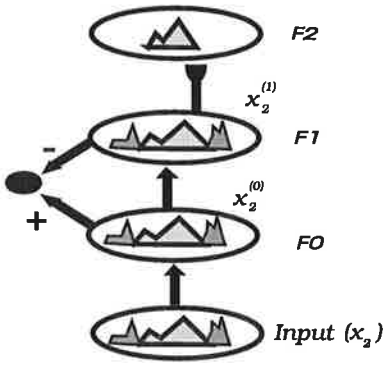
3.4.2 Weakness of ART's Attentional Subsystem

Although it has been claimed that ART models can self-stabilize learned codes in an "arbitrary input environment" (Grossberg, 1987) or "arbitrary complex input environments" (Grossberg 1988), we will show that this is not so. In fact, it can be very easily demonstrated that none of the ART neural networks developed to date can recognise a familiar input when it is embedded in cluttered and complex backgrounds. It should be very obvious that if a learning system cannot recognize a learned input when it is embedded in a cluttered background then the system cannot learn in those types of backgrounds. Simple computer simulations in section 3.4.3 of ART-3's neural layers (processing Fields) will clearly demonstrate this. The problem that we are exposing below, however, is not a peculiarity of any of the ART based networks but is inherent in all because of the shortcoming of the theory itself. ART does not propose how such problems may be dealt with by a real-time neural system. Hence it is first necessary to provide a theoretical argument, which actually happens to be very simple, to support our claim that ART model lacks the generality needed to be used as a physical theory of biological cognitive information processing in complex, cluttered and arbitrary input environments, primarily because of its rigid attentional subsystem.

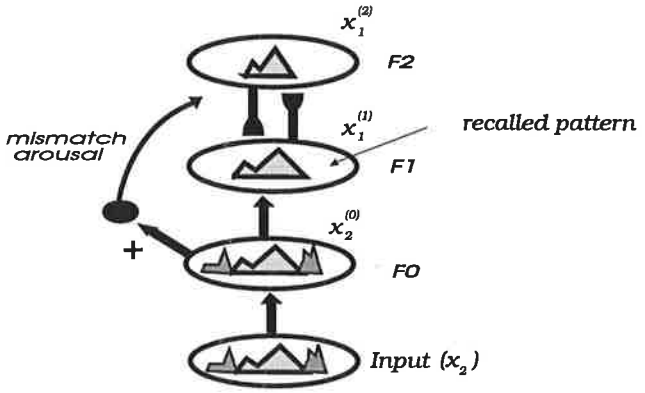
The exposed problem, schematized in Figure 3.2, can best be understood if one ignores the signal amplitudes that are fed into the system and simply considers the input to be a pattern in 2-D space that excites neurons in a 2-D lattice, such that each activated neuron corresponds to a location of the space that is occupied by the input pattern.



(a) learn pattern 1



(b) pattern 1 embedded in clutter



(c) failure to recognize pattern 1 when embedded in cluttered background

FIGURE 3.2. Revised thought experiment that reveals the inadequacy of ART's attentional subsystem: (a) shows the case during the learning of an input pattern (pattern 1); (b) and (c) show the processing stages of the ART model when the learned pattern is embedded in a cluttered background. Note that ART's attentional subsystem is unable to selectively attend to and recognize a familiar pattern (input x_1) when it is embedded in the cluttered background of input pattern x_2. The network fails to recognize the familiar stimulus because the attentional subsystem is not capable of paying attention to those parts of the input that can be matched by the top-down memory.

We are now led to ask whether the attentional subsystem, as proposed in ART, is sufficient to explain how we can recognise a previously learned visual stimulus when that stimulus is embedded in a complex background. Let us for the sake of the argument assume that a person is looking at the familiar visual stimulus and that this stimulus is embedded in a cluttered background. Let as also assume that the early visual layers of the person's brain are not able to separate (segment) the familiar stimulus from its background. Since the stimulus is at the centre of gaze and has not changed in size or orientation, there is no reason to suppose why its features will not be aligned with the corresponding memory pathways that have on previous occasion encoded these features. Therefore we should expect that the recognition nodes of the familiar stimulus will be

activated and its memory will be recalled and transferred via the top-down memory pathways into the short-term memory at Field F1. Figure 3.2 illustrates this simple scenario.

As shown in Figure 3.2(c), strong top-down memory traces are able to completely override any activity in F1. Any bottom-up initiated activity in F1 that does not match the recalled top-down activity will be annihilated by competitive interactions in F1. However, this has no effect on the activity in F0. Therefore if F0 is simultaneously activated by several external stimuli, as shown, its activity will not match that of F1. Thus, although the correct memory may be initially activated and recalled, ART suggests that because of the resultant mismatch between patterns in F1 and F0, the recalled memory should be reset and the memory search engaged. Hence the familiar object cannot be recognised. Accordingly, we will never be able to recognise a familiar object when not presented in isolation. This obviously is not the case for we are able segment and recognise a familiar object in a multitude of complex environments.

We note that ART suggests that in such case one of the following can occur: (i) once the recalled memory is reset and temporarily biased against further activation, the input stimulus can reactivate the system and eventually get stored in memory as a new stimulus; or (ii) the vigilance may be reduced to accommodate higher degrees of mismatches. However, we argue that neither of these is appropriate. In the former case the familiar stimulus is not recognised, while in the latter case the memory can never stabilize since it will be continuously modified whenever it is activated by non-matching inputs.

We conclude that the weakness of the ART model is its rigid attentional system. This rigidity is due to the fact that the system always pays attention to the whole neural activity across its input layer rather than to parts of the input and, as a result, does not allow resonant activity to occur with a portion of the input that may in fact be very familiar. That is, the theory does not suggest how a neural system can selectively attend to and recognize a familiar stimulus in complex sensory stimulation. Thus, the same neural mechanism that has empowered the ART model with stable self-organised learning in randomly presented but simple input environments now prevents it from recognising a previously learned stimulus in complex and cluttered environments.

ART model therefore does not have sufficient processing capability and flexibility to offer a thorough rationalization of cognitive information processing, but it does provide a strong theoretical foundation for a more general theory. Since ART is the only theory developed to date that has the necessary ingredients that correlate with the psychologically and neuropsychologically observable phenomena of biological cognition, it is imperative that we seek a more refined solution to real-time learning. This solution should simultaneously cater for the mechanism of stable evolution of memories as well

as selective attention. Before we proceed to propose a more general solution in subsequent chapters, it is worthwhile to consider the original thought experiment around which the theory was developed.

Grossberg's thought experiment (Grossberg, 1980) considered an array of cells in two neural Fields (F1 and F2). However, because the input was not treated as a 2-D spatial pattern of activity, the arguments that led to the proposal of ART's attentional mechanism were developed in a one-dimensional domain. Consequently, some problems that are obvious in the two dimensional spatial vision have not been considered and have therefore led to a restricted neural theory. For example, in 2-D visual pattern recognition problems, the amplitude of signals is not as important in defining an object as are the spatial locations of the activated neurons. Although ART was not aimed to be a theory of vision but a theory of self-organised learning regardless of the sensory modality, it may be argued that our thought experiment is only applicable to vision. However, since selective attention is not unique to vision but permeates every other sensory modality, our claim about ART's deficiency is independent of the sensory modality.

Thus, although ART models are applicable to simple and uncluttered sensory environments, they are inadequate in complex, cluttered, noisy and arbitrary environments. This is well exemplified in a recent application of the ART-2 neural network where it is embedded in a vision hierarchy for adaptive 3-D object recognition by a mobile robot (Seibert and Waxman, 1992). The problem faced by such a system is that it only succeeds if the objects to be recognized are presented in a manner that allows easy and clean segmentation by feedforward computations. Objects were either painted matte black to ensure significant contrast with the background or were represented by a suitable arrangement of lit light globes. The architecture used by these authors is strictly feedforward, although some parts of the system, such as the ART-2 network, do have internal feedback. Since the internal feedback in the ART model provides it with one of the most powerful cognitive neural concepts i.e., attention, why should not that feedback be also used to enhance processing at lower layers?

One of ART's predictions is that whenever a novel stimulus appears then the arousal of the system should increase. However, the theory does not propose how a self-organising neural learning system knows whether the input is familiar or novel. It is assumed that whenever a top-down expectancy (or recalled memory) mismatches the bottom-up input that an orienting subsystem triggers a non-specific arousal mechanism, thus resetting the activated top-down memory. But does memory mismatch and reset always imply novelty? Is increased arousal and memory reset always needed for familiar inputs? In all the current neural network implementations of the theory (e.g., in ART based networks such as ART-2 or ART-3), a familiar input can cause a

mismatch reset when it is presented after another input. Should a familiar input be able to activate its memory without a prior memory reset by having a direct route to the memory (i.e., direct memory access or DMA), which by-passes some neural layers?

3.4.3 ART-3 Neural Network

We now review one of the implementations of ART called ART-3 neural network (Carpenter and Grossberg, 1990) and present several computer simulations that demonstrate the weakness of ART in cluttered background. ART-3, schematized in Figure 3.3, is the first ART based neural network whose adaptive bottom-up and top-down memory pathways are based on a simple model of chemical synapses. The primary purpose of using a model of chemical synapses in the memory pathways was to implement a mechanism for parallel memory search of learned pattern recognition codes, thus enabling ART-3 to be embedded in network hierarchies.

Although ART-3's learning capability has not been demonstrated nor were its learning equations specified in the published article of Carpenter and Grossberg, our investigations of ART-3 have indicated that its learning does not exceed that of its predecessors (such as ART-2, Carpenter and Grossberg, 1987b) on 1-D analog spatial patterns. We have also found (Lozo, 1993) that ART-3 neural network can end up in a never ending search cycle (where it continuously activates the same set of LTM nodes, followed by a mismatch reset) and that this memory search is in fact serial. That is, the network is likely to activate a number of non-relevant (partly or poorly matching) memory nodes before it activates and settles down into a resonant state with the best matching node. The serial nature of memory activation and search is also evident in the published computer simulations (see Figures 23-25, Carpenter and Grossberg, 1990). Nevertheless, ART-3 is of particular interest because it demonstrates that it is feasible to design artificial neural network architecture whose signals flow through dynamic pathways that have a close correlate with the neurobiological form of signal transmission.

Each of ART-3's processing Fields are identical with the exception that Field F_c is modelled by a winner-take-all competition (although not strictly necessary in ART-3). In Figure 3.4 we reproduce one of the processing Fields, whose implementation approximates the competitive feedback equation (3.1). Large black circles represent signal normalization across each layer (i.e., approximation of lateral competition). Reverberatory loops in each of the Fields maintains a short-term memory (STM) of their input. Within each Field, the gain in the top-down directions is typically chosen to be at least ten times the gain of the bottom-up signals. The purpose of the larger

top-down gain is to ensure that when a given spatial input pattern establishes a reverberatory activity within each Field, then this reverberatory activity maintains a contrast enhanced memory of the input after the input offset.

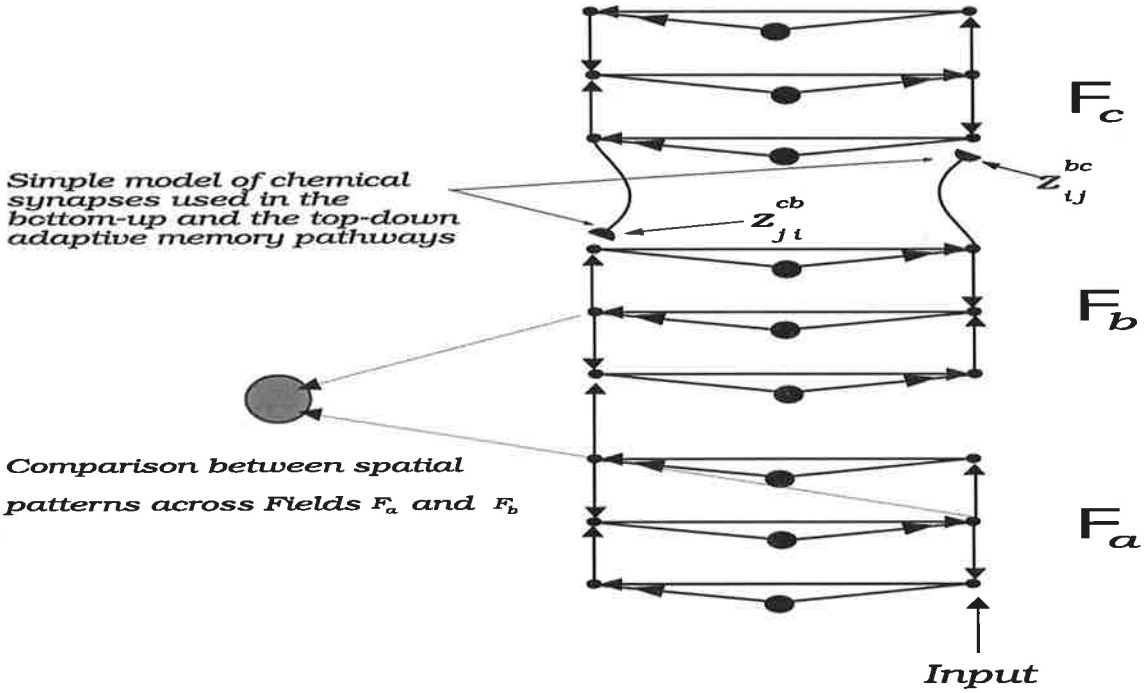


FIGURE 3.3. Implementation of ART-3, with simple model of chemical synapses in the adaptive memory pathways (from Carpenter & Grossberg, 1990).

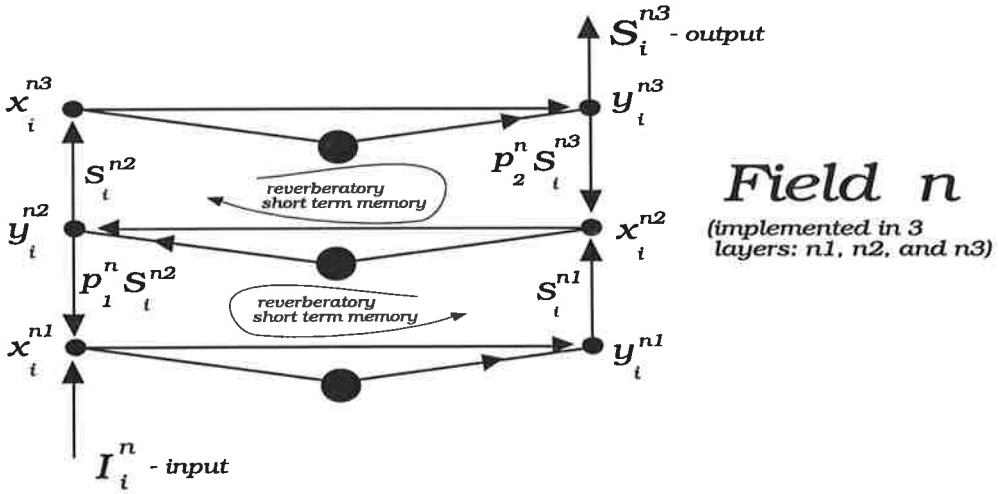


FIGURE 3.4. Implementation of shunting competitive neural Fields in ART-2 and ART-3 neural networks (Carpenter & Grossberg, 1987b, 1990).

These reverberatory Fields have the interesting property that when activated, their activity prevents subsequent inputs to enter (unless the reverberations are reset). That is, because of the higher top-down gain, the established activity prevents subsequent inputs to enter and corrupt the integrity of the spatial pattern that the Field is 'remembering' via the reverberatory loops. The system is reset when the match between Fields F_a and F_b (measured by an indirect evaluation of the cosine of the angle between the two multidimensional vectors that are represented by the spatial activity across the two Fields) falls below a preset level (called vigilance).

The following six equations represent the implementation of the shown Field.

$$x_i^{n1} = I_i^n + p_1^n S_i^{n2} \tag{3.13}$$

$$y_i^{n1} = \frac{x_i^{n1}}{p_3^n + \| X^{n1} \|} \tag{3.14}$$

$$x_i^{n2} = S_i^{n1} + p_2^n S_i^{n3} \tag{3.15}$$

$$y_i^{n2} = \frac{x_i^{n2}}{p_3^n + \| X^{n2} \|} \tag{3.16}$$

$$x_i^{n3} = S_i^{n2} \tag{3.17}$$

$$y_i^{n3} = \frac{x_i^{n3}}{p_3^n + \| X^{n3} \|} \tag{3.18}$$

where $S_i^{nk} = G[y_i^{nk} - \Theta]^+ = G \max(y_i^{nk} - \Theta, 0)$; G is the gain at the output of each layer; Θ is the threshold on the output stage of each layer and is chosen so that a uniform spatial pattern across the whole layer is quenched below the threshold, i.e., $\Theta = 1/\sqrt{(n)}$, where n is the number of neurons in the layer; $\| X^{nk} \|$ is the L_2 norm; p_3^n (= 0.0001) is a computational convenience for zero inputs.

Model of Chemical Synapses in ART-3's Bottom-up Memory Pathways

Although both the bottom-up and the top-down memory pathways are modelled by adaptive transmitter dynamics, here we will consider only the bottom-up pathways. Equation (3.19) represents the dynamics of the bottom-up synapses between Fields F_b and F_c.

$$\frac{du_{ij}^{bc}}{dt} = (z_{ij}^{bc} - u_{ij}^{bc}) - p_1^c S_i^{bc} u_{ij}^{bc} (p_2^c + x_j^{c1}) \tag{3.19}$$

where z_{ij}^{bc} represents the bottom-up transmitter production level (learned long term memory); u_{ij}^{bc} represents the stored transmitter (medium term memory); S_i^{bc} is the bottom-up input from cell i of Field F_b to cell j of Field F_c; p_1^c and p_2^c are constants. This equation says that the medium term memory (MTM) tracks the long term memory (LTM) and that the stored transmitter (MTM) is reduced by the action of the input signal S_i^{bc} and by the correlated firing of the input signal and the postsynaptic feedback signal x_j^c from the active cell in Field F_c. The excitatory postsynaptic voltage, v_{ij}^{bc}, is given by

$$\frac{dv_{ij}^{bc}}{dt} = -v_{ij}^{bc} + p_5^c S_i^{b3} u_{ij}^{bc} (x_j^{c1} + p_6^c) \tag{3.20}$$

where p_5^c and p_6^c are constants. The pre-postsynaptic interactions in the bottom-up memory pathways are schematized in Figure 3.5.

As shown in Figure 3.5, the postsynaptic signal x_j^{c1} interacts multiplicativelly with the input signal S_i^{b3} to provide a further boost in the EPSP. As mentioned above, although Carpenter and Grossberg (1990) did not provide learning equations for ART-3, it can be assumed that the transmitter production rates in the bottom-up and the top-down pathways are adaptive. As in the previous ART neural networks, the bottom-up pathways learn their driving signal (i.e., the presynaptic learning rule), while the top-down pathways learn the signals on their postsynaptic target cells (i.e., the postsynaptic learning rule).

The learning equation for the bottom-up memory pathways may be written as follows

$$\frac{dz_{ij}^{bc}}{dt} = R\delta(S_i^{b3} - z_{ij}^{bc}) \tag{3.21}$$

where δ is the learning rate, while R is a gating signal defined by

$$R = \begin{Bmatrix} 1 & if & M \geq \rho & \text{AND} & Steady State \\ & & 0 & & otherwise \end{Bmatrix} \tag{3.22}$$

In the above, M is the degree of match between the relevant neural layers while ρ is a dimensionless parameter (vigilance) that controls the degree of the approximate match that needs to be satisfied before learning is enabled. It is thus necessary to track the system dynamics, so that the memory update is enabled whenever a steady state is reached and, when the match between the relevant Fields is above the preset vigilance level of ρ.

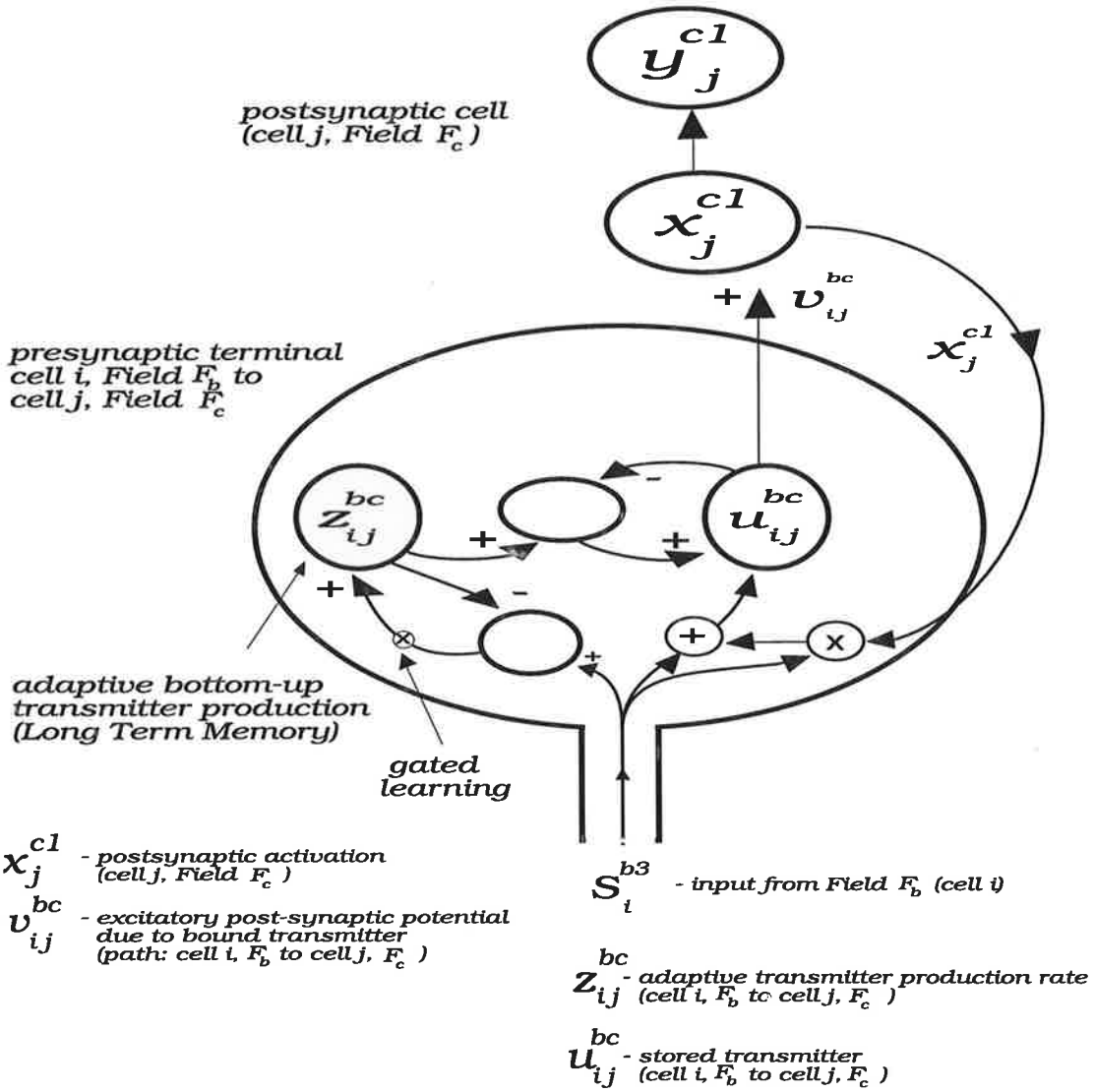


FIGURE 3.5. Schematic of a simple model of chemical synapses in ART-3's bottom-up memory pathways. Note that the transmitter production rate (z_{ij}^{bc}) adapts to the level of the presynaptic signal S_i^{b3}.

In addition to the problems of serial and cyclic memory search cycles as discussed above (which are reasonably trivial), ART-3 neural network (as well as its predecessors, such as ART-2) has a far more serious problem that limits its application to simple problems. As claimed in section 3.3.2, none of the ART based neural networks that have been developed to date can recognize a familiar pattern when it is embedded in a cluttered background. To reveal this problem it is not necessary to provide a computer simulation of the whole network, but its individual processing Fields. Below we provide six computer simulations that clearly reveal the inadequacy of ART neural networks in more complex inputs.

ART-3's Processing Field: Simulation I

The first simulation uses 2-D shapes shown in Figure 3.6. These shapes are pre-aligned in the input to maximize their spatial overlap. Input 1 is presented for two iterations of the equations (3.13)-(3.18), followed by Input 2 for further 15 iterations. Note that the Field is not reset upon the presentation of the second input. The purpose of the simulation is to show that an established reverberatory activity in the Field prevents subsequent inputs to enter into the Field. To expose the property of the Field, we will show the state of each variable in the Field in the form of the resultant steady state activity. The Figure below shows the steady state activity of the Field after the presentation of the inputs.

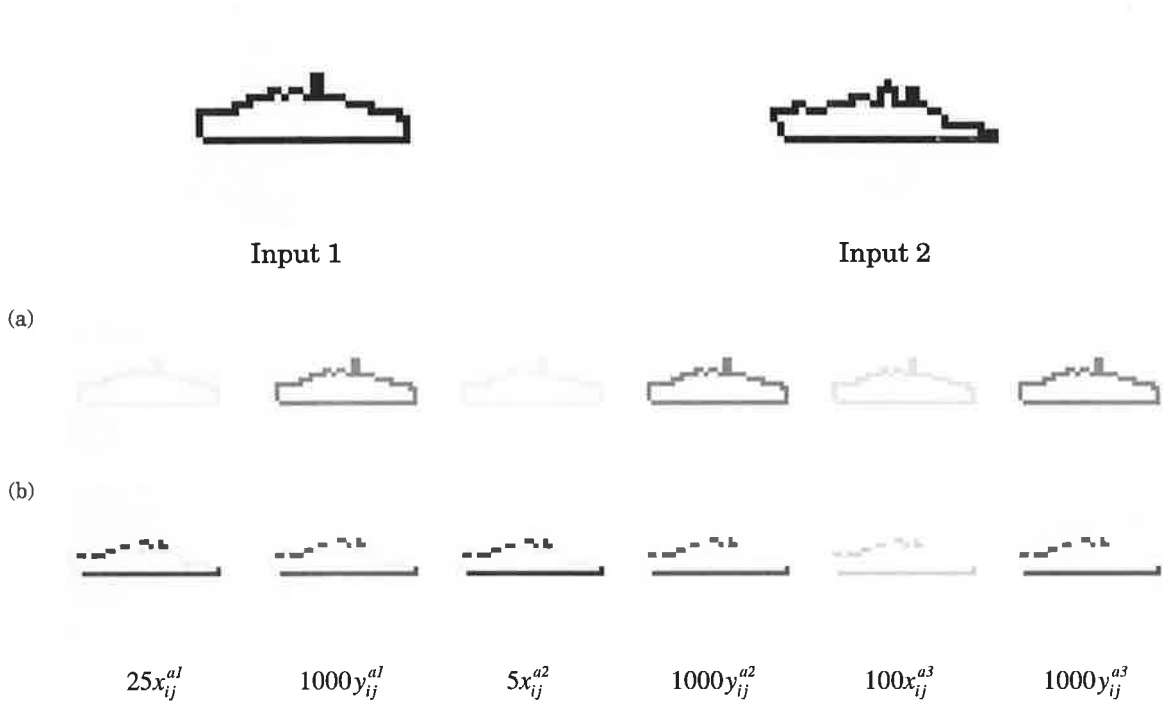


FIGURE 3.6. ART-3's processing Fields: Simulation 1. (a) steady state due to input 1; (b) steady state due to input 2 (without prior reset). Field parameters: $p_1^n = p_2^n = 10.0$; $G = 5.0$; $\Theta = 0.03125$, number of neurons in each layer = 1024; $p_3^n = 0.0001$. Convention: dark = 1; white = 0;

The simulation in Figure 3.6 shows that when a reverberating spatial pattern of activity is established across one of ART-3's processing Fields, then this activity has strong suppressive effects on subsequent inputs that do not match the reverberating pattern. Hence, unless all activities within the Field are reset to zero, a new input cannot enter into the processing Field. Since the established reverberation continues after the input offset and for as long as there is no active reset, it can be said that the established reverberation represents a memory of the spatial pattern that has caused it (the strength of which depends on the gain in the top-down direction). It should also

be noted that this reverberatory memory acts as a filter. That is, when a new input shares some common pathways with the reverberatory activity within the Field, then these common pathways will be amplified while the competition within the Field will prevent all other inputs from entering into the Field.

As can be seen in the above figure, the activity of the established reverberatory pattern will increase in some places and decrease in others. The reason for this is that the initially activated cells that no longer receive a bottom-up input will be suppressed by all the other cells that are still receiving a bottom-up input.

ART-3's Processing Field: Simulation II

The filtering action of the reverberatory pattern is further exemplified below where the first 2-D shape is embedded in the cluttered background of the second input, as shown in Figure 3.7.

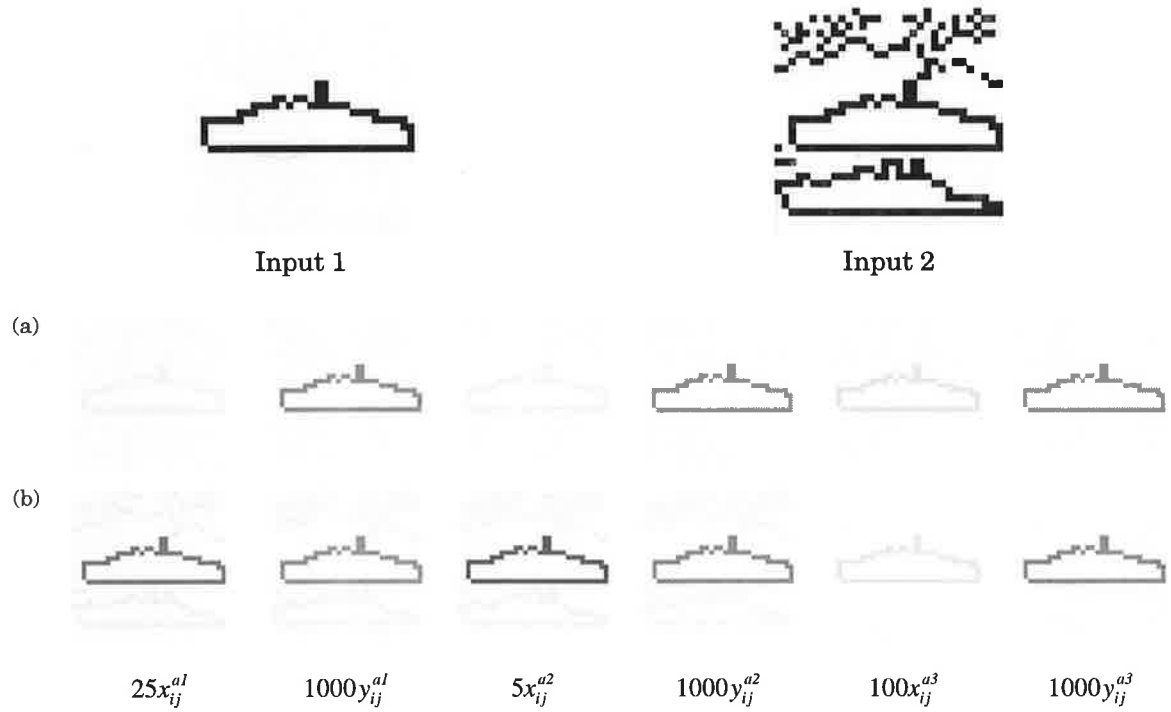


FIGURE 3.7. ART-3's processing fields: Simulation II. Field parameters as for Figure 3.6.

ART-3's Processing Field: Simulation III

In the next simulation, the first shape also appears in the cluttered background, but some of its boundary elements are weaker in strength than the background elements. The purpose of this simulation is to demonstrate that while the non-relevant inputs cannot get into the Field, the competition within the Field also annihilates weaker inputs of the relevant pattern.

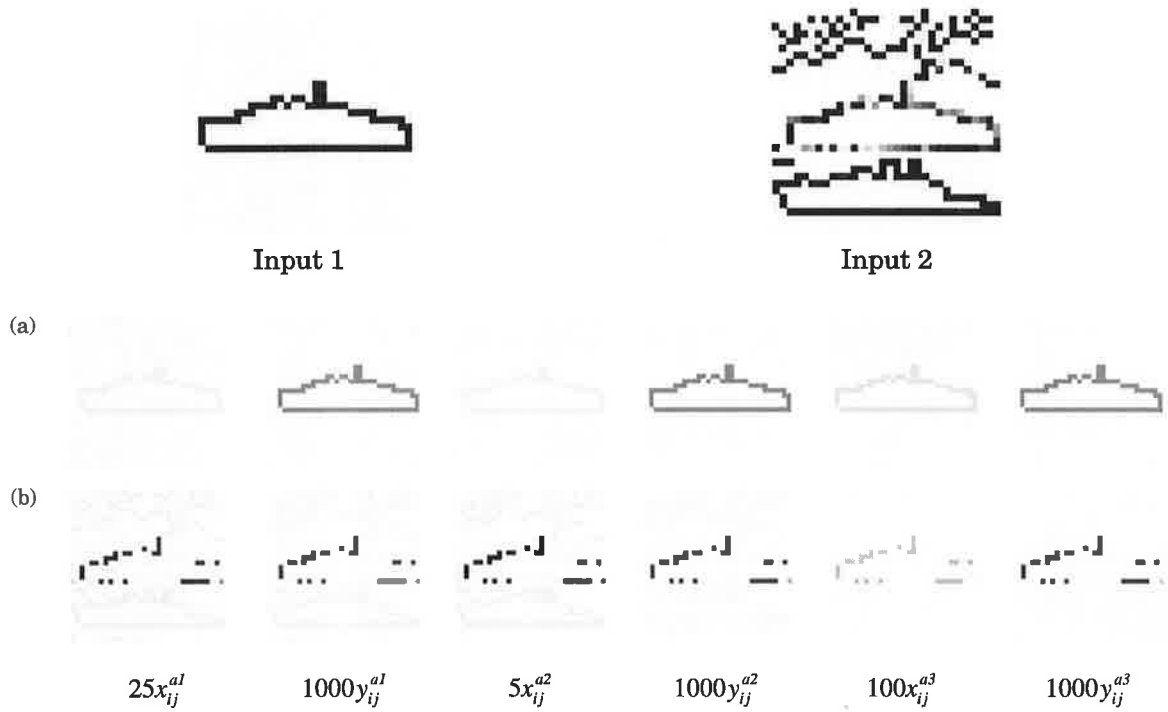


FIGURE 3.8. ART-3's processing fields: Simulation III. Field parameters as for Figure 3.6.

ART-3's Processing Field: Simulation IV and V

The next two simulations show the resultant steady state when the input is cluttered with several shapes.

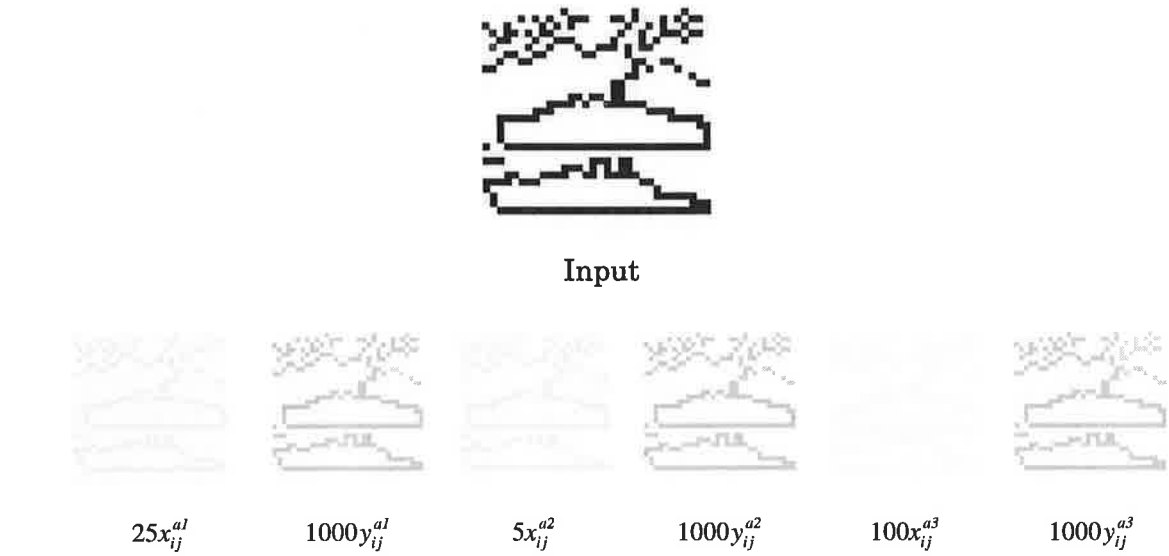


FIGURE 3.9. ART-3's processing fields: Simulation IV. Field parameters as for Figure 3.6.

Variable input strength along object's boundary

$25x_{ij}^{a1}$ $1000y_{ij}^{a1}$ $5x_{ij}^{a2}$ $1000y_{ij}^{a2}$ $100x_{ij}^{a3}$ $1000y_{ij}^{a3}$

FIGURE 3.10. ART-3's processing fields: Simulation V. Field parameters as for Figure 3.6.

In Simulation IV all inputs have equal intensity whereas in Simulation V some parts of one of the shapes varies in strength along its boundary. The purpose of these two simulations is to show that when no prior activity is reverberating then a new input pattern, regardless of its complexity, will register itself, but only in locations where the input can drive its target cell above the threshold.

ART-3's Processing Field: Simulation VI

Below we present computer simulations of ART-3's Field F_b when it is driven by the bottom-up and the top-down inputs, as shown in Figure 3.11. The top-down input is assumed to represent the recalled memory of a previously learned 2-D shape.

Instead of modelling the top-down synapse into Field F_b, for simplicity we use equation (3.23) to represent the activity of node in Field F_b where J_i^{b3} is the top-down input into the node. The larger gain for the top-down input ensures that the final steady state of the Field will be due to the top-down input. This type of suppression of all initial activities in F_b and its replacement by the recalled memory generally occurs in all ART based neural networks when the top-down memory is reasonably large (or has converged).

$$x_i^{b3} = S_i^{b2} + 10J_i^{b3} \tag{3.23}$$

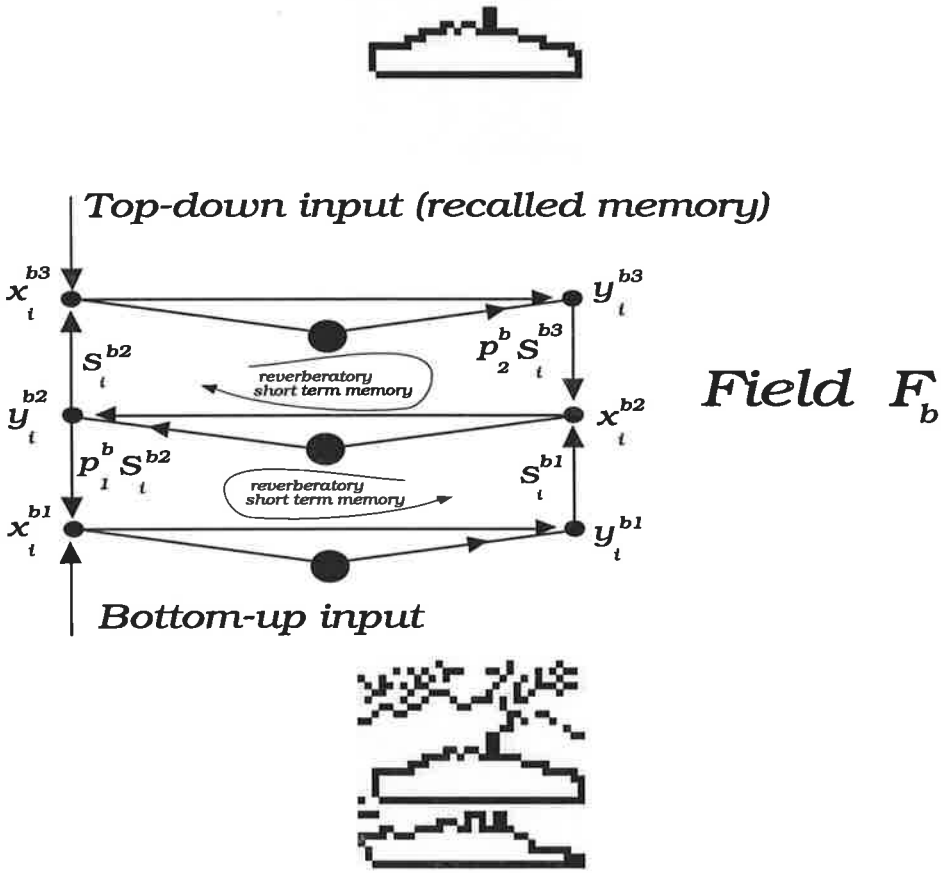


FIGURE 3.11.Mixing of the bottom-up input and a recalled top-down memory in ART-3's Field F_b.

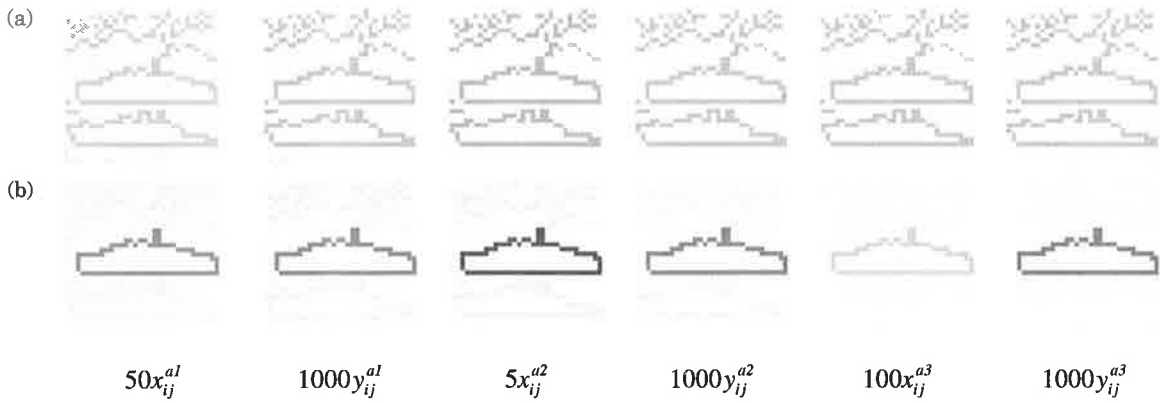


FIGURE 3.12. ART-3's processing Fields: Simulation VI. (a) steady state before the activation of the top-down input; (b) steady state after the introduction of the strong-top down input. Field parameters as for Figure 3.6.

As can be seen in the above simulation data, the top-down input has suppressed the activity of all other cells. Thus, in a steady state, Field F_b will contain an active memory of the recalled shape. However, as has been seen demonstrated in the previous simulations, the input Field F_a will in addition contain other non-relevant shapes and clutter. This will cause the match between Fields F_a and F_b to be low and below the vigilance level that is typically used to discriminate between similar looking shapes. It can thus be concluded that ART-3 (as well as ART-2) neural network would not in general recognize a previously learned 2-D shape when that shape is embedded in a cluttered background or when its boundary edges are non-uniform in intensity.

3.5 The Concept of Selective Attention Adaptive Resonance Theory (SAART)

In this section we return to the thought experiment of Figure 3.2 and derive the basic theoretical concept of Selective Attention Adaptive Resonance Theory (SAART), which is then fully developed in the subsequent chapters of the thesis.

As in section 3.4.2, we assume that a memory of a given sensory stimulus is stored in the bottom-up and the top-down adaptive memory pathways. We now embed the stimulus in a cluttered background, such that it remains un-obscured. If we can somehow filter out the non-relevant information (the background stimulation) from the input Field F_0 so as to cause a match between F_0 and the recalled memory at the short term memory (STM) Field F_1, then we will enable the system to recognize the familiar stimulus. The problem then is to figure out how to filter out the irrelevant inputs in order to cause a match between the recalled memory and a portion of the input pattern. But how can the system know which bits of the bottom-up input are irrelevant? Rather than attempting to answer this problem first, let us see whether the system has the means of knowing which bits of the input are relevant. But in order for it to have any knowledge of what is relevant it must have an established memory. Hence if a particular memory is activated and recalled, then the system can use it to attend to the known bits of the input. Thus it seems intuitive to suggest that the activated memory across the STM Field F_1 should be used to enhance the corresponding bottom-up input pathways into the input Field F_0. Lateral competition in the input Field F_0 will then suppress the activity of all neurons whose inputs are not enhanced by the corresponding neurons in the STM Field F_1. Figure 3.13 shows our proposed solution to how ART's attentional subsystem may be extended to enable a self-organising neural network to pay attention, and hence to recognize, familiar inputs in the cluttered bottom-up neural activity.

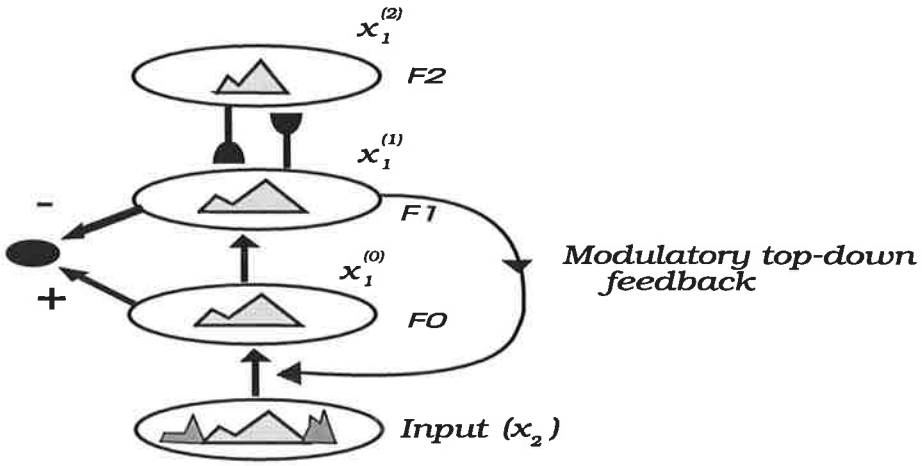


FIGURE 3.13. Selective Attention Adaptive Resonance Theory (SAART) concept. The top-down modulatory feedback from recalled memory amplifies the corresponding bottom-up inputs into the input Field F_0. Lateral competition in the input Field F_0 will annihilate the activity of all cells whose bottom-up signal transmission gain is not amplified by the active top-down memory pathways. These feedforward-feedback interactions enable resonance to occur between the recalled memory and a selected portion of the input.

Thus rather than resetting the system as soon as the match falls below the required threshold (as in the ART models) which would lead to recognition failure, we can alter the course of processing by providing a top-down facilitatory feedback to the input Field (F_0), enabling the network to test whether the recalled memory can be matched with a portion of the input. Clearly this feedback must act on the presynaptic pathways of the input Field F_0 in order to extract the familiar stimulus while suppressing all the other inputs. The mechanisms by which the relevant neural activity across the input Field F_0 can be enhanced at the expense of all the other activity is referred to a *top-down selective attention*. Top-down selective attention in the above example can be achieved by projecting the output activity of the STM Field F_1 to act on the neurotransmitter dynamics at the presynaptic terminal of the input Field F_0. This feedback pathway acts to selectively facilitate bottom-up signal transmission by enhancing the transmitter levels of corresponding bottom-up pathways. This simple but a very crucial step marks our departure from the ART model.

Because the facilitatory signals and the competitive interaction in F_0 do not act instantaneously but take some time, the resonant steady state develops over a period of time during which the network may be in a highly dynamic state. To follow the progress of these interactions, it becomes necessary to measure the degree of match between the spatial patterns across the Fields F_0 and F_1 as well as the time rate of change of this match. To protect the long term memory from unwarranted modification by non-matching inputs during these rapid changes, long term memory is updated when

the system is in a stable resonant state. Similarly, the certainty of the network's pattern recognition response increases as the steady state is approached. Thus, what may initially be taken as a bad mismatch may eventually end up as a perfect match but with a selected portion of the input pattern. Refer to Appendix C to see how ART-3 (as well as ART-2 and various other ART based networks) may be modified to model top-down memory guided selective attention.

The above concept of memory guided selective attention forms the main theoretical foundation of the thesis and is further developed in the subsequent chapters.

3.6 Conclusions

This chapter has provided a review of major theoretical concepts and mathematical tools upon which the thesis is based. A qualitative analysis of ART and the related computer simulations of ART-3's neural Fields have demonstrated that ART based neural networks have a serious limitation when the input is cluttered. We have identified this limitation to be primarily due to the fact that the attentional subsystem as proposed in ART is too rigid. The rigidity of ART's attentional subsystem does not allow resonant activity to occur between the recalled memory and a portion of the input with which it can match. This leads to the recognition failure of the familiar (previously learned) object when it is embedded in a cluttered background.

It is interesting to note that some of the most recent views on vision, which conveniently appear as three chapters in the same book, (Churchland *et al.,* 1994; Mumford, 1994; Ullman, 1994) have one thing in common: they all suggest an important role for the feedback pathways in the brain. Given that ART as a theory is already two decades old, it is quite remarkable that its concepts do not feature much in the work of others. Although Mumford has mentioned the ART neural networks of Carpenter and Grossberg (1987a, b), this was only to emphasise a matching process between a synthesized pattern and the input pattern. Even more surprising is that none of the above authors have mentioned the Boundary Contour System, even though it embeds feedforward-feedback interactions that the others seem to consider as being of great importance. If one were to predict where all these three views are likely to lead, one cannot escape noticing the similarity between their main propositions and ART (as well as the BCS and FCS neural network models). Mumford (1994), for example, has proposed the need for top-down pathways whose role would be to duplicate (or synthesize the stimulus). Isn't that exactly what the top-down memory as proposed in ART is all about? It would be interesting to see, as time goes by, whether an ever-increasing number of researchers converge onto the concepts that are already embedded in ART. Since ART addresses the most fundamental problem that is faced by real-time learning systems, i.e., stable evolution of memories, one can safely assume that it can be used

as a bench-mark against which future theories and models of the brain are to be compared. Thus rather than rediscovering ART, why not accept it and build onto it? This thesis attempts to build onto the power of ART, but from the perspective of a designer of a neurocomputational machine vision and object recognition system in cluttered visual scenes.

In section 3.5 we have shown how a simple extension to Grossberg's thought experiment has led us to derive the basic theoretical concept behind our Selective Attention Adaptive Resonance Theory (SAART). We have thus theoretically predicted the existence of axo-axonal interactions (where one axonal terminal modulates the synaptic terminal of another axon). We also predict the existence of modulatory feedback pathways in the primate visual system, whose role is to selectively facilitate the transfer of neuronal signals from lower to higher cortical layers. In the following chapters we use the concepts of selective attention and modulation to propose novel neural layers, mechanisms, design principles, neural circuits and networks for complex and cluttered visual scenes.

Chapter 4

Novel Neural Layers, Mechanisms and Design Principles

"The visual system is impressively efficient at extracting useful information, discarding irrelevant information, and being minimally disrupted by spurious data contained in the barrage of images that it must continuously process. This efficiency did not arise instantaneously during evolution, and to incorporate it into neural models will require careful attention to basic principles of good engineering".

D.C. Van Essen, D.J. Felleman, E.A. DeYoe, J. Olavarria and J. Knierim (1990)

4.1 Introduction and Overview

In this chapter we propose neurobiologically inspired fundamental neuro-engineering design principles and synaptic mechanisms for non-linear visual artificial neural circuits. The proposed design principles and mechanisms are based primarily on the experimental data from the gill withdrawal reflex of the sea snail *Aplysia* (Kandel, 1979; Kandel and Schwartz, 1982, Hawkins *et al.*, 1983; Carew, 1987), the mathematical models of chemical synapses with transmitter mobilization (Grossberg, 1968, 1969; Carpenter and Grossberg, 1981), the mathematics of non-linear neural networks (Grossberg, 1973, 1988); the concepts of the Adaptive Resonance Theory (Grossberg, 1976, 1980) and the theoretical concept of Selective Attention Adaptive Resonance Theory (SAART) introduced in section 3.5 of Chapter 3. We also derive a family of new shunting competitive neural layers, called Presynaptically Modulated Shunting Competitive Neural Layers (PM-SCNLs) and a powerful new 2-D circuit called Feedforward Excitation-Feedback Presynaptic Facilitation (FFE-FBPF).

The proposed neuro-engineering design principles and mechanisms form the foundation for the design of cognitive and perceptual real-time artificial neural circuits and networks of subsequent chapters. It is assumed that the fundamental neuro-computational unit of cognitive biological information processing is a 2-D Presynaptically Modulated Shunting Competitive Neural Layer. Inputs to each cell in a PM-SCNL are gated by dynamic neural model of transmitter based chemical synapses whose signal transmission gain may in general be subject to various modes of modulation. The inhibitory interaction coefficients between neurons in a 2-D layer of competitive cells may in general be distance modulated and, unless otherwise stated, are presently assumed to be distance independent and symmetric. Symmetric interaction coefficients in a competitive layer of neurons provide stable pattern processing (Cohen and Grossberg, 1983).

We begin by generalising the model of a chemical synapse of Chapter 3 and consider how the internal synaptic dynamics may be altered and/or modulated by the following three types of interactions:

(i) influence of the synaptic input signal;

(ii) modulatory influences by the facilitatory/inhibitory gain control signals;

(iii) influence of pre-postsynaptic interactions.

The gain of a chemical synapse is defined as the amount of chemical neurotransmitter that is in the mobilized state and hence available for release. Initially we consider an idealised model of a chemical synapse in isolation and suggest how it may be extended and then we generalise the model by embedding it at the input stage of competitive neurons and consider the effect of feedback signal from the postsynaptic cells. The layer of neurons thus formed is called Presynaptically Modulated Shunting Competitive Neural Layer, whose properties are revealed through various applications and computer simulations. The basic structure of a PM-SCNL is initially derived in Appendix A and is then developed in section 4.3 of this chapter. The parameter design for layer stability is presented in Chapter 5.

The chapter addresses the various mechanisms and synaptic interactions for selective information transfer; mechanisms for synchronization of pulsating neurons; mechanisms for cooperative linking of non-pulsating neurons; top-down memory guided selective attention and the recognition of objects in cluttered visual backgrounds; synaptic mechanisms for the modulation and the fine-tuning of simple and complex receptive field profiles, and the recognition of distorted 2-D shapes. In subsequent chapters we will consider how the neuro-engineering design principles established here

can be further evolved and integrated into design logic for real-time cognitive and perceptual neural circuits and adaptive networks that can learn and recognize (perceive) a familiar object shape in cluttered and complex bottom-up inputs.

4.2 Modulation of Synaptic Transmission Gain

An idealised chemical synapse is a synapse that can be considered in isolation. The internal dynamics of an idealised chemical synapse is therefore independent of the postsynaptic activity. The Carpenter-Grossberg model discussed in Chapter 3 (section 3.3.1) is an example of an idealised chemical synapse that we now wish to extend by suggesting that signals other than those along the synaptic pathway may have influence on the transmitter dynamics. Specifically, the gain of a chemical synapse, which we have defined as the quantity of chemical neurotransmitter that is in the mobilized state and above the release threshold, may be altered by another signal that acts directly on the transmitter mobilization process. In biological systems, such signals are due to synapses that abut and modulate neural signal transmission of other synapses by releasing a neuromodulatory type of a chemical transmitter (as in the *Aplysia* circuit, section 2.3).

Below we present two complementary models, a facilitated synapse whose signal transmission gain is increased and an inhibitory model whose gain is decreased by the external (or non-synaptic) signal.

4.2.1 Model of Facilitated Chemical Synapses

Figure 4.1 illustrates the simplest model of a facilitated chemical synapse whose signal transmission gain is increased by the facilitatory signal F_i. The following two equations represent the internal dynamics of a facilitated synapse

$$\frac{du_i}{dt} = \alpha_u(z_i - u_i) - (\beta_u + K_u J_i)(u_i - y_i) \tag{4.1}$$

$$\frac{dy_i}{dt} = (\beta_y + F_i + J_i)(u_i - y_i) - \rho_y J_i[y_i - Y_i]^+ - \gamma_y y_i \tag{4.2}$$

where α_u, β_u and β_y represent the tonic activities of the synapse. The term $\alpha_u(z_i - u_i)$ in (4.1) says that $u_i(t)$ attempts to maintain a level z_i via the transmitter accumulation (production) rate $\alpha_u z_i$ and the feedback inhibition rate $-\alpha_u u_i$. The term $-(\beta_u + K_u J_i)u_i$, where K_u is a constant, says that the level of the stored transmitter is reduced by the input signal (some of which is lost to the mobilization process), which is counterbalanced by $(\beta_u + K_u J_i)y_i$. The term $(\beta_y + F_i + J_i)u_i$ in (4.2) says that the stored transmitter u_i is

transported (mobilized) to the release sites by the action of the facilitatory input F_i and by the synaptic input signal J_i. The term $-\rho_y J_i[y_i - Y_i]^+$, where ρ_y is a constant, says that the mobilized transmitter y_i is released by the input signal but only when $y_i > Y_i$, where $[y_i - Y_i]^+ = \max(y_i - Y_i, 0)$ and Y_i is a threshold for transmitter release. The last term in (4.2) represents the passive decay of the mobilized transmitter that is due to the enzymatic destruction of the transmitter at the release sites. Figure 4.2 shows the simulated response of the facilitated synapse.

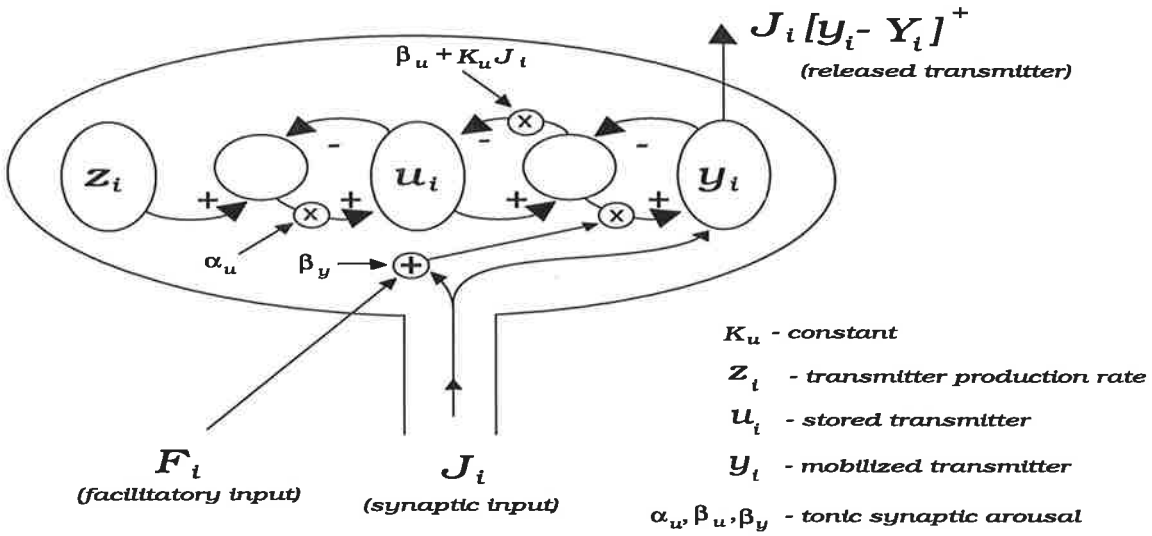


FIGURE 4.1. Simplest model of a facilitated chemical synapse.

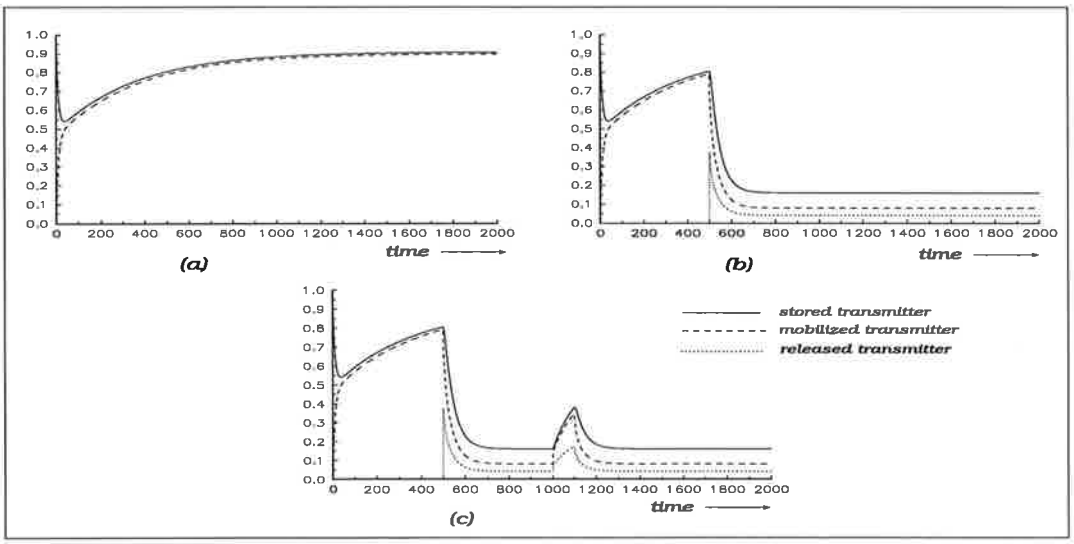


FIGURE 4.2. Response of the facilitated chemical synapse. (a) tonic synaptic activity; (b) synaptic input switched on at t = 500; (c) response of the synapse described by (4.1) and (4.2) when the facilitatory input is switched on at time t = 1000 and then switched off at t = 1100; (threshold for transmitter release in all the shown simulations

was set to zero). Parameters: $\alpha_u = 0.005$, $\beta_u = \beta_y = 0.05$, $K_u = 0.0005$, $\rho_y = 0.1$, $\gamma_y = 0.0005$, $F_i = 10$, $J_i = 1$, Y = 0. Equations (4.1) and (4.2) are iterated using Euler's first approximation method with $\Delta t = 1$.

The initial conditions in the above simulations are $z_i = 1$, $u_i(0) = 0$, $y_i(0) = 0$. For a useful set of synaptic parameters refer to Chapter 5 for the parameter design procedure.

It can be shown (see Chapter 5) that the steady state solution to equation (4.2), for zero thresholds (and constant inputs), is given by

$$y_i(\infty) = \frac{z_i a d}{ad + ae + ce} \tag{4.3}$$

where $a = \alpha_u$, $d = \beta_y + F_i + J_i$, $c = \beta_u + K_u J_i$, $e = \rho_y J_i + \gamma_y$. If the facilitatory signal F_i is small compared to β_y and J_i, then it is not going to have much of an effect on the synaptic gain. However, if F_i is very much larger, then it will dominate the synaptic gain.

When active, the facilitatory presynaptic signal F_i increases the transmitter mobilization rate and hence boosts the synaptic signal transmission gain. In the above model the facilitatory input F_i does not interact with the synaptic input J_i. If we assume that these two signals do interact to boost the level of transmitter even further, but under the condition of correlated firing of the facilitatory and the synaptic inputs (as is found in the *Aplysia* circuit), then we can write equation (4.1) in the following way

$$\frac{du_i}{dt} = (\alpha_u + F_i J_i)(z_i - u_i) - (\beta_u + K_u J_i)(u_i - y_i) \tag{4.4}$$

Figure 4.3 shows a more general model of a facilitated synapse whose transmitter levels are increased by the facilitatory signal F_i alone and by the multiplicative interaction of F_i and the synaptic input signal J_i. Figure 4.4 shows the response of the synapse.

Comparison of Figures 4.2 (c) and 4.4 shows that the correlated firing of the synaptic input and the facilitatory input has boosted the levels of transmitter above the levels provided by the action of the facilitatory signal alone. Whether this type of processing is required in more complex neural circuits will most likely depend on other system requirements. Throughout most of this chapter (and the rest of the thesis) we will in fact ignore this component and will concentrate primarily on the facilitation of the transmitter mobilization rate.

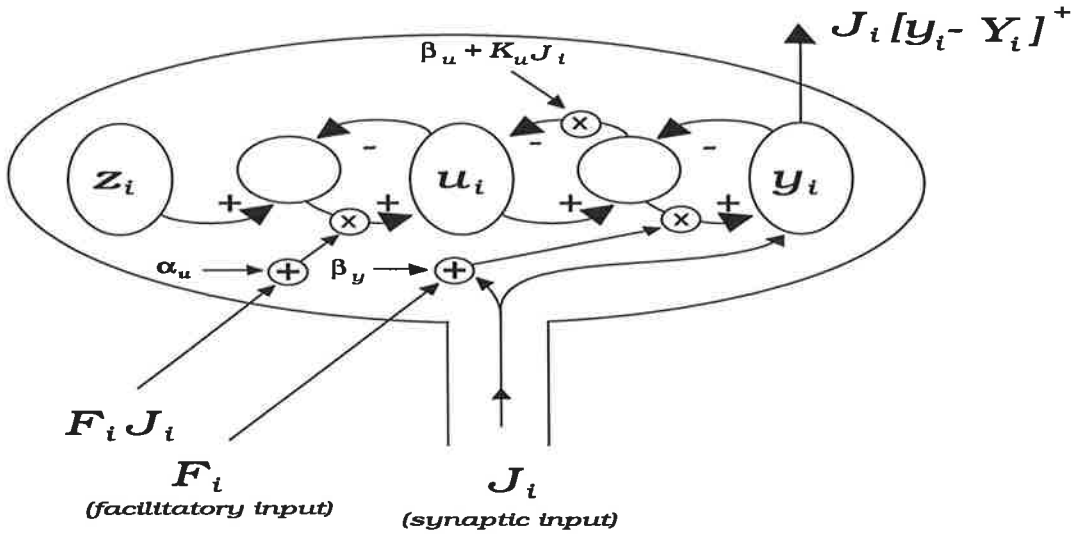


FIGURE 4.3. More general model of a facilitated chemical synapse.

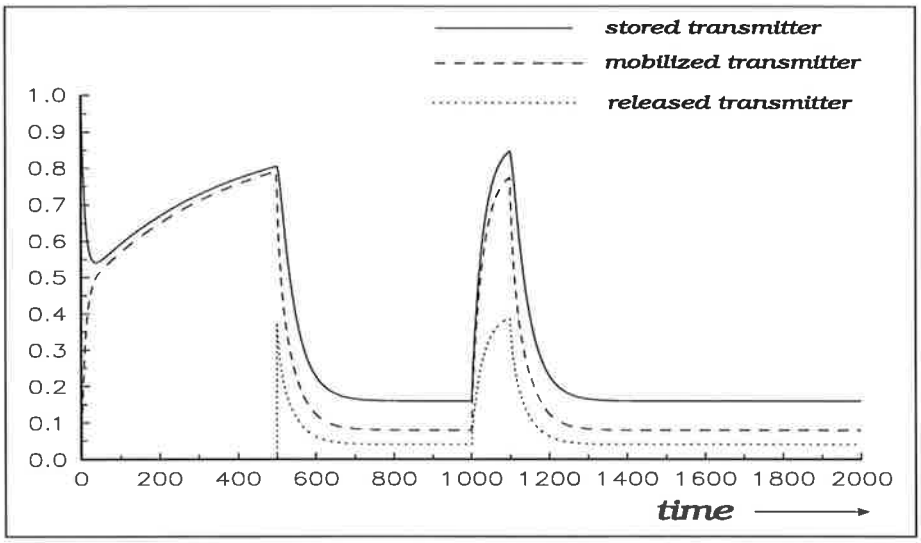


FIGURE 4.4. Response of the more general model of a facilitated chemical synapse. The synaptic input is switched on at t = 500 (and remains on), the facilitatory input is switched-on in the time interval 1000-1100. Parameters as for Fig. 4.2.

4.2.2 Model of Inhibited Chemical Synapses

The neural counterpart of the facilitated synapses is a synapse whose transmitter levels are reduced by the modulatory signals. The following two equations represent the internal dynamics of an inhibited synapse whose gain is decreased by an inhibitory signal I_i.

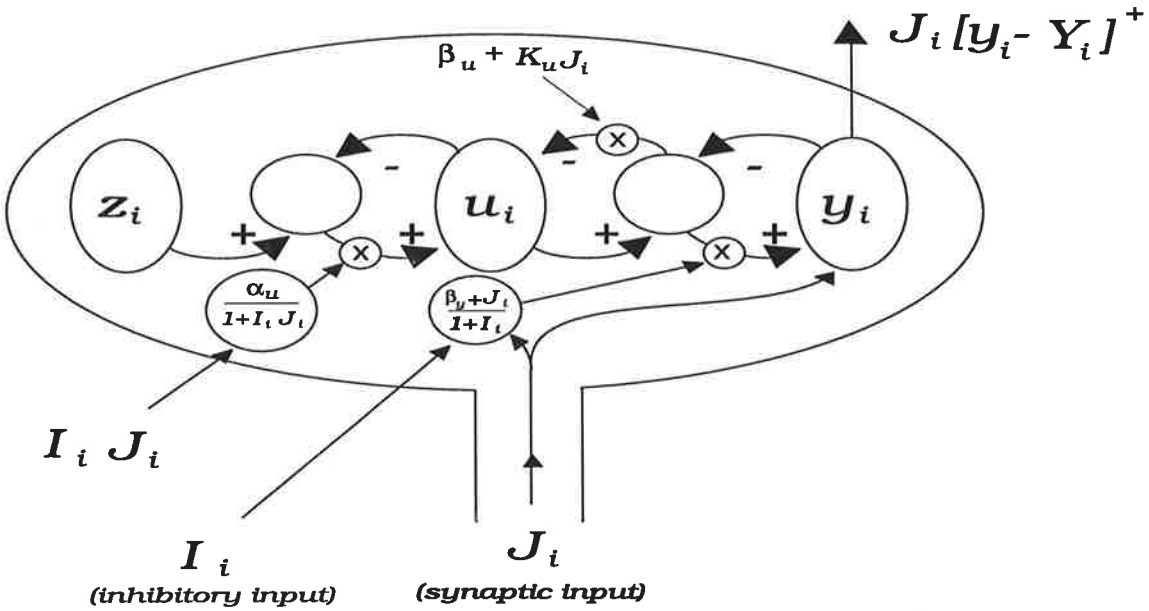


FIGURE 4.5. Model of an inhibited chemical synapse.

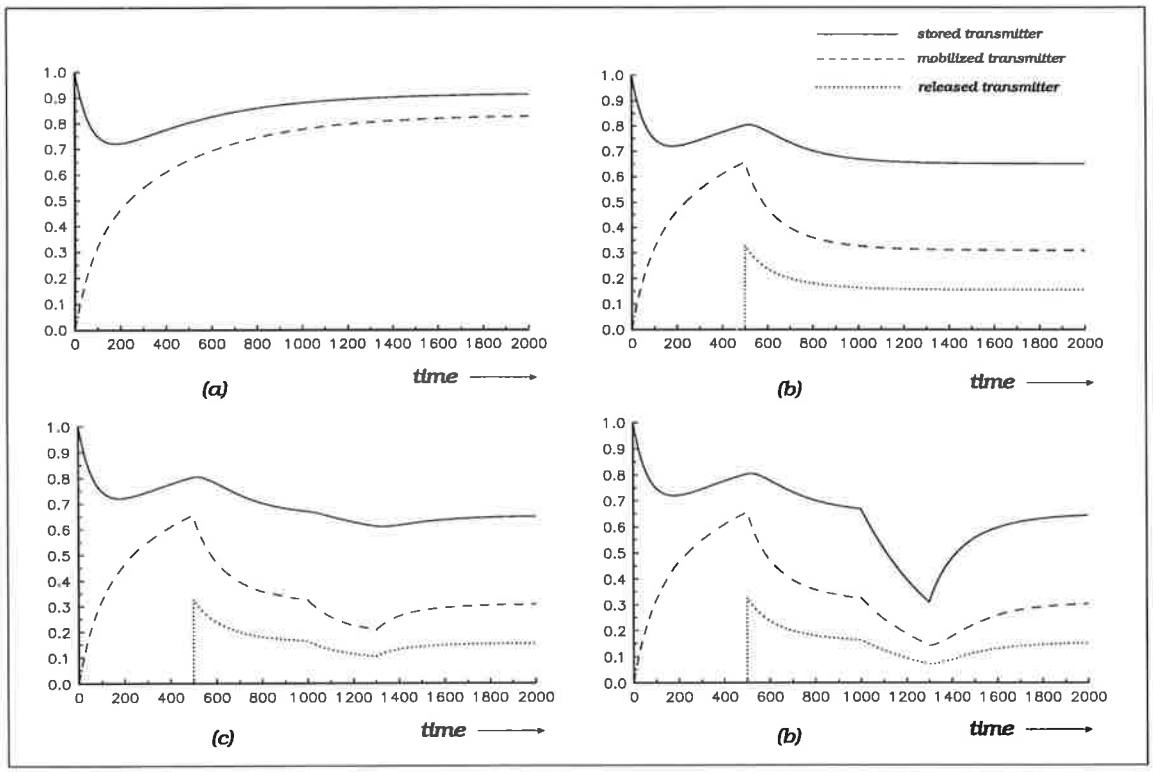


FIGURE 4.6. Response of the inhibited chemical synapse. (a) tonic synaptic activity; (b) synaptic input switched on at t = 500 (and remains on); (c) response of the synapse described by (4.3) and (4.4) when the inhibitory synaptic modulation is switched on at time t = 1000 and then switched off at t = 1300; (d) response under the additional condition of correlated firing.

$$\frac{du_i}{dt} = \frac{\alpha_u}{(1+I_iJ_i)}(z_i - u_i) - (\beta_u + K_uJ_i)(u_i - y_i) \tag{4.5}$$

$$\frac{dy_i}{dt} = \frac{(\beta_y + J_i)}{(1+I_i)}(u_i - y_i) - \rho_y J_i[y_i - Y_i]^+ - \gamma_y y_i \tag{4.6}$$

The inhibitory signal I_i affects the synapse by decreasing the rate of transmitter storage in (4.5) under the condition of correlated firing with the synaptic input J_i and is involved alone in the decrease of the mobilization rate in (4.6).

4.2.3 Synaptic Potentiation and Depression

In above sections we have assumed that the modulatory signal acts directly and instantaneously on the transmitter storage and mobilization rates. However, in biological systems such as in the neural circuit of the gill withdrawal reflex of *Aplysia* (and particularly in the mammalian hippocampus), the effect of modulatory signals tends to last for some time. Synaptic potentiation (depression) describes the durable effect of a facilitatory (inhibitory) gain control signal on the transmitter levels. In order to model the lasting modulatory effects on the synapse, we introduce two intermediate variables, $\alpha_i(t)$ and $\beta_i(t)$, whose dynamics is described by the following shunting equations:

$$\frac{d\alpha_i}{dt} = -A_\alpha\alpha_i + (1-\alpha_i)(A_0 + F_iJ_i) \tag{4.7}$$

$$\frac{d\beta_i}{dt} = -B_\beta\beta_i + (1-\beta_i)(B_0 + F_i) \tag{4.8}$$

where A_0 and B_0 are the tonic adaptation constants. Using these equations, we can rewrite the models of modulated synapses in the following way:

Long-term facilitated (potentiated) synapse

$$\frac{du_i}{dt} = A_u^+\alpha_i(z_i - u_i) - (\beta_u + K_uJ_i)(u_i - y_i) \tag{4.9}$$

$$\frac{dy_i}{dt} = A_y^+\beta_i(u_i - y_i) - \rho_y J_i[y_i - Y_i]^+ - \gamma_y y_i \tag{4.10}$$

Long-term inhibited (depressed) synapse

$$\frac{du_i}{dt} = \frac{\alpha_u(z_i - u_i)}{(1 + A_u^- \alpha_i)} - (\beta_u + K_u J_i)(u_i - y_i) \tag{4.11}$$

$$\frac{dy_i}{dt} = \frac{\beta_y(u_i - y_i)}{(1 + A_y^- \beta_i)} - \rho_y J_i [y_i - Y_i]^+ - \gamma_y y_i \tag{4.12}$$

Figure 4.7 below shows the simulated dynamics of the potentiated and the depressed synapse during various input conditions. Comparison of these results with those in Figs. 4.4 and 4.6 demonstrates the longer lasting nature of synaptic modulation by the potentiating and the depressing agents.

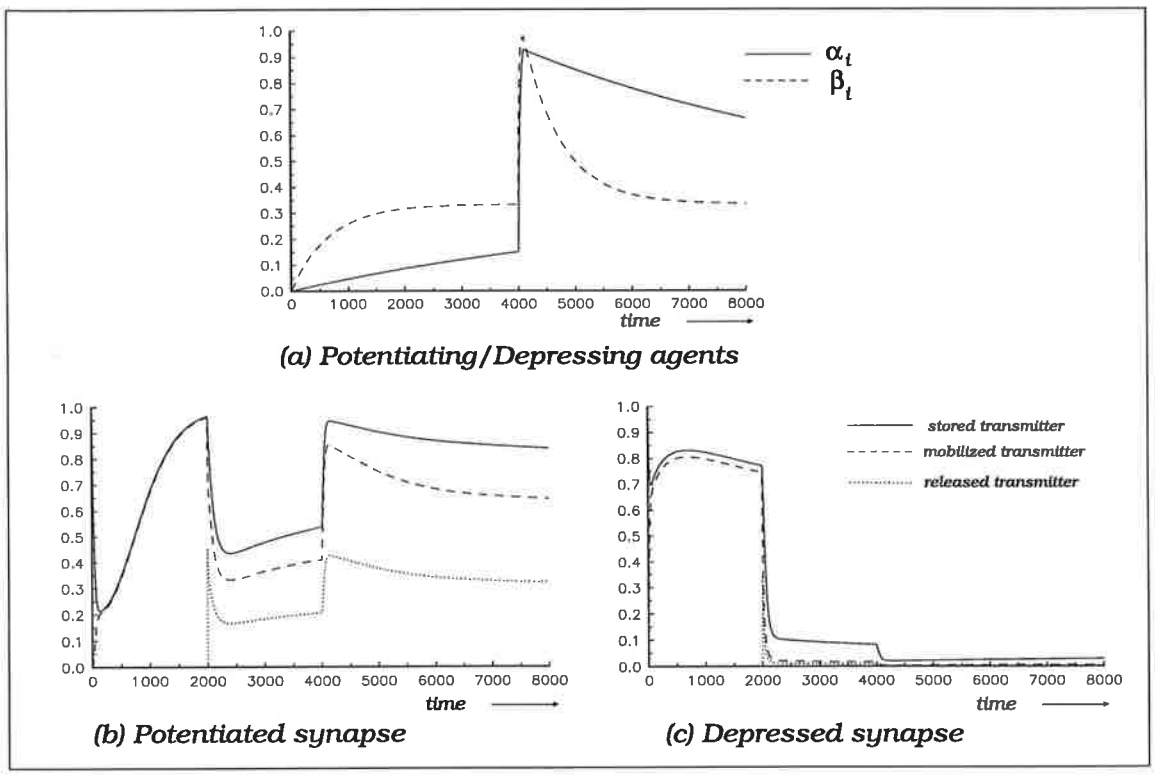


FIGURE 4.7. Simulation of long term synaptic potentiation and depression. (a) dynamics of the potentiating and the depressing agents (α_i and β_i) before($0 < t < 4000$), during ($4000 \leq t < 4100$) and after ($t \geq 4100$) the introduction of the synaptic modulatory signal; (b) response of a potentiated synapse (synaptic input switched on at $t = 2000$); and (c) response of a depressed synapse. Parameters: $A_\alpha = 0.0001$; $A_0 = 0.00005$; $B_\beta = 0.001$; $B_0 = 0.0005$; $F_i = 1$; $J_i = 0.5$; $\alpha_u = 0.005$; $\beta_u = \beta_y = 0.025$; $K_u = 0.005$; $\rho_y = 0.1$; $\gamma_y = 0.0005$; $A_u^+ = 10$; $A_y^+ = 10$; $A_u^- = 10$; $A_y^- = 10$, $Y = 0$.

The parameters for the above simulation were simply chosen to demonstrate the dynamics of long term synaptic effects rather than the robustness of the model. The many parameters that are involved may be chosen on the basis of the required temporal dynamics which can only be determined from the more global system requirements. In the next section we propose how the model of modulated chemical synapses may be embedded at the input stage of a shunting competitive neural layer to provide new computational capabilities.

4.3 Presynaptically Modulated Neural Layers

We now use the above mechanisms of synaptic modulation to provide a mathematical model and the circuit implementation for a new family of shunting competitive neural layers called Presynaptically Modulated Shunting Competitive Neural Layers (PM-SCNLs). PM-SCNLs are characterized by non-linear differential equations that represent the dynamics of a layer of shunting competitive neurons whose inputs are gated by dynamic models of modulated chemical synapses. PM-SCNLs represent our neurobiologically motivated generalisation of S. Grossberg's shunting competitive neural layers (Grossberg, 1973, 1988) by incorporating the synaptic modulatory mechanism to provide new computational capabilities.

To begin with, we initially ignore the postsynaptic feedback and propose a simple feedforward modulated neural layer with single excitatory synapse into each cell. We then extend the layer through several stages, first by incorporating the post-synaptic feedback and then bipolar synaptic inputs into each cell. Throughout the section we will address one-dimensional layers. Extensions to two-dimensions will be done in the latter part of the chapter (section 4.10).

The mathematical and an intuitive procedure for the parameter design of presynaptically modulated shunting competitive neural layers that ensures stability, wide dynamic range and invariance to the size of the layer is described in Chapter 5.

4.3.1 Simple Model of Presynaptically Modulated Competitive Neural Layers

Figure 4.8 illustrates an example of simple Presynaptically Facilitated Excitatory Shunting Competitive Neural Layer (PFE-SCNL) that is actually implemented in two layers, a layer of fast excitatory cells and a layer of slow inhibitory interneurons (large black circles). The inhibitory neural layer mediates the shunted competition in the fast excitatory layer. Throughout the subsequent sections and chapters we will refer to these two layers as a *Field* of competitive neurons.

Postsynaptic cellular activity

Cellular activity of the i^{th} neuron in the excitatory layer is described by the following non-linear differential equation

$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)Gv_i^+ - (C + x_i)(\overline{G}\overline{v}_i + \Gamma) \tag{4.13}$$

where Gv_i^+ is the amplified transmitter gated excitatory postsynaptic potential, given by (4.14), G is the amplification factor; $\overline{G}\overline{v}_i$ is the amplified lateral feedback inhibition that is mediated by slow inhibitory interneurons, given by (4.15); Γ is the tonic level of inhibition; A is the passive decay rate, B is the upper saturation level and C is the lower saturation level (A, B, C > 0). This equation thus represents shunted competition of a layer of neurons with the on-centre off-surround anatomy whose cellular activity is restricted to range (-C, B).

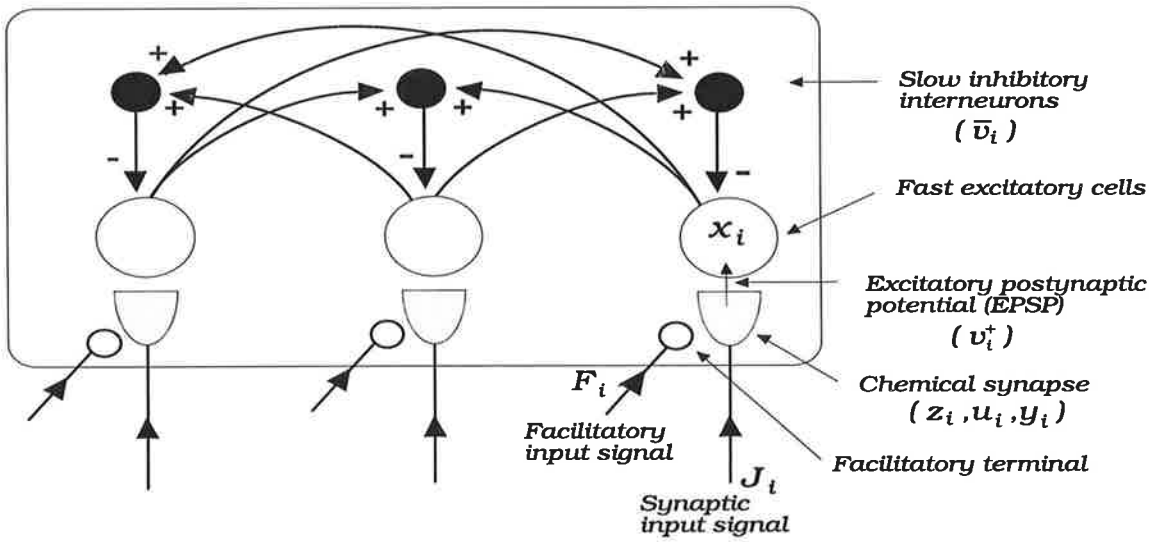


FIGURE 4.8. Simplest implementation of a Presynaptically Facilitated Excitatory Shunting Competitive Neural Layer. Note that this model does not include the postsynaptic feedback from a cell to its synapse.

Each neuron in the excitatory layer receives its excitatory input through a facilitated synaptic pathway whose internal dynamics is slower than the postsynaptic dynamics. The following set of equations describe the dynamics of the layer:

Excitatory postsynaptic potential

The excitatory postsynaptic potential (EPSP) acting on a cell is due to the bound transmitter on the postsynaptic cell and is given by

$$\frac{dv_i^+}{dt} = -Dv_i^+ + \rho_v J_i [y_i - Y]^+ \tag{4.14}$$

where D and ρ_v are constants. The term $\rho_v J_i [y_i - Y_i]^+$, where $[y_i - Y_i]^+ = \max(y_i - Y_i, 0)$, says that the EPSP is charged by the released (and bound) transmitter, but only when the synaptic input J_i is active and when the mobilized transmitter is above the release threshold of Y.

Lateral feedback inhibition

The lateral feedback inhibition ($\overline{v}_i$) is assumed to be due to slowly charging inhibitory interneurons whose potential is given by

$$\frac{d\overline{v}_i}{dt} = -\overline{A}\,\overline{v}_i + \overline{B} \sum_{j \neq i} f(x_j) \tag{4.15}$$

where $\overline{A}$ and $\overline{B}$ are positive constants ($\overline{A} < A, \overline{B} < B$); $f(x_j) = \max(x_j - \Theta, 0) = [x_j - \Theta_j]^+$ is a thresholding function which says that the postsynaptic cellular activity must be above the threshold Θ before the cell fires, with the assumption that each cell has the same threshold.

Postsynaptic cellular activity at equilibrium

Since the excitatory postsynaptic potential ($v_i^+(t)$) results from the gating of fast input neural signals by slow chemical synapses and because the lateral feedback inhibition is mediated by slowly charging inhibitory neurons that have fast interactions with their target cells, we can represent (4.13) by its equilibrium state (4.16). Thus, by assuming that the postsynaptic cell reaches its equilibrium before a significant change occurs in its input synapses, we obtain the following approximation by letting $\frac{dx_i}{dt} = 0$ in (4.13).

$$x_i = \frac{BGv_i^+ - C(\overline{G}\overline{v}_i + \Gamma)}{A + Gv_i^+ + \overline{G}\overline{v}_i + \Gamma} \tag{4.16}$$

which has a range $(-C, B)$. If we restrict the operating range of (4.16) to positive values by letting $C = 0$ and if $\Gamma = 0$, then (4.16) can be simplified to

$$x_i = \frac{BGv_i^+}{A + Gv_i^+ + \overline{G}\overline{v}_i} \tag{4.17}$$

To complete the minimum model of a simple presynaptically modulated competitive neural layer, below we write the two equations that represent the dynamics of the simplest facilitated chemical synapse whose internal dynamics is independent of the postsynaptic cellular activity (a more general model is presented in the next section).

Stored transmitter

$$\frac{du_i}{dt} = \alpha_u(z_i - u_i) - (\beta_u + K_u J_i)(u_i - y_i) \tag{4.18}$$

Mobilized transmitter

$$\frac{dy_i}{dt} = (\beta_y + F_i)(u_i - y_i) - \rho_y J_i[y_i - Y_i]^+ - \gamma_y y_i \tag{4.19}$$

We have presently left out the facilitation of the storage transmitter. This may be included when needed and, as already mentioned, most of the modelling work presented in subsequent sections and chapters is based primarily on the facilitation of the transmitter mobilization process. We have also left out the contribution to transmitter mobilization that may be due to the synaptic input signal. This component may be incorporated when required, but is presently assumed to be very small compared to the facilitatory signal F_i (see Chapter 5 for the rationale of large facilitatory gain). The transmitter production level (z_i) may also be subject to temporal adaptation, but is presently assumed to be constant ($z_i = 1$).

Figure 4.9 shows the feedforward (unilateral) interactions between an excitatory input synapse and its postsynaptic cell, while Figure 4.10 shows the equilibrium response of one neuron in the layer (and its lateral feedback inhibition) as the number of active inputs (all equal in intensity) in a 1024 neuron layer is increased from 1 to 1024. As can be seen, the equilibrium activity of the cell decreases with an increase in the number of active synaptic inputs. This is due to the greater lateral feedback inhibition provided by the extra active cells.

Note that in the shown simulation, the parameters were not chosen in any predetermined way and the emphasis is on showing the effect of competition on the cellular activity as a function of the number of active inputs. The simple model of a presynaptically modulated neural layer described above does not take into account the possibility that the excitatory layer may also receive inhibitory synaptic inputs.

Furthermore, the model has ignored the effect of postsynaptic feedback signals on the transmitter release process. In the next section we generalise the model to take into account the bidirectional pre-postsynaptic interactions and bipolar synaptic inputs.

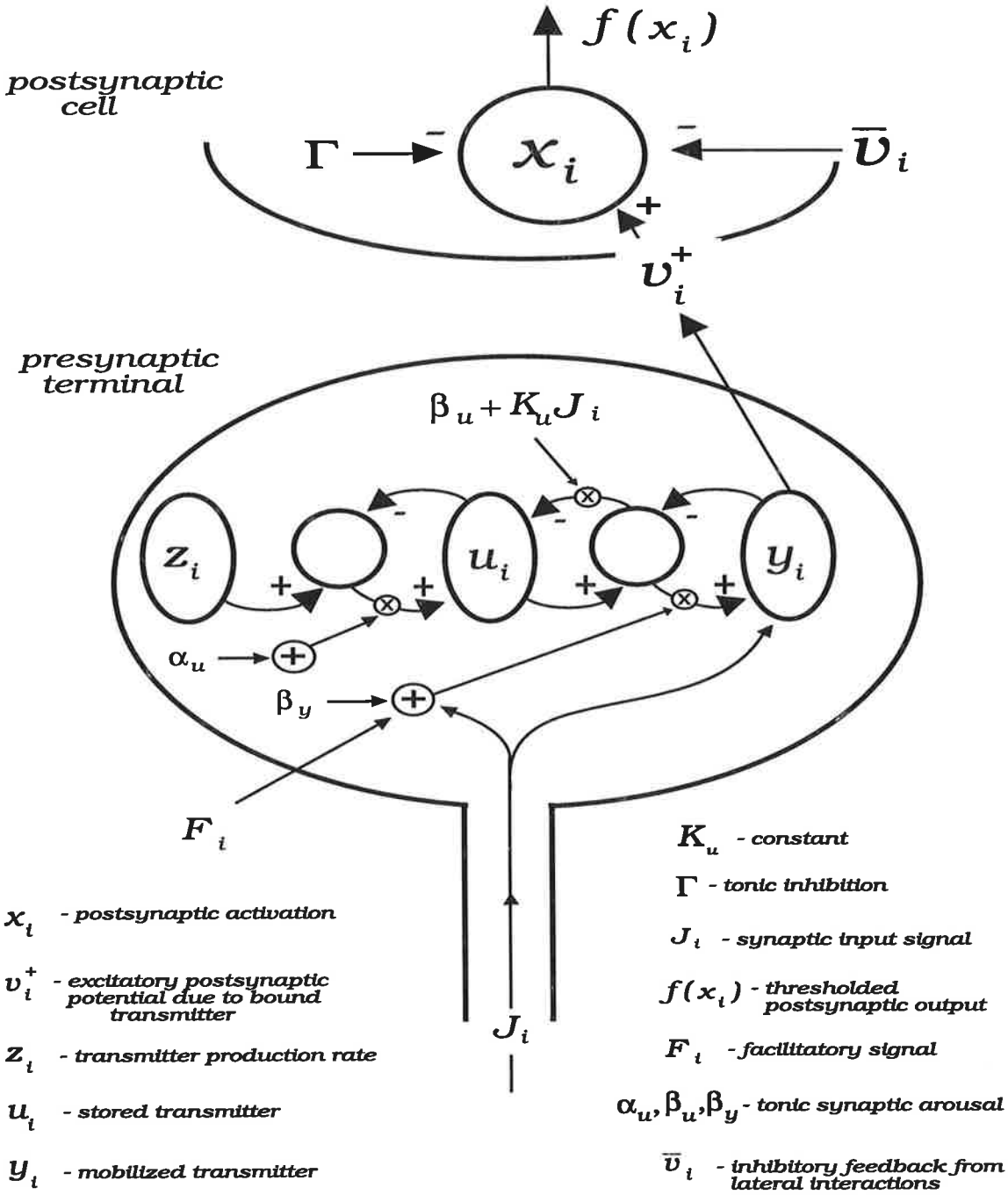


FIGURE 4.9. Feedforward interactions between a synapse and its postsynaptic cell.

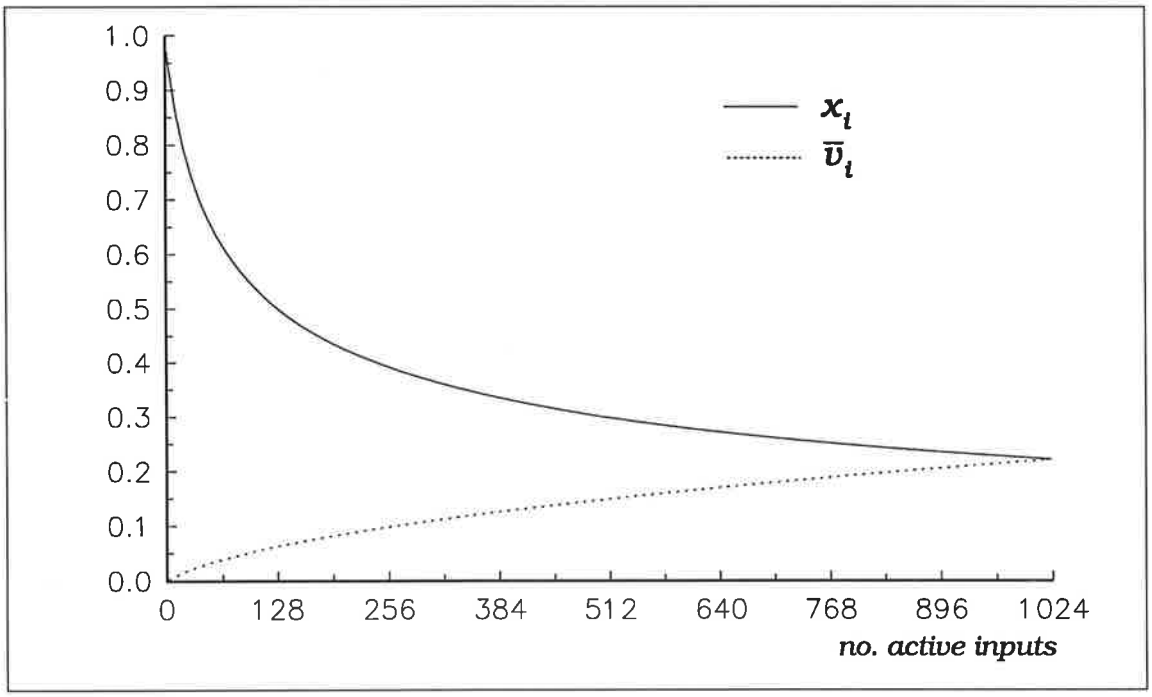


FIGURE 4.10. Property of the simple transmitter gated shunting competitive neural layer. Each plot is obtained after the layer reached a steady state (typically after 20 iterations). Parameters: A = B = 1; C = 0; D = 0.5; $\rho_v = 0.25$; $Y_i = 0$; $\Theta_i = 0$; $\Gamma = 0$; $\alpha_u = 0.1$; $\beta_u = \beta_y = 0.2$; $K_u = 0.2$; $F_i = 0.2$; $\rho_y = 0.5$; $\gamma_y = 0.05$; G = 100; $\overline{G} = 1000$; $J_i = 1$; $\overline{A} = 0.2$; $\overline{B} = 0.2/n$, where n is the number of neurons in the layer (= 1024), Y = 0.

4.3.2 General Model of Presynaptically Modulated Competitive Neural Layers

In the previous section we have presented an idealised model of a modulated chemical synapses whose internal dynamics is independent of postsynaptic cellular activity. We now consider the effect of feedforward-feedback interactions between a synapse and its target cell and extend the model to bipolar synaptic inputs and bipolar synaptic modulation (synaptic facilitation and synaptic inhibition).

Figure 4.11 illustrates a presynaptically modulated competitive neural layer whose neurons have one excitatory and one inhibitory input synapse. The excitatory synapse is represented by an unfilled semi-ellipse, while the inhibitory synapse is shown shaded. Pre-postsynaptic interactions are represented by the arrow from the post-synaptic cell to its input synapse.

In general, each neuron in a PM-SCNL may have a number of excitatory and inhibitory synaptic inputs whose combined effect determines the cellular response. Each of the input synapses may also be modulated by facilitatory/inhibitory gain control signals. In order to develop a general mathematical model of an input gated and presynaptically modulated shunting competitive neural layer, we first consider the qualitative properties of the model described by equation (4.20). This is a version of the shunting cooperative-competitive feedback neural layer with the on-centre off-surround anatomy that was introduced by S. Grossberg (1973).

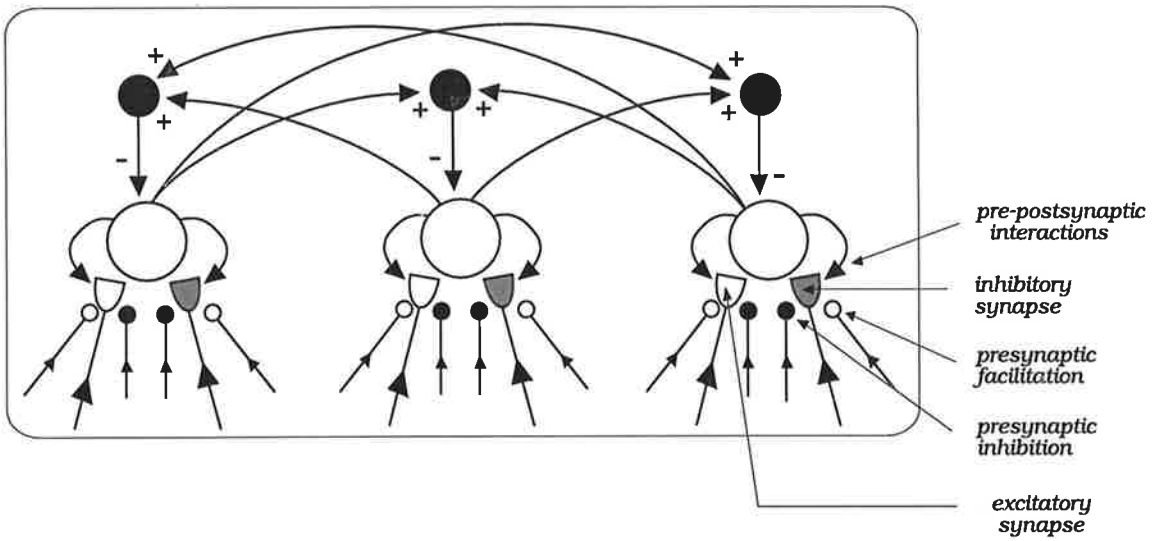


FIGURE 4.11. General model of a Presynaptically Modulated Shunting Competitive Neural Layer.

$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)[J_i^+ + f(x_i)] - (C + x_i)\left[J_i^- + \sum_{j=1}^{n} F_{ji}g(x_j)\right] \tag{4.20}$$

In Appendix A we show that a neural layer described by the above equation has the undesirable characteristic where the inhibitory synaptic inputs do not effectively regulate the cellular response. For example, it can be shown that if the total excitatory synaptic input J_i^+ (acting on the i^{th} cell) is exactly matched in amplitude by the total inhibitory synaptic input J_i^- acting on the same cell, then it is possible for the cellular activity x_i to exceed the threshold for self-excitation and become highly active. This is not very useful and would make it awkward to engineer a neural network whose layers use both excitatory and inhibitory input synapses. In order to overcome this undesirable characteristic of the system described by equation (4.20), we propose the following equation (see Appendix A) as having more desirable characteristics for the neuro-engineering design of circuits whose neural layers use bipolar synaptic inputs.

$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)[G(V_i^+ - V_i^-)]^+ - (C + x_i)\left([G(V_i^- - V_i^+)]^+ + \overline{G}\overline{v}_i + \Gamma\right) \tag{4.21}$$

where $[G(V_i^+ - V_i^-)]^+ = \max(G(V_i^+ - V_i^-), 0)$ is the net excitatory postsynaptic potential that drives the cell towards its positive saturation limit of B; $[G(V_i^- - V_i^+)]^+ + \overline{G}\overline{v}_i + \Gamma$ is the net shunting inhibition that drives the cell towards its negative saturation level of -C; G is the gain of excitatory and inhibitory synaptic inputs (assumed equal); Γ is the tonic level of inhibition.

With the exception of the self-excitatory feedback term that appears in (4.20), which we will incorporate below in a different way, the main difference between (4.20) and (4.21) is that in the latter the excitatory and the inhibitory inputs interact additively before they affect the cell. The equilibrium solution of (4.21) is given by

$$x_i = \frac{B[G(V_i^+ - V_i^-)]^+ - C\left([G(V_i^- - V_i^+)]^+ + \overline{G}\overline{v}_i + \Gamma\right)}{A + [G(V_i^+ - V_i^-)]^+ + [G(V_i^- - V_i^+)]^+ + \overline{G}\overline{v}_i + \Gamma} \tag{4.22}$$

Note that when $V_i^+ = V_i^-$, the equilibrium state is independent of the synaptic inputs and is given by

$$x_i = -\frac{C(\overline{G}\overline{v}_i + \Gamma)}{A + (\overline{G}\overline{v}_i + \Gamma)} \tag{4.23}$$

Since $x_i < 0, \quad \forall i$, the lateral inhibition will decay to zero (i.e., $\overline{v}_i \to 0, \quad \forall i$). Hence the resting state is given by

$$x_i = -\frac{C\Gamma}{A + \Gamma} \tag{4.23a}$$

Throughout this chapter (and the rest of the thesis), we will base the design of presynaptically modulated shunting competitive neural layers on a version of (4.21) with C = 0, shown below.

$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)[G(V_i^+ - V_i^-)]^+ - x_i\left([G(V_i^- - V_i^+)]^+ + \overline{G}\overline{v}_i + \Gamma\right) \tag{4.24}$$

In order to generate the model of (4.21) we have ignored the positive self-excitation term ($f(x_i)$) that appears in (4.20). The model described by (4.20) has the property of retaining a contrast enhanced memory of its inputs after the input offset, while the model described by (4.21) does not have this property. Since biological neurons are oscillators (i.e., they fire at some frequency in response to a fixed driving signal), it is not possible for one layer of neural oscillators to retain a memory of their driving signals once their inputs are removed. Hence, the system described by (4.20) cannot be a neurobiologically plausible method of storing short term memory. In Chapter 6 we will provide several alternate solutions to how the short term memory may be maintained between several interacting neural layers. However, neurobiological data does suggest that postsynaptic feedback signals do interact with the transmitter dynamics by increasing transmitter release. To take this into account, we suggest that the post-synaptic cellular activity interacts with the dynamics of the synaptic variables to enhance transmitter release thus providing self-excitation. There are two possible means by which postsynaptic signal may contribute to cellular self-excitation: (i) direct release of the mobilized synaptic transmitter; or (ii) release of the mobilized presynaptic transmitter under the condition of correlated firing of the synaptic input and the postsynaptic cell.

Below we write a set of equations that define a general model of presynaptically modulated shunting competitive neural layers with the on-centre off-surround anatomy and distance independent competitive interactions. We will initially consider a pres-ynaptically facilitated neural layer whose cells have one excitatory and one inhibitory synapse. The modulated neural layer presented in this section will be extended to multiple and spatially distributed synapse in section 4.10.

Postsynaptic Cellular activity at Equilibrium

The equilibrium postsynaptic cellular activity of (4.24) is given by

$$x_i = \frac{B[G(V_i^+ - V_i^-)]^+}{A + [G(V_i^+ - V_i^-)]^+ + [G(V_i^- - V_i^+)]^+ + \overline{G}\overline{v}_i + \Gamma} \tag{4.25}$$

Since we are presently addressing the case of a single excitatory and a single inhibitory synapse, we can use the same indices for synaptic variables as for the cellular variables. However we need to be able to distinguish between the two types of synapses. Excitatory synaptic variables (and the associated constants) will therefore be labelled by a positive superscript, whereas the inhibitory synaptic variables will be labelled by a negative superscript. If we assume that all cells have the same postsynaptic threshold of Θ and that all synapses have the same threshold for transmitter release, then we can write the synaptic dynamics as follows:

Excitatory postsynaptic potential

The excitatory postsynaptic potential (EPSP) acting on a cell is due to the bound transmitter on the postsynaptic cell and is given by

$$\frac{dv_i^+}{dt} = -D^+ v_i^+ + J_i^+ [y_i^+ - Y]^+ [\rho_v^+ + K_v^+ f(x_i)] \tag{4.26}$$

where D^+, ρ_v^+ and K_v^+ are constants in the excitatory synaptic pathway (ρ_v^+ and K_v^+ determine the relative contributions to the EPSP from the uncorrelated and the correlated pre-postsynaptic signals respectively). The equation says that when $K_v^+ > \rho_v^+$, the major contribution to the EPSP is from the correlated firing of the synaptic input and the postsynaptic cell. The term $[y_i^+ - Y^+]^+ = \max(y_i^+ - Y^+, 0)$ says that the transmitter in the excitatory synaptic pathways can be released only if it is above the release threshold level of Y; while the function $f(x_i) = \max(x_i - \Theta, 0)$ is linear above a threshold. That is, the postsynaptic feedback signal needs to be above the threshold of Θ before it can affect modify the EPSP.

In equation (4.26) we have assumed that once activated above a firing threshold, each cell contributes to its self-excitation but under the condition of correlated firing with the synaptic input signal J_i^+. The alternate case where the activated cell can self-excite by direct release of its presynaptic transmitter is given by (4.27).

$$\frac{dv_i^+}{dt} = -D^+ v_i^+ + [y_i^+ - Y_i^+]^+ [\rho_v^+ J_i^+ + K_v^+ f(x_i)] \tag{4.27}$$

The EPSP modelled by the above equation produces oscillatory (pulsating) behaviour in a layer of neurons. With the exception of section 4.5 where we will propose several synaptic mechanisms for the synchronization of a layer of pulsating neurons, this model (equation (4.27)) will be largely ignored in the thesis.

Dynamics of pre-postsynaptic interactions in an excitatory synapse

The gain of an excitatory chemical synapse (which we have defined as the amount of transmitter that is in the mobilized state and above the release threshold of Y) can be increased by the facilitatory presynaptic signals F_i (or decreased by the inhibitory presynaptic signal I_i). The following two equations represent the transmitter variables in a facilitated excitatory chemical synapse.

$$\frac{du_i^+}{dt} = \alpha_u^+(z_i^+ - u_i^+) - [\beta_u^+ + J_i^+ K_u^+ f(x_i)](u_i^+ - y_i^+)$$

(4.28)

which says that the transmitter storage rate in the excitatory synapse is depleted by the correlated firing of the input signal J_i^+ and the postsynaptic feedback signal $f(x_i)$ (refer to the parameter design procedure and computer simulations in Chapter 5).

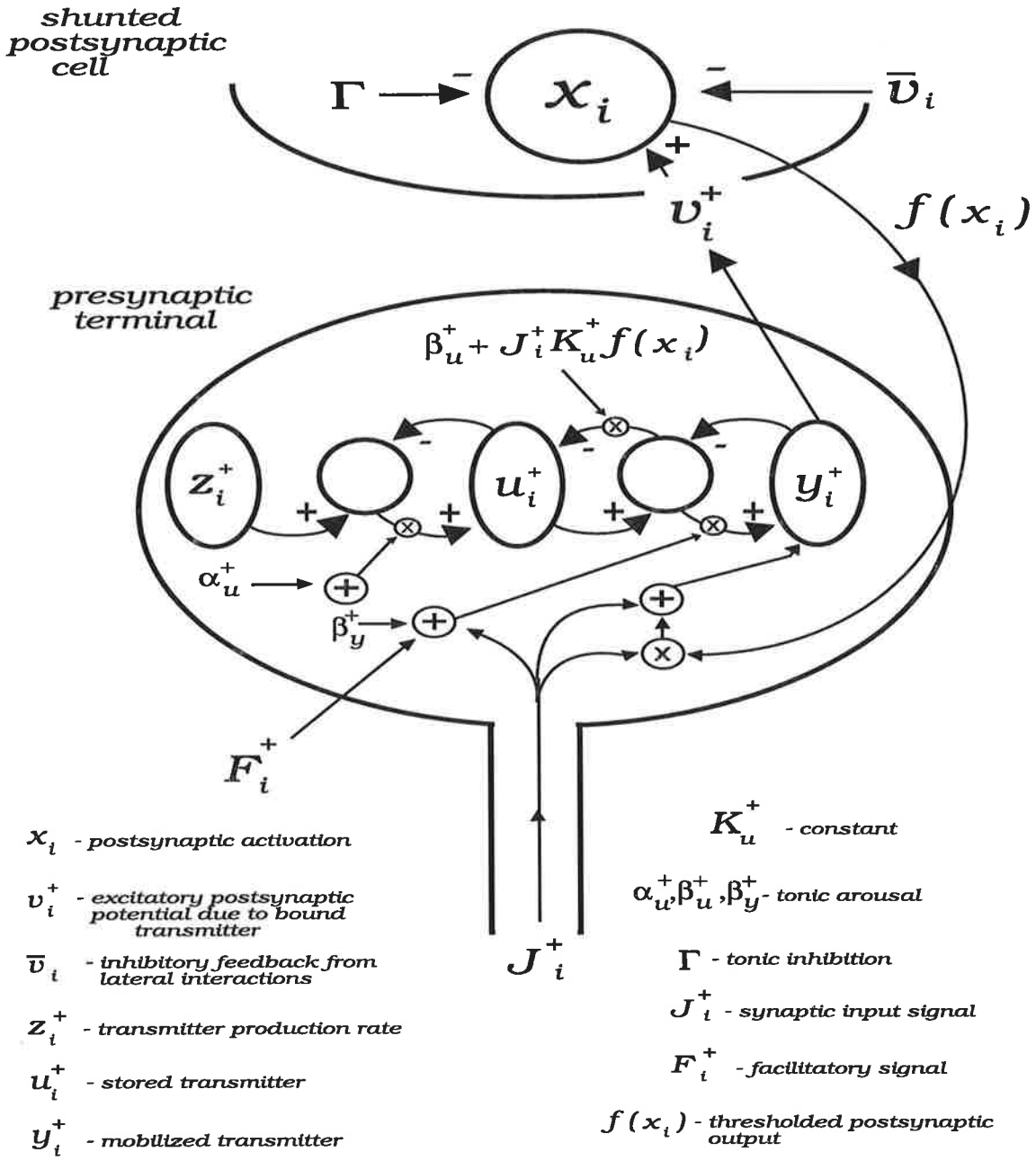


FIGURE 4.12. Model of pre-postsynaptic interactions in a facilitated excitatory chemical synapse. The transmitter production rate (z_i^+) is assumed to be constant.

$$\frac{dy_i^+}{dt} = (\beta_y^+ + F_i^+)(u_i^+ - y_i^+) - J_i^+[y_i^+ - Y]^+[\rho_y^+ + K_y^+ f(x_i)] - \gamma_y^+ y_i^+ \tag{4.29}$$

where β_y^+ is a tonic level of transmitter mobilization. This equation says that the transmitter mobilization rate is increased by the facilitatory signal F_i^+ and that the transmitter is released (and hence depleted) by the input signal J_i^+ and by the correlated firing of J_i^+ and the postsynaptic feedback signal $f(x_i)$ (provided that the level of mobilized transmitter is above the release threshold of Y. The postsynaptic cellular activity must be above the threshold of Θ before it can influence the synapse.

Note that the transmitter production rate (z_i^+) is presently assumed to be constant. In subsequent chapters we will generalise our present model of chemical synapses by taking into account the possibility that the transmitter production rate may also be subjected to adaptation (i.e., learning) but on a slower time scale. The above two equations do not include the dynamics of the long term potentiating/depressing agents (α_i and β_i) of section 4.1.3. These may be modelled when required and, unless otherwise stated, will be generally ignored in subsequent sections and chapters.

Dynamics of pre-postsynaptic interactions in an inhibitory synapse

For completeness we represent the dynamics of inhibitory synapses by a similar set of equations, with an understanding that its tonic synaptic activity does not necessarily have to be the same as for the excitatory synapses.

Inhibitory postsynaptic potential

The inhibitory postsynaptic potential (IPSP) due to an inhibitory input synapse acting on the cell is given by

$$\frac{dv_i^-}{dt} = -D^- v_i^- + J_i^-[y_i^- - Y]^+[\rho_v^- + K_v^- f(x_i)] \tag{4.30}$$

where D^-, ρ_v^- and K_v^- are the time constants in the inhibitory synaptic pathway.

Transmitter variables

The following two equations represent the transmitter variables in a facilitated inhibitory chemical synapse.

$$\frac{du_i^-}{dt} = \alpha_u^-(z_i^- - u_i^-) - [\beta_u^- + K_u^- J_i^- f(x_i)](u_i^- - y_i^-)$$

(4.31)

$$\frac{dy_i^-}{dt} = (\beta_y^- + F_i^-)(u_i^- - y_i^-) - J_i^-[y_i^- - Y]^+[\rho_y^- + K_y^- f(x_i)] - \gamma_y^- y_i^-$$

(4.32)

Presynaptically Inhibited Shunting Competitive Neural Layer (PI-SCNL) is represented by similar set of equations, with the exception that the excitatory and inhibitory synaptic variables are presynaptically attenuated by an inhibitory signal I_i.

The general model of presynaptically facilitated shunting competitive neural layers presented above (equation (4.15) of section 4.3.1 and equations (4.25) - (4.32)) assumes that each cell in a layer interacts competitively with all the other cells through distance independent interaction coefficients. Extension to distance dependant (but symmetric) interactions can be easily envisaged. In the next section we briefly discuss the computational properties of two-dimensional shunting competitive neural layers.

4.4 Properties of 2-D Competitive Neural Layers

A two dimensional neural layer of competitive cells whose dynamics is described by the shunting competitive equation has some important pattern processing properties. The two most basic properties are illustrated below. Figure 4.13 illustrates how the steady state cellular activity of a layer varies with the input signal strength. At low levels of the input, the steady state response of the layer is also low in magnitude and increases with the increase in the input strength, but in a converging fashion. The non-linear convergence is due to the competitive inhibition of cells as well as the inhibition from term $(B - x_{ij})$ (see equation (4.13)). This ensures that the steady state activity of each cell in the layer converges to magnitude that is below the saturation point.

Thus if the input signal strength is held fixed, then the complexity of the input (number of active input pathways) will determine the magnitude of the cellular activity. Figure 4.14 illustrates how the amplitude of the cellular activity is increased when the density of the input is decreased (i.e., when the number of active input pathways is decreased, while keeping the strength of the other inputs constant). The reason for the increase in the cellular activity is due to a decrease in the level of lateral inhibition.

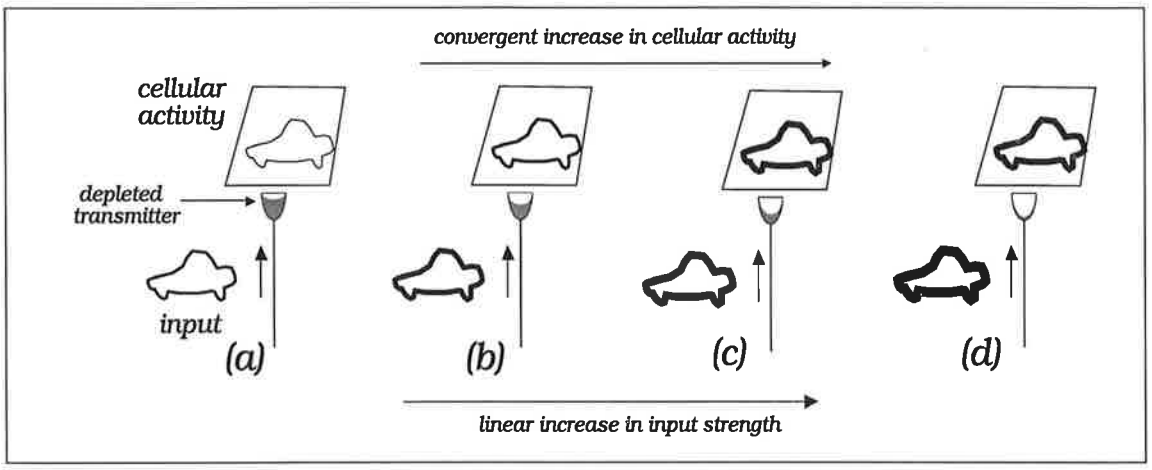


FIGURE 4.13. Effect of input signal strength on the steady state cellular activity in a shunting competitive neural layer (inputs are assumed to be excitatory). The thickness of the object's boundary in a particular location represents the magnitude of the associated steady state (equilibrium) cellular activity at that location (thicker boundary implies stronger activity). Transmitter levels are depleted by the correlated pre-postsynaptic activity.

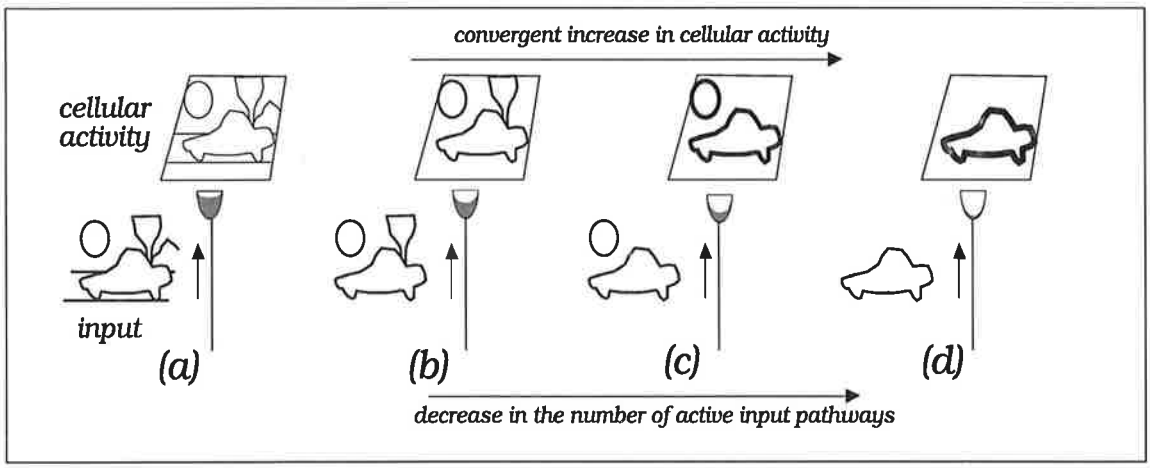


FIGURE 4.14. Effect of input signal density on the steady state cellular activity in a shunting competitive neural layer.

In Figure 4.15 we illustrate how the activity of one shunting competitive layer can be altered when it receives signals from two different sources. We wish to consider what happens to the activity of this layer when the signals from one of its sources is steadily increased in strength, while the signals from the second source are held constant. Two two source Fields, F0 and F2, contain a common spatial pattern of neural activity (boundary of a car that is perfectly aligned in all respects), but the source Field

F0 also includes other patterns of neural activity. When the signals from Field F2 are very much stronger than the signals from Field F0, the competitive interactions in Field F1 will suppress the activity that was initially due to the inputs from Field F0.

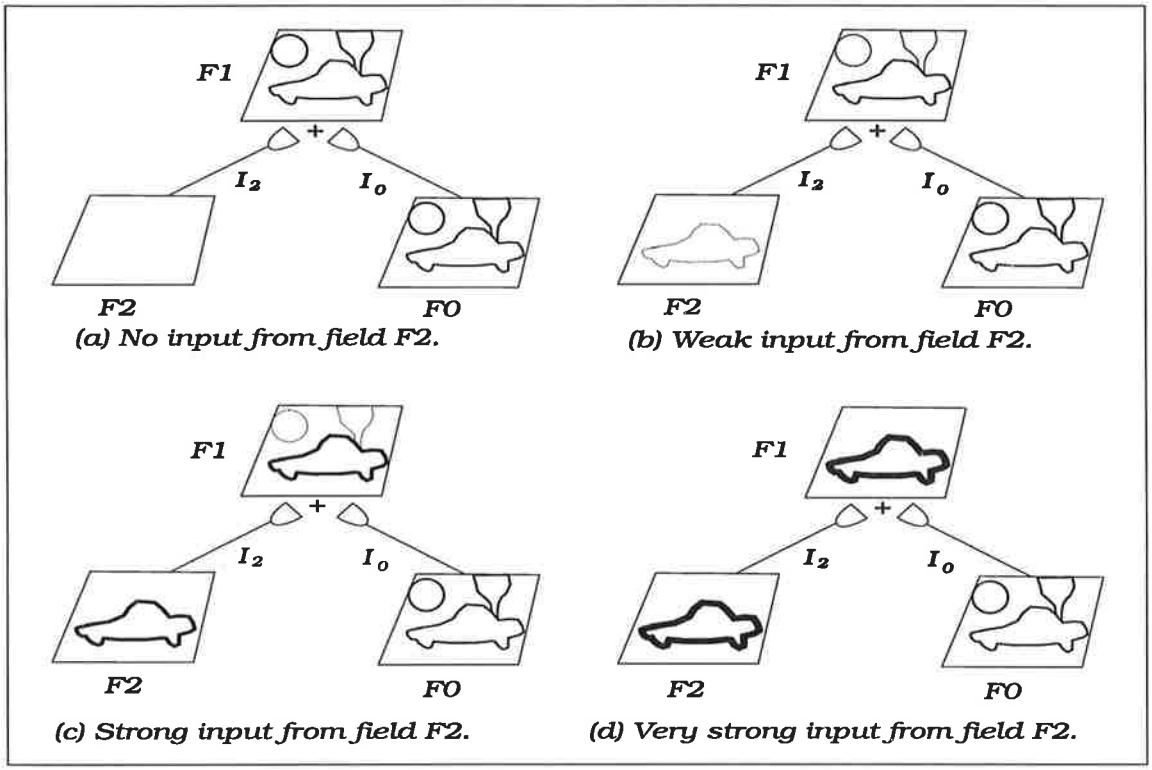


FIGURE 4.15. Cooperative combination of two 2-D shunting competitive neural layers.

Each of the shown Fields obeys the shunting competitive equation (4.24) of section 4.3. Since there are no inhibitory inputs in Figure 4.15, the cellular activity of each Field can be written as

$$\frac{dx_{ij}^{Fn}}{dt} = -Ax_{ij}^{Fn} + (B - x_{ij}^{Fn})Gv_{ij}^{Fn+} - x_{ij}^{Fn}(\overline{G}\overline{v}_{ij}^{Fn} + \Gamma_{ij}^{Fn}) \tag{4.33}$$

where the superscript Fn identifies the layer. Although we can write this equation in a 1-D form (i.e., by changing indices i,j to a single index, it is more convenient to analyse the behaviour of the network as a 2-D array of cells). For example, the cellular activity of Field F1 is given by

$$\frac{dx_{ij}^{F1}}{dt} = -Ax_{ij}^{F1} + (B - x_{ij}^{F1})G\left(v_{ij}^{F0F1+} + v_{ij}^{F2F1+}\right) - x_{ij}^{F1}(\overline{G}\overline{v}_{ij}^{F1} + \Gamma_{ij}^{F1}) \tag{4.34}$$

where v_{ij}^{F0F1+} is the excitatory postsynaptic potential due to the synaptic pathways $F0 \rightarrow F1$, while v_{ij}^{F2F1+} is the excitatory postsynaptic potential due to pathways $F2 \rightarrow F1$.

In the example shown, the strong input from Field F2 acts to suppress the non-matching inputs from F0. This can be achieved effectively only if the cells in field F1 are not initially activated by F0 at levels close to their saturation point and if the input from F2 is very much stronger than the input from F0. Otherwise, the addition of signals from source F2 will not have much effect on F1 since highly active cells in F1 cannot increase their activities much further in order to competitively annihilate the activity of cells whose input is only from source F0. Although the diagram shows that the resultant activity in F1 contains elements that are common to both F2 and F0, this need not be so. When the signal strength from F2 is much stronger than that from F0, the resultant activity in F1 will only represent source F2, regardless of what was initially transferred from F0. It is only when the magnitude of signals from both source Fields are approximately equal (as in examples 4.15(b) and (c)), that the activity in F1 will contain information from both sources.

In order to completely suppress the non-matching cellular activity in F1 by strong inputs from F2, it is desirable to design F1 to be more sensitive to the excitation from F2 than F0. In addition, the tonic level of transmitter mobilization in synaptic pathways $F0 \rightarrow F1$ should not be too high so as to excite the cells in Field F1 close to their saturation point. To emphasise the fact that the relative sensitivities may be different, we write equation (4.34) as

$$\frac{dx_{ij}^{F1}}{dt} = -Ax_{ij}^{F1} + (B - x_{ij}^{F1})\left(G_0 v_{ij}^{F0F1+} + G_2 v_{ij}^{F2F1+}\right) - x_{ij}^{F1}(\overline{G}\overline{v}_{ij}^{F1} + \Gamma_{ij}^{F1}) \tag{4.35}$$

where G_0 and G_2 are the constants (postsynaptic gains) that determine the relative sensitivities to the two input synapses. If $G_2 \gg G_0$, then Field F1 is more sensitive to inputs from F2 than it is to inputs from F0. The steady state solution to this equation is given by

$$x_{ij}^{F1} = \frac{B\left(G_0 v_{ij}^{F0F1+} + G_2 v_{ij}^{F2F1+}\right)}{A + (G_0 v_{ij}^{F0F1+} + G_2 v_{ij}^{F2F1+}) + (\overline{G}\overline{v}_{ij} + \Gamma_{ij}^{F1})} \tag{4.36}$$

Whereas in the above example we have assumed that both F0 and F2 provide excitatory inputs to F1 and hence interact cooperatively in F1 (i.e., both are involved in excitation of F1), below we consider a competitive interaction of F0 and F2 in F1 where the two input Fields provide inputs of opposite polarity. Figure 4.16 illustrates the result of competitive combination of the source two Fields, where one provides excitatory synaptic inputs while the other provides inhibitory synaptic inputs.

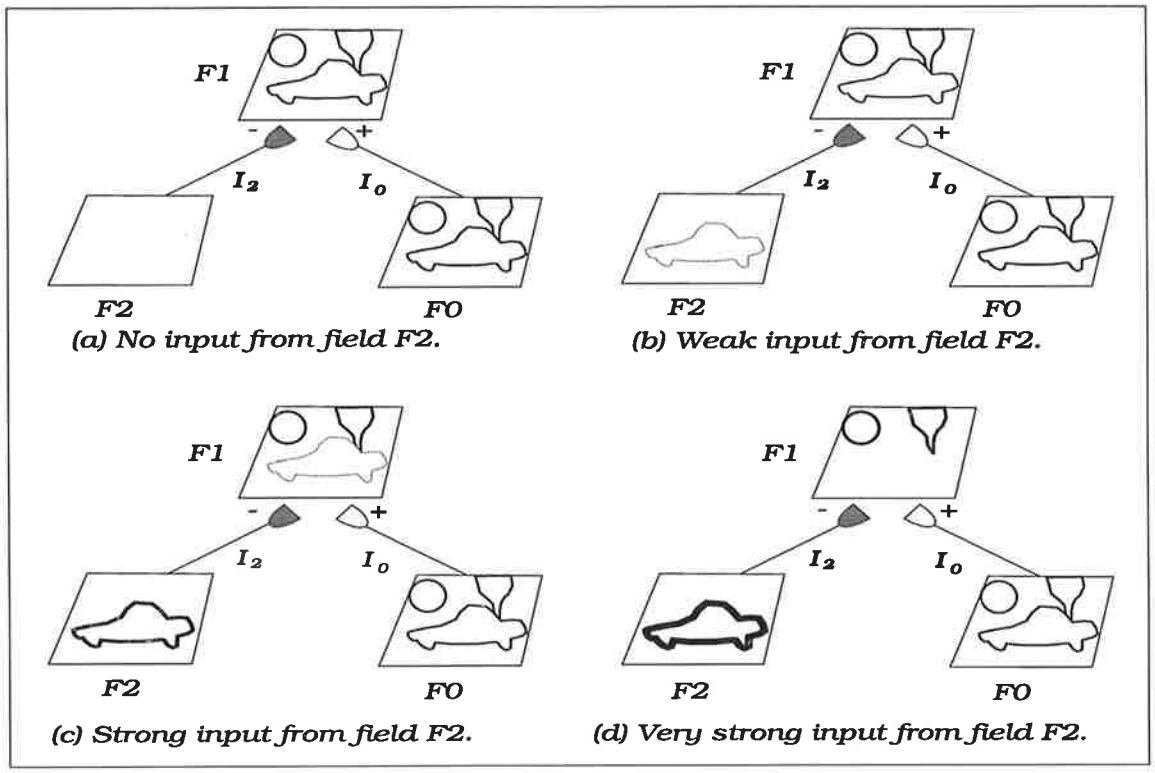


FIGURE 4.16. Competitive combination of two 2-D shunting competitive neural layers.

The dynamics of Field F1 is now described by the following equation

$$\frac{dx_{ij}^{F1}}{dt} = -Ax_{ij}^{F1} + (B - x_{ij}^{F1})\left[G_0 v_{ij}^{F0F1+} - G_2 v_{ij}^{F2F1-}\right]^+ - x_{ij}^{F1}\left(\left[G_2 v_{ij}^{F2F1-} - G_0 v_{ij}^{F0F1+}\right]^+ + \overline{G}\overline{v}_{ij}^{F1} + \Gamma_{ij}^{F1}\right) \tag{4.37}$$

whose equilibrium solution (for the case $G_0 = G_2 = G$) is given by

$$x_{ij}^{F1} = \frac{BG[v_{ij}^{F0F1+} - v_{ij}^{F2F1-}]^+}{A + G[v_{ij}^{F0F1+} - v_{ij}^{F2F1-}]^+ + G[v_{ij}^{F2F1-} - v_{ij}^{F0F1+}]^+ + \overline{G}\overline{v}_{ij}^{F1} + \Gamma_{ij}^{F1}} \tag{4.38}$$

As shown in Figure 4.16, Field F2 suppresses the cellular activity of Field F1 through its inhibitory synapse, but only in those cellular locations where F2 and F0 have common activity. The degree of suppression is dependant on the relative strengths of the corresponding activities of F2 and F1 (and the postsynaptic gains G_0 and G_2), as well as the levels of tonic transmitter mobilization. During the course of this suppression in F1, the total lateral inhibition on the other active cells in Field F1 will temporarily drop, forcing the Field to seek a new equilibrium state by increasing the activity of the other active cells. However, if Field F1 also receives excitatory inputs from other sources, then the resultant state will be determined by the combined effect of all the excitatory and inhibitory inputs.

In addition to the two modes of interactions between neural layers as shown in Figures 4.15 and 4.16, it is also possible for the modulatory signals to mediate inter-layer interactions by selectively facilitating or inhibiting synaptic pathways. In the following five sections we consider how presynaptic facilitation and presynaptic inhibition may be used in various combinations to provide a rich set of simple design principles and building blocks for two-dimensional neural circuits.

4.5 Synchronization of Oscillatory Neural Layers

The experimental data from stimulus evoked synchronized oscillations in the primary visual cortex (Eckhorn *et al.*, 1988; Gray and Singer, 1989; Gray *et al.*, 1989) has led to several neural network models of how such synchronized oscillations may arise. Although the existence of these synchronized oscillations or resonances and their role in the development of cortical feature-detectors was theoretically predicted by Grossberg (in the context of Adaptive Resonance Theory, Grossberg, 1976, 1978, 1982), well before they were detected in the visual system, it is only recently that neural network modelers began to propose models and architectures that could explain the mechanisms of coherent oscillatory behaviour in widely distributed cellular assemblies. The interest in oscillatory synchronization of feature detecting cells has also attracted the attention of applications oriented researchers since it is believed that such mechanisms may be useful in image texture segmentation. Below we first review two very different models of synchronized oscillations that were recently published and then we propose several alternative models that are based on the mechanisms of presynaptic modulation by lateral coupling.

Grossberg and Somers, (1991), have demonstrated robust synchronization in a 1-D layer of neurons with several different coupling architectures (Cooperative Bipole Coupling, Adaptive Filter Coupling, Nearest Neighbour Coupling and Random Coupling).

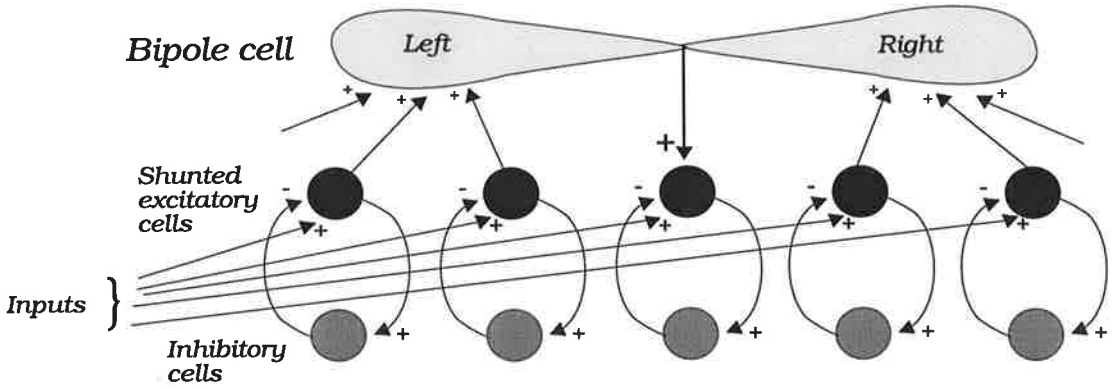


FIGURE 4.17 Synchronized oscillations via *bipole* coupling in the Grossberg and Somers model (1991).

$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)(C[x_i - \Gamma]^+ + \alpha C[Z_i - \Gamma]^+ + I_i) - Dx_i[y_i - \Gamma]^+ \tag{4.39}$$

and

$$\frac{dy_i}{dt} = -Ey_i + Fx_i \tag{4.40}$$

where α determines the size of the excitatory coupling with respect to the self-excitatory term $C[x_i - \Gamma]^+$ and Z_i is the activity of the ith coupling unit. For the cooperative bipole coupling architecture, Z_i is given by the following equations:

$$Z_i = \left[\frac{P(\text{Right}_i)^n}{Q^n + (\text{Right}_i)^n} + \frac{P(\text{Left}_i)^n}{Q^n + (\text{Left}_i)^n} - \Gamma_{cpl} \right]^+ \tag{4.41}$$

where

$$\text{Right}_i = \frac{1}{width} \sum_{j=1}^{width} [x_{i+j} - \Gamma]^+ \tag{4.42}$$

and

$$\text{Left}_i = \frac{1}{width} \sum_{j=-1}^{-width} [x_{i+j} - \Gamma]^+ \tag{4.43}$$

Parameter Γ_{cpl} is the coupling threshold. The other parameters (P, Q and n) represent a nonlinear summation within each cell compartment. For bipole coupling, both compartments of the *bipole cell* need to be sufficiently activated for Z_i to activate and thus provide the necessary excitatory coupling feedback. Since this coupling architecture forces coherent cellular synchronization to emerge as a result of direct cellular excitation by the feedback from the *bipole cell*, it is possible to induce oscillatory activity at cellular positions that do not correspond to any inputs. Other coupling architectures (such as the Adaptive Filter) do not introduce oscillations in cells whose inputs are zero. Although Grossberg and Somers (1991) were able to demonstrate robust synchronization with several different coupling architectures, their simulation data

shows that they have used one dimensional binary patterns. It is therefore not clear how their model would behave in a more general case when the inputs are of non-uniform intensity, such as what is typically obtained when a visual image is preprocessed by oriented contrast detectors (edge detectors). Since their model does not embed competition across the input layer and because the coupling feedback to the input neurons is positive, it can be concluded that those cells that receive maximum input will fire first, followed by neurons that receive the next highest excitation, etc. The model is therefore restricted to special cases and cannot be of much practical use in linking collinear edges whose magnitudes may differ widely.

The neural model of Eckhorn *et al.* (1990), whose fundamental component is schematized in Figure 4.18, is also of particular interest. Unlike the above model that uses excitatory coupling, the model by Eckhorn *et al.* (1990) shares some basic principles to the models that we are proposing below. This model not only provides means for synchronizing two interacting neural layers but it uses two functionally different synapses, referred to as *feeding synapses* and *linking synapses*. Their models of synapses are not as extensive as those discussed in this chapter and are simply represented by leaky integrators. However, the idea of *linking synapses* is similar in principle to our presynaptic facilitation. That is, the integrated signals from the *linking inputs* (together with a constant offset term) interact multiplicativelly with the integrated signals from the *feeding inputs*.

The membrane voltage $U_{m,k}(t)$ of the kth neuron is given by

$$U_{m,k}(t) = F_k(t)\,[1 + L_k(t)] \tag{4.44}$$

where $F_k(t)$ is the total contribution via the feeding inputs and $L_k(t)$ is the total contribution via the linking inputs. The neuron is modelled by a spike encoder whose output is high if the membrane potential is higher than a time variant threshold. As well as being able to synchronize neurons in one layer, this model is also able to synchronize two successive layers of neurons by the use of top-down feedback from the higher layer. However, as in the model of Grossberg and Somers (1991), the reported simulations have only demonstrated the model on binary inputs. Since there is no means in this model by which highly active neurons can be suppressed to compensate for weakly active neurons, the model cannot deal with inputs of non-uniform intensity.

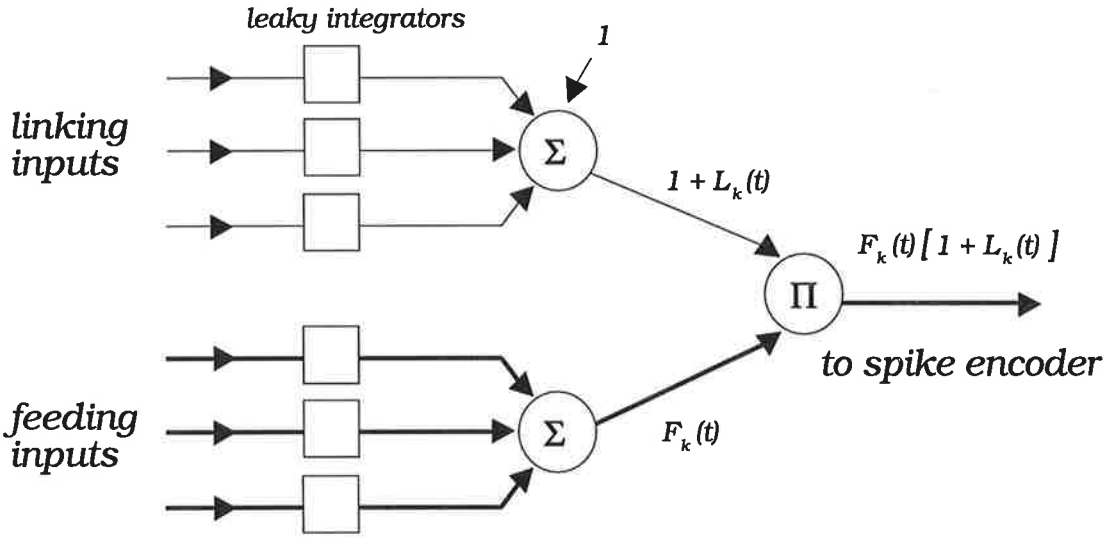


FIGURE 4.18. Synchronized oscillations via *linking* and *feeding* inputs in the model by Eckhorn *et al.*, (1990).

Below we propose three basic presynaptic mechanisms that may be used to synchronize cells within a presynaptically modulated shunting competitive neural layer, even when the spatial pattern feeding the layer is non-uniform in input strength. The basic mechanisms introduced here will be further extended to multiple neural layers in sections 4.8, while the neural circuit that is capable of detecting whether two interacting and pulsating neural layers are locked into synchronized oscillations (resonance) will be proposed in Chapter 6 (section 6.2).

4.5.1 Synchronization via Lateral Presynaptic Facilitation

The simplest presynaptically modulated shunting competitive neural layer is one whose input synapses are facilitated by the neighbouring cells within the layer via *lateral presynaptic facilitation*, shown in Figure 4.19. An oscillatory neural layer, coupled via lateral presynaptic facilitation has the interesting property of achieving synchronization of its cellular response when the layer is given random initial conditions and fixed inputs (random initial cellular activity $x_i(t_0)$ and random initial lateral inhibition $\overline{v}_i(t_0)$). Figure 4.20 shows the simulation results obtained with a layer whose cells are (a) uncoupled; and (b) coupled via nearest-neighbour presynaptic facilitation. For the case of nearest neighbour coupling all inputs were held fixed at equal magnitudes ($J_i = 1.0$) while the cellular activity and the lateral feedback inhibition were given random initial values.

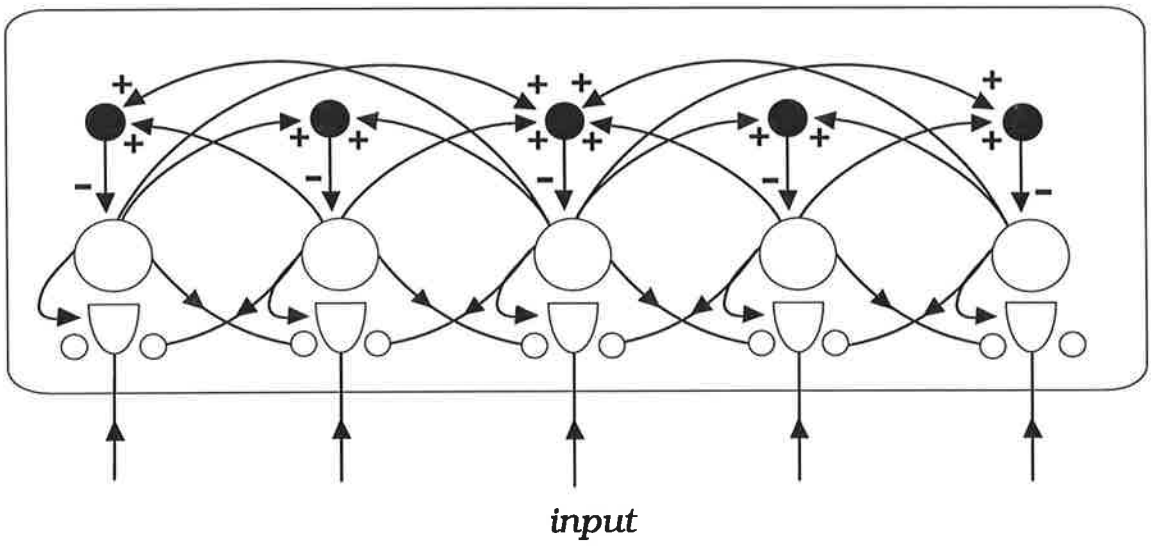


FIGURE 4.19. Transmitter gated shunting competitive neural layer coupled via *nearest neighbour lateral presynaptic facilitation*.

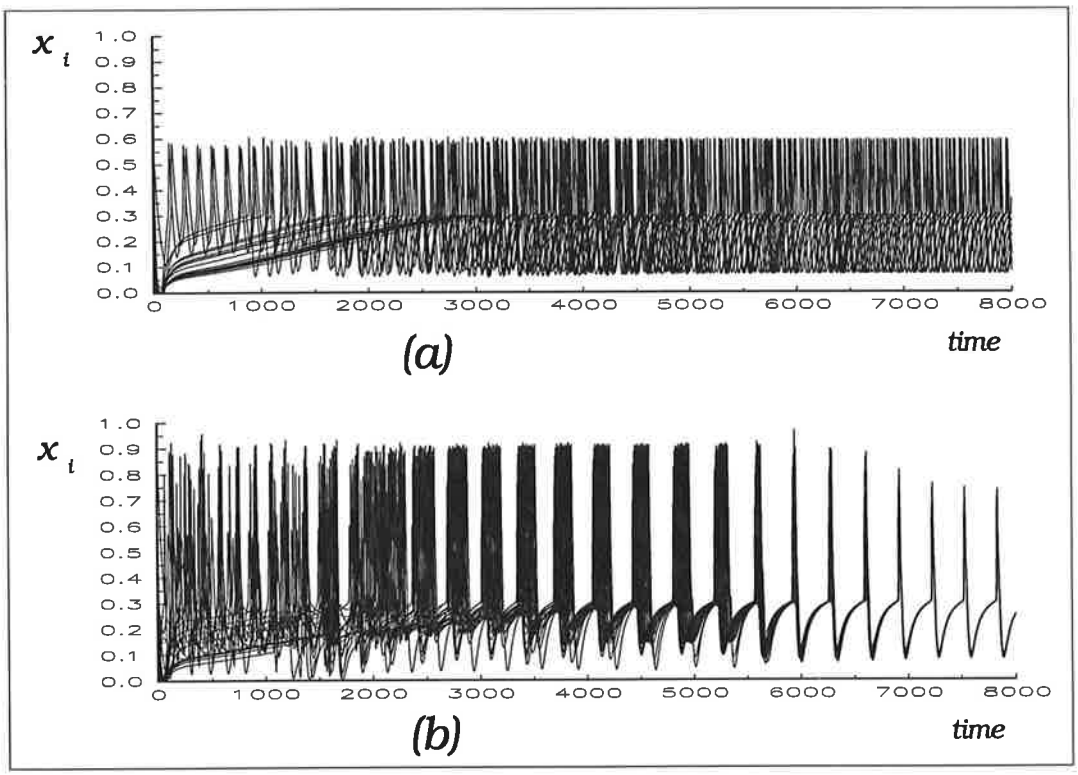


FIGURE 4.20. Simulation results of lateral presynaptic facilitation: (a) uncoupled case; and (b) coupled case. Layer equations and parameters are given in Appendix B.1.

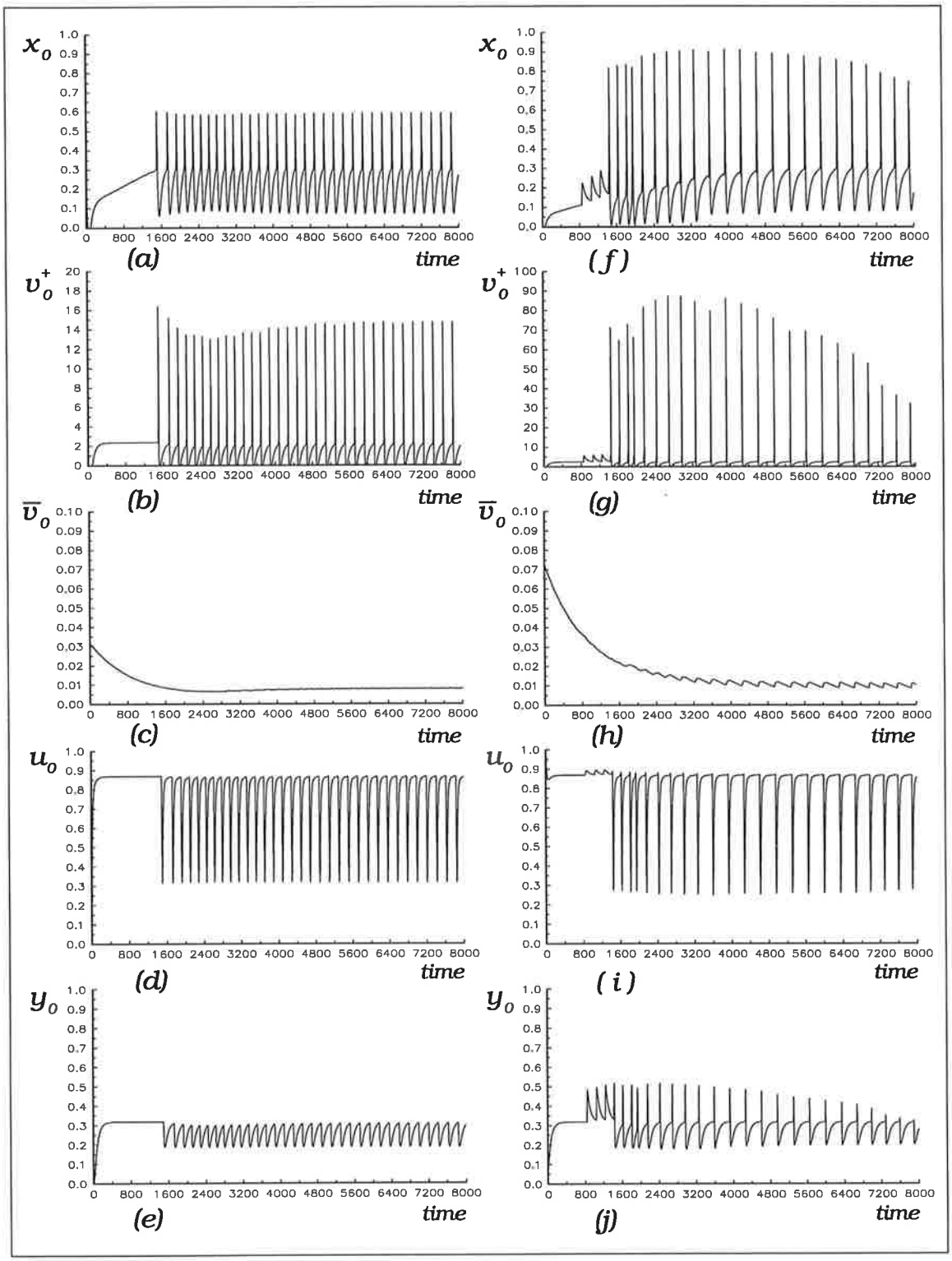


FIGURE 4.21. Dynamics of cellular and synaptic variables in a lateral nearest neighbour presynaptically facilitated neural layer: (a) - (e) uncoupled layer; (f) - (j) coupled layer.

The following equation describes the dynamics of the mobilized transmitter to each excitatory cell in the layer

$$\frac{dy_i}{dt} = [\beta_y + L_y c_{ji} f(x_j)](u_i - y_i) - J_i[y_i - Y]^+[\rho_y + K_y f(x_i)] - \gamma_y y_i \tag{4.45}$$

where $L_y c_{ji} f(x_j)(u_i - y_i)$ is the additional component to the transmitter mobilization process that is provided by the active neighbours; L_y is the gain of lateral presynaptic facilitation, while c_{ji} is the coupling coefficient. For the case of nearest neighbour coupling (as in the present case), c_{ji} is given by

$$c_{ji} = \begin{Bmatrix} 1 & if & \|j - i\| = 1 \\ 0 & otherwise & \end{Bmatrix} \tag{4.46}$$

Simulation results shown in Fig. 4.20 demonstrate that a layer of oscillatory neurons can effectively synchronize via nearest neighbour lateral presynaptic facilitation, but over a large number of cycles (> 20). The rate of synchronization depends on the lateral extent of the coupling neighbourhood. Synchronization is very rapid (within the first two cycles) when the coupling neighbourhood is comparable to the size of the layer. Figure 4.21 shows the dynamics of cellular and synaptic variables for one cell (cell 0) in the uncoupled and the coupled case.

Note that because the coupling mechanism of lateral presynaptic facilitation does not provide direct cellular excitation, it cannot induce oscillatory dynamics in cells whose synaptic inputs are zero (which may be required in some applications).

4.5.2 Synchronization via Later Presynaptic Excitation

In addition to the neural mechanism of lateral presynaptic facilitation, it is also possible to synchronize neurons within the layer via the neural mechanisms of *lateral presynaptic excitation*. In lateral presynaptic excitation, each cell within the layer provides excitatory inputs to its neighbour, but through the neighbour's input synapse. That is, each cell in a layer may cause the release of its neighbour's presynaptic transmitter. Below we consider two possible mechanisms by which an active cell may cause the release of its neighbour's excitatory synaptic transmitter.

(i) Lateral Presynaptic Correlated Transmitter Release

The postsynaptic feedback signal from an activated cell in the layer may multiplicatively interact with the synaptic input of its neighbour and cause transmitter release under the condition of correlated firing. Figure 4.22 shows a layer of competing neurons that can be synchronized via the correlated firing of the synaptic signal and the nearest neighbour coupling. This interactions is represented by the following equation.

$$\frac{dv_i}{dt} = -Dv_i + J_i[y_i - Y]^+ \left\{ \rho_v + K_v f(x_i) + L_v^c \sum_{j \neq i}^{n} c_{ji} f(x_j) \right\} \tag{4.47}$$

where L_v^c is the gain of lateral correlated transmitter release while c_{ji} is a coefficient that defines the coupling neighbourhood. In a one dimensional layer of competing neurons, where the coupling is between the nearest neighbours, the coupling coefficient is also specified by (4.46).

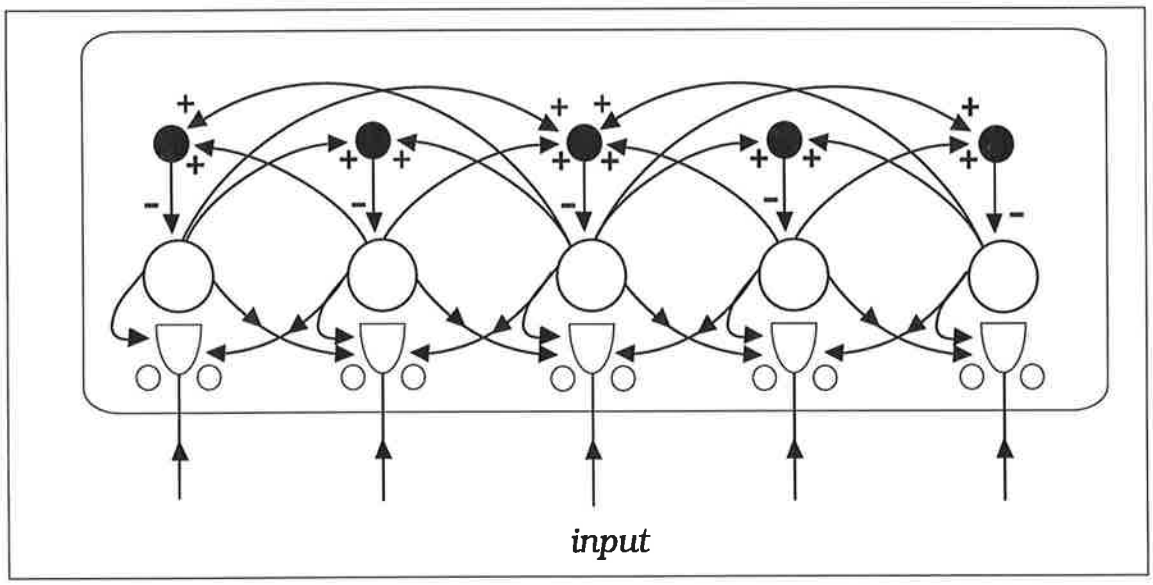


FIGURE 4.22. Shunting competitive neural layer coupled via *nearest neighbour lateral presynaptic excitation and correlated transmitter release.*

Simulation results for a one dimensional layer of 20 competing neurons is shown in Figure 4.23. This simulation shows that for the same size of coupling neighbourhood, coupling via correlated lateral presynaptic excitation is more efficient in synchronizing the layer than the mechanism of lateral presynaptic facilitation (i.e., the layer becomes synchronized over a smaller number of cycles when the coupling is

via correlated lateral presynaptic excitation). Correlated presynaptic excitation also provides more robust means of synchronizing a layer of competing neurons when their synaptic inputs are non-uniform in strength.

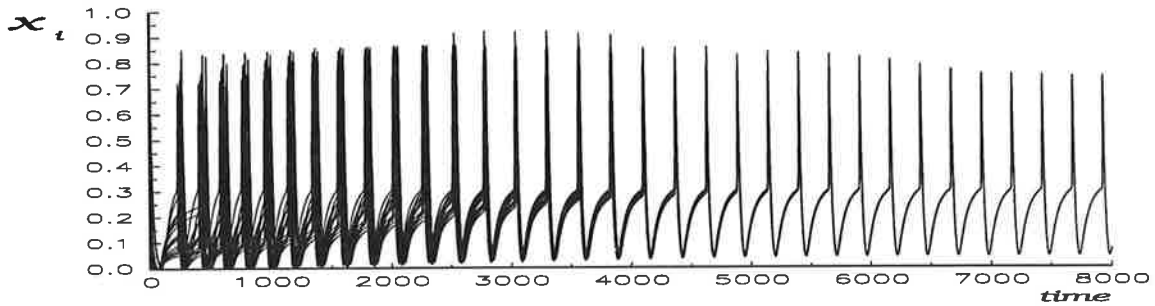


FIGURE 4.23. Simulation results of synchronized oscillations via nearest neighbour correlated transmitter release. Layer equations and parameters are given in Appendix B.1.

Figure 4.24 below shows an input signal waveform that is used in the simulation whose results are shown in Figure 4.25. The coupling neighbourhood in this simulation was extended to include all the cells in the layer.

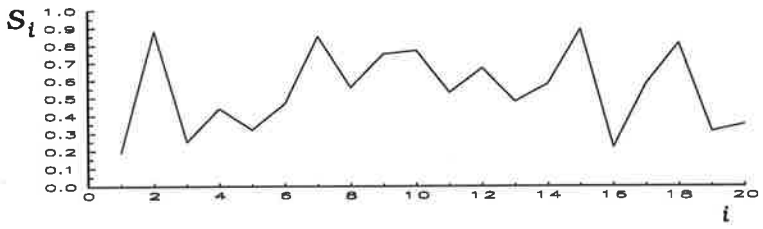


FIGURE 4.24. Non-uniform synaptic input into a layer of competing neurons coupled via correlated lateral presynaptic excitation.

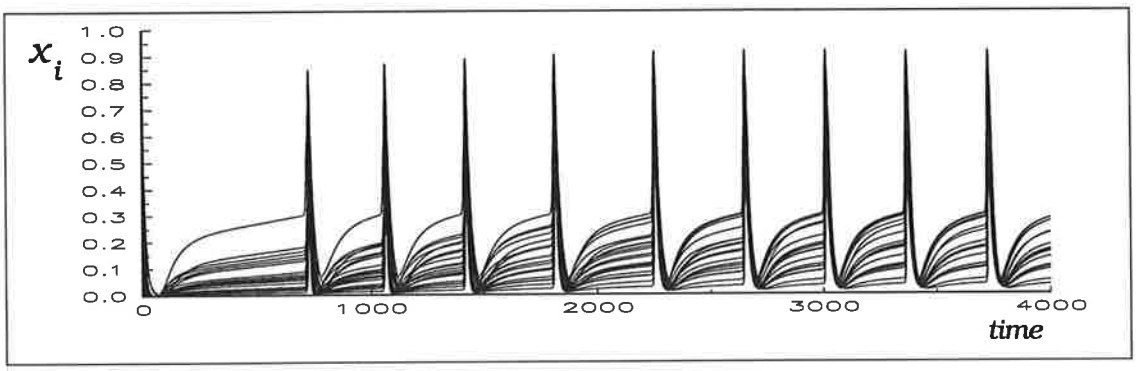


FIGURE 4.25. Synchronized oscillation in a competitive neural layer whose inputs are non-uniform and whose coupling neighbourhood includes all the neurons in the layer.

The above simulation data (Figure 4.25) shows that when the coupling neighbourhood is increased, then the layer can be quickly synchronized even when the synaptic input signal is non-uniform in strength. However, only those neurons whose bottom-up synaptic input strength is large enough to drive the cell above its firing threshold will be synchronized. Below we present an alternate synchronization method that is effective in firing and synchronizing a selected portion of neurons that may not receive any synaptic inputs other than via lateral interactions.

(ii) Lateral Presynaptic Excitation via Direct Transmitter Release

The third coupling method that can synchronize a layer of competing neurons is where each active cell in the layer is directly engaged in the release of its neighbour's presynaptic transmitter, i.e., via *direct lateral presynaptic excitation*. Figure 4.26 represents a layer of shunting competitive neurons that are coupled via direct lateral presynaptic excitation. To expose the capability of this synchronization method, each alternate input was set to zero. Thus half of the input synapses are driven by a non-zero input signal (equal to 1) while the other half are driven only by their neighbours. For the case of nearest neighbour coupling, the following equation describes the dynamics of the excitatory postsynaptic potential acting on each cell.

$$\frac{dv_i}{dt} = -Dv_i + [y_i - Y]^+ \left\{ \rho_v J_i + K_v J_i f(x_i) + L_v^d \sum_{j \neq i}^{n} c_{ji} f(x_j) \right\} \tag{4.48}$$

where L_v^d is the gain of the direct lateral transmitter release, c_{ji} is the coupling coefficient, defined as in (4.46). Simulation results, shown in Figure 4.27, are plotted for each of the two groups of neurons.

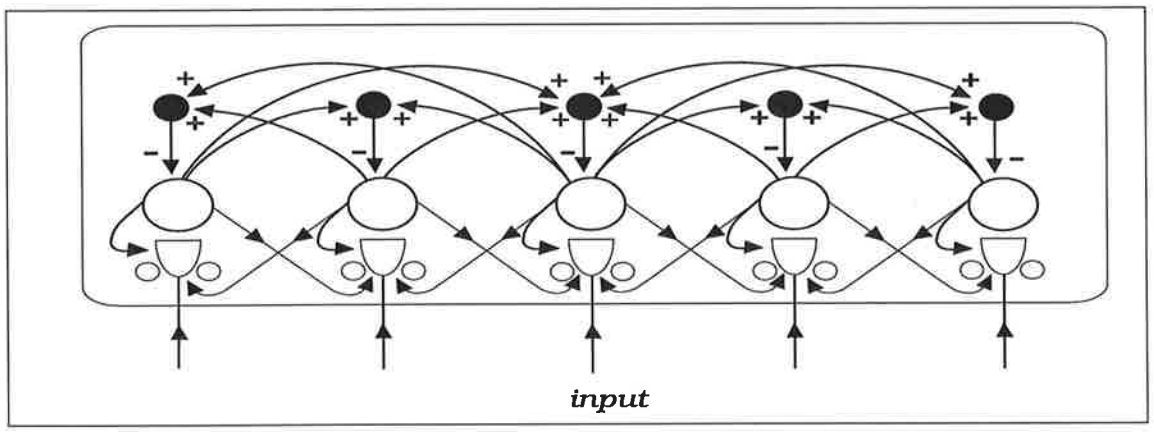


FIGURE 4.26. Shunting competitive neural layer coupled via *direct lateral presynaptic excitation*.

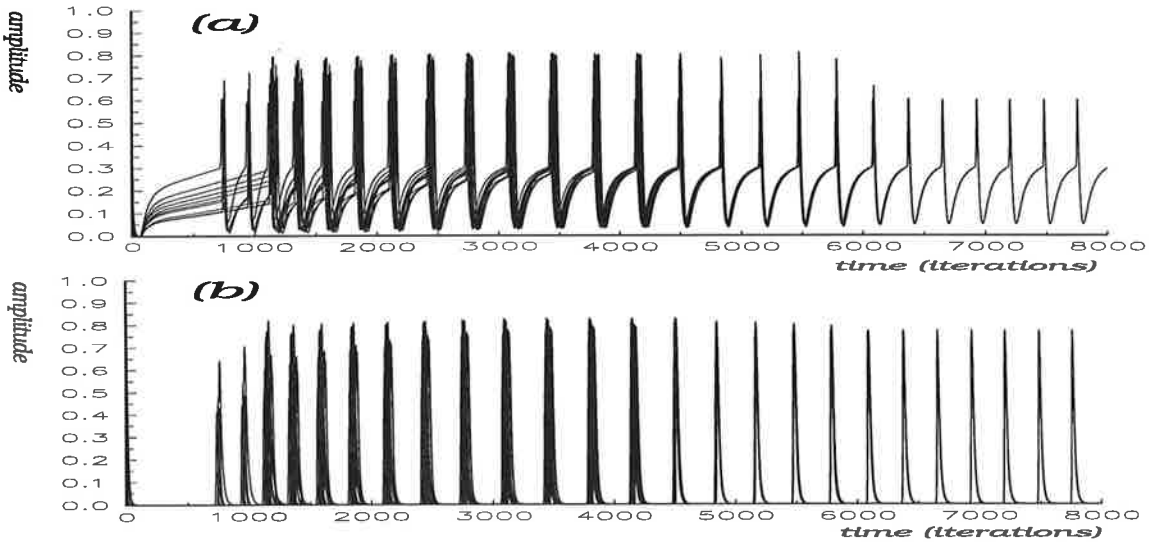


FIGURE 4.27. Simulation results of synchronized oscillations in a competitive neural layer coupled via direct lateral presynaptic excitation and transmitter release: (a) response of the neurons that receive a steady synaptic input signal; (b) response of the neurons that receive only the presynaptic lateral excitation. Layer equations and parameters are given in Appendix B.1.

The above simulation demonstrates that synchronized oscillations via direct lateral presynaptic excitation is a very effective method of synchronizing a layer of shunting competitive neurons, even if some of the neurons in the layer do not receive any direct synaptic inputs. This form of synchronization is very useful in applications where it is desired to generate a neural pattern of activity in a sub-population of cells that do not receive a direct synaptic input but need to be activated in order to match other system requirements. A good example of this application is in the Boundary Contour System (Grossberg and Mingolla, 1985a,b) where, under certain input conditions, illusionary 2-D contours may be generated.

In all the above simulations of synchronized oscillations, we have assumed that the lateral coupling coefficient is distance independent. However, in general applications to 2-D visual information processing problems, synchronization may have to be achieved across distant cells and perhaps across several interacting neural layers. The coupling coefficient should thus be distance dependant as well as orientation dependant. The neural mechanism of presynaptic modulation may thus be also used in modulating the extent and the orientation of the coupling neighbourhood. That is, since the cells in the visual cortex are generally sensitive to some optimum spatial frequency and orientation, cells of like orientation may have to be cooperatively coupled more strongly than cells whose optimum orientation sensitivity differs (as in the BCS model of Grossberg and Mingolla). However, unlike the fixed receptive field of the theoretical *bipole cell*, which is used in the CC Loop of the BCS neural network model

of preattentive vision to cooperatively link perceptual features into emergent boundary segmentation, we propose that presynaptic modulation be used as means of modulating its receptive field profile. Presynaptic neural mechanisms for the modulation of cellular receptive fields are described in section 4.10.

4.6 Lateral Presynaptic Facilitation in Non-oscillatory 2-D Neural Layers

Presynaptic modulation has thus proven to be an extremely useful neural mechanism in synchronizing a layer of oscillatory neurons and is likely to play a significant role in many other applications. Below we present the simulation data of cooperative linking in a two dimensional non-oscillatory presynaptically modulated shunting competitive neural layer whose cells are cooperatively linked to their nearest neighbours, via several coupling topologies shown in Figure 4.28.

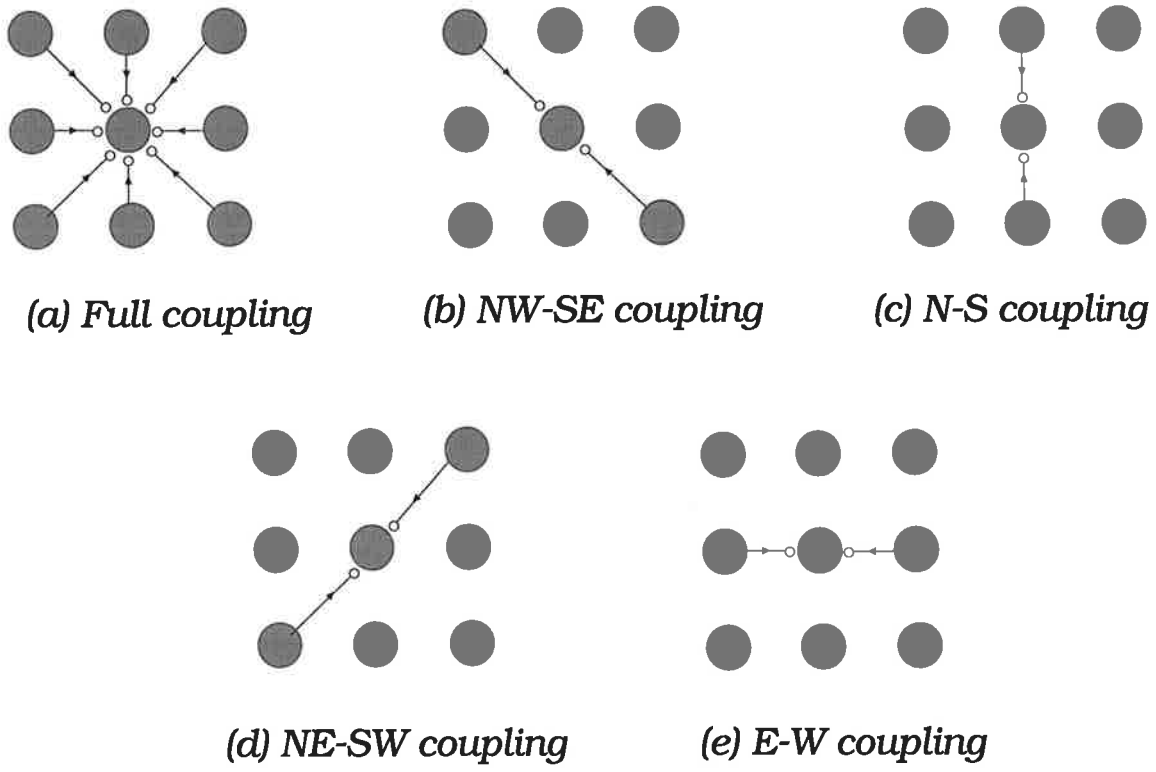
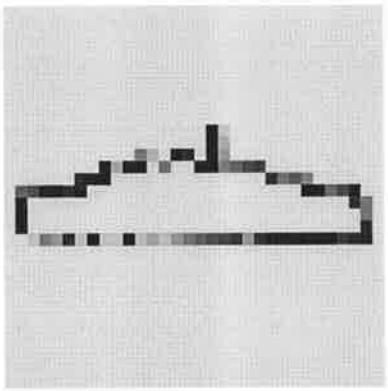
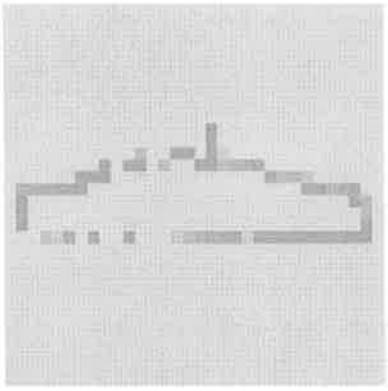


FIGURE 4.28. Simple coupling topologies for laterally presynaptically modulated shunting competitive 2-D neural layers.

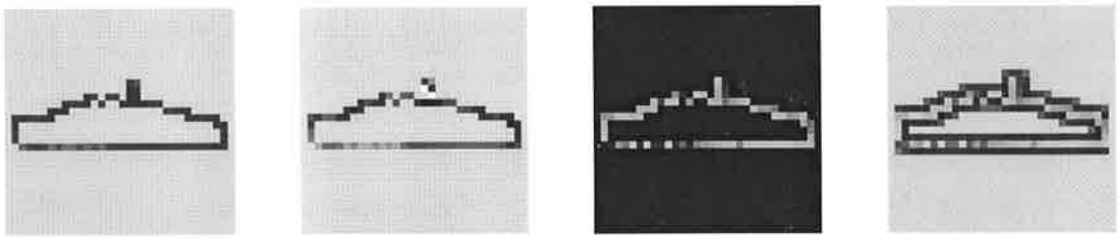


(a) Test input



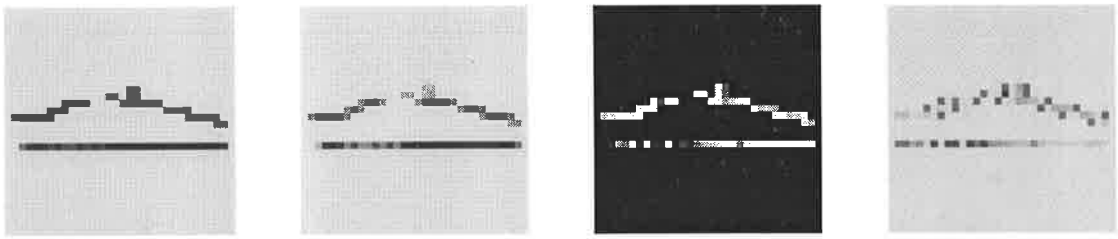
(b) Steady state results for an uncoupled layer

FIGURE 4.29. Test input for 2-D cooperative coupling in a presynaptically modulated shunting competitive neural layer.



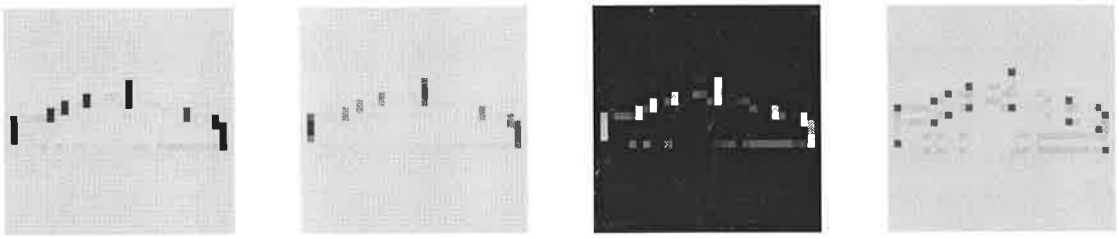
$\hat{x}_{ij}$ $\hat{v}_{ij}$ $\hat{u}_{ij}$ $\hat{y}_{ij}$

FIGURE 4.30. Steady state simulation results with the full coupling topology.



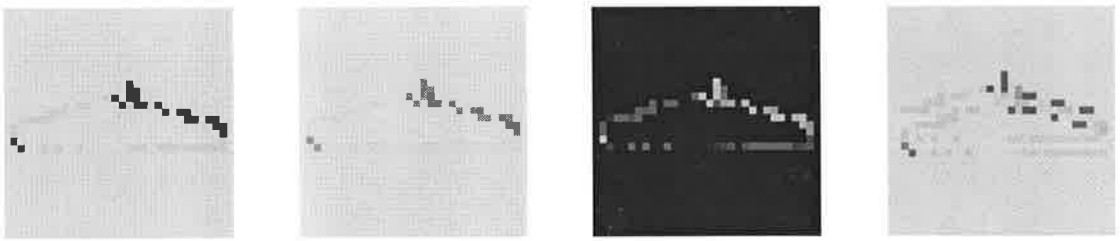
$\hat{x}_{ij}$ $\hat{v}_{ij}$ $\hat{u}_{ij}$ $\hat{y}_{ij}$

FIGURE 4.31. Steady state simulation results with the E-W coupling topology.



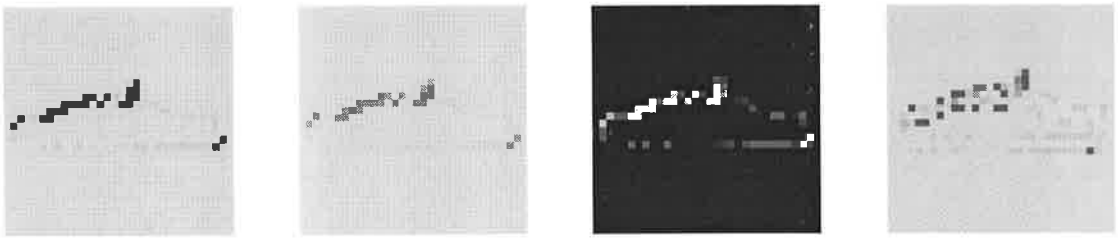
$\hat{x}_{ij}$ $\hat{v}_{ij}$ $\hat{u}_{ij}$ $\hat{y}_{ij}$

FIGURE 4.32. Steady state simulation results with the N-S coupling topology.



$\hat{x}_{ij}$ $\hat{v}_{ij}$ $\hat{u}_{ij}$ $\hat{y}_{ij}$

FIGURE 4.33. Steady state simulation results with the NW-SE coupling topology.



$\hat{x}_{ij}$ $\hat{v}_{ij}$ $\hat{u}_{ij}$ $\hat{y}_{ij}$

FIGURE 4.34. Steady state simulation results with the NE-SW coupling topology.

The above simulation data shows that the laterally presynaptic coupling topology can be very effective in amplifying weak inputs of certain cells if their coupling neighbours are strongly active. Thus even when some inputs are much smaller than others by a factor of 10 and would therefore not enter into the circuit (as exemplified above in Figure 4.29(b)), their synaptic signal transmission gain can be increased to compensate for this. Below we provide simulation data that was obtained from edge processed images of several objects.

The simulation data on more realistic inputs shows that the simple coupling topology considered here is not sufficiently robust to be of much practical use. The problem is due to the fact that we have lumped all input edges into one array and have thus ignored the edge orientation. Since the competitive interactions in the simulated neural layer are distance independent and cover the whole extent of the layer, the first output that begins to form via cooperative coupling are those from an input region with lots of strong input signals. Therefore, a weak input edge that does not have many strong neighbours is unlikely to enter into the circuit. Nevertheless, this simple simulation supports the view that image edges should be distributed into different edge arrays such that edge magnitudes within each array share common spatial orientation. The experimental data from the cat and monkey visual cortex does suggest that the early visual filters (edge and contrast detectors) are in fact arranged in edge maps of similar orientation.

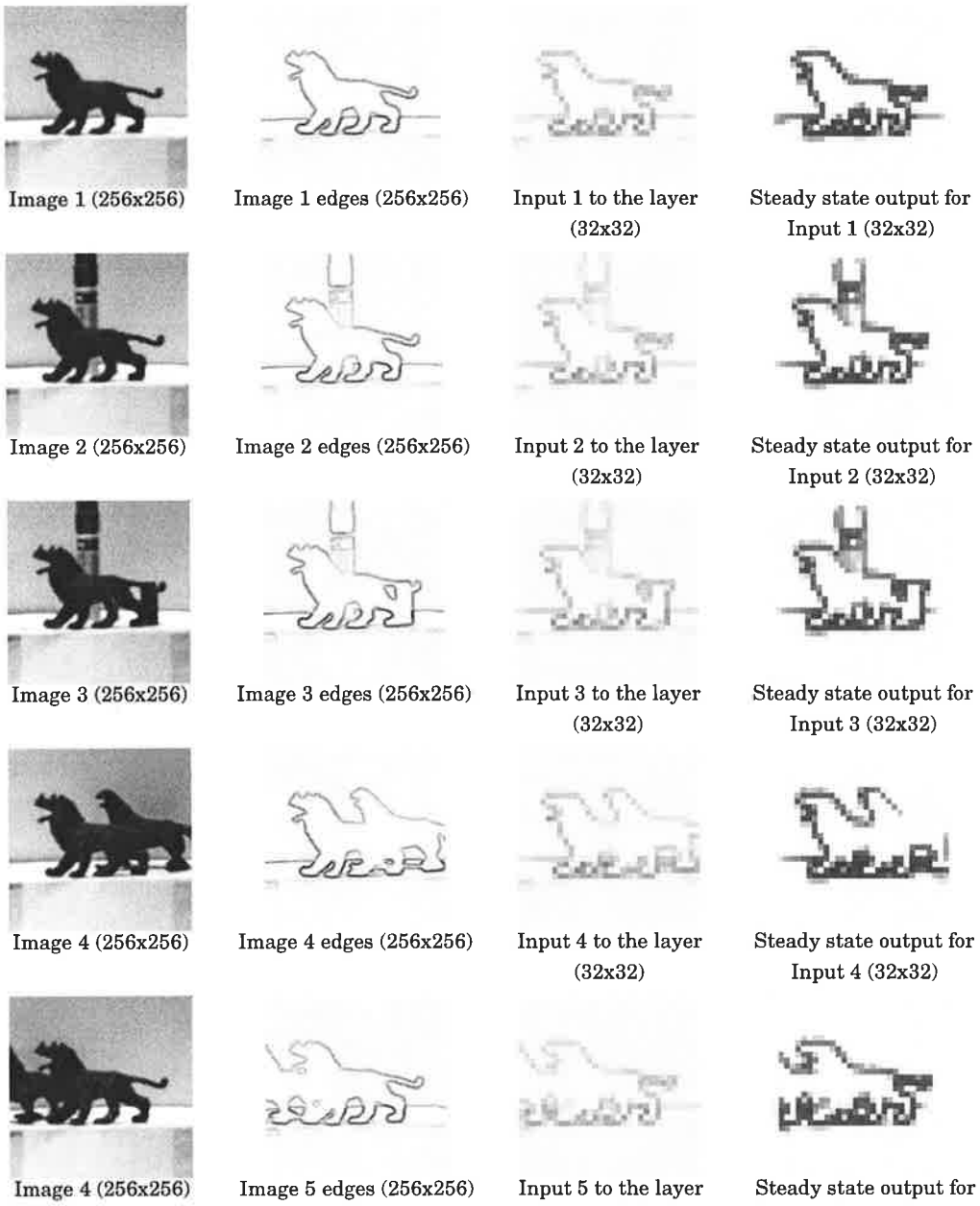


FIGURE 4.35. Steady state simulation results of the full coupling topology when applied to edge processed images of objects.

In the next section we consider how presynaptic facilitation and presynaptic inhibition may be used to selectively transfer 2-D spatial patterns from one neural layer to another.

4.7 Selective Information Transfer

We now propose how the presynaptic facilitatory/inhibitory gain control signals can be used to selectively modulate the transfer of two dimensional spatial patterns of neural activity between several transmitter gated and presynaptically modulated

shunting competitive neural layers. In all the examples discussed below, it is assumed that the common 2-D neural patterns of activities across each layers are spatially aligned in all respects (position, size and 2-D orientation). Generalization to multiple size, orientation and position will be presented in subsequent chapters.

To expose the new properties of PM-SCNLs, we will assume that each layer is modelled by its steady state. This enables us to represent the neural patterns of activity by drawing various 2-D shapes within a given layer (such that the thickness of a shape in a particular spatial location denotes the strength of the steady state cellular activity). Modelling a neural layer by its steady state also enables us to achieve faster simulation times and to represent the simulation results as a spatial pattern.

Figure 4.36 depicts three neural Fields whose steady state cellular activity in various spatial locations is represented by the drawn shapes within each 2-D layer. Fields F0 and F2 receive inputs I_0 and I_2 respectively, such that I_2 is a subset of I_0 (i.e., Fields F0 and F2 share common activity). If we assume that Field F0 provides excitation to Field F1, then there are two ways that signals from F2 can influence the neural signal transmission in the pathway $F0 \rightarrow F1$. That is, F2 can either facilitate or inhibit this transmission. If we allow the signals from F2 to facilitate the signal transmission from Field F0 to Field F1, then the resultant activity in F1 will be determined by the common spatial pattern of activity between F0 and F2 (the car boundary). If on the other hand F2 has an inhibitory effect on the presynaptic pathways $F0 \rightarrow F1$, then what results in F1 is the activity of F0 minus the activity of F2 (i.e., the neural pattern of activity that is in $F0$ but not in $F2$).

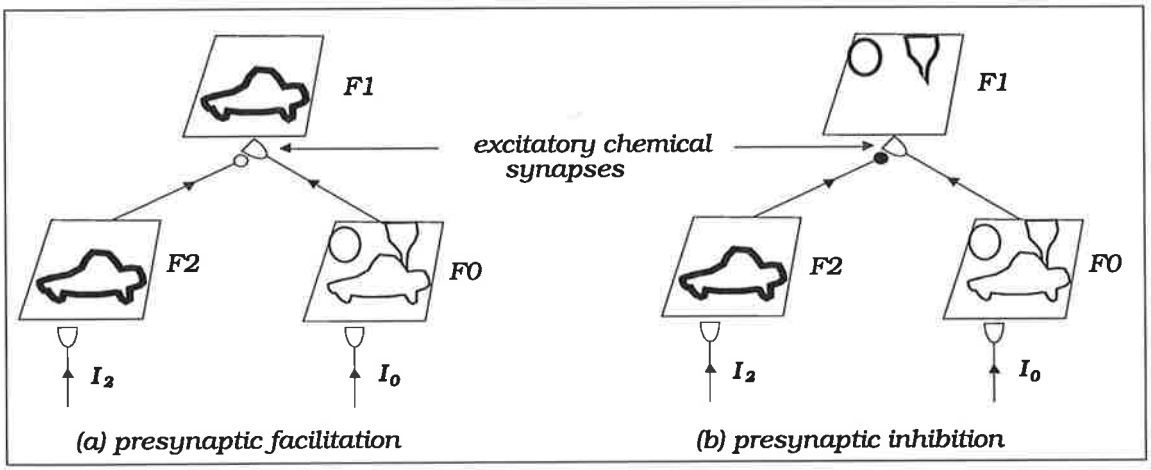


FIGURE 4.36. Presynaptic facilitation and inhibition in selective transfer of neural signals.

Unless otherwise stated, empty circles abutting another synapse (as in the above diagram) will represent facilitatory presynaptic modulation, while the black-filled circles will represent the inhibitory modulation of synapses. Similarly, empty semi-ellipses will represent excitatory synapses whereas shaded synapses will denote inhibitory synapses.

This facilitatory/inhibitory modulation of presynaptic signals is a useful form of selective transmission of neural information between various layers. The above two examples illustrate only some very simple but powerful neural design principles that we can use in order to *neuro-engineer* a neural network model with ability to selectively attend to portions of the bottom-up neural activity or the sensory input. Figure 4.37 shows a one dimensional schematic of the simplest 2-D neural circuit capable of selectively transferring neural signals in the synaptic pathways $F0 \rightarrow F1$. Simulation result in Figure 4.38 show the cellular activities of the two layers and the synaptic variables $(u_{ij}^{F0F1}, y_{ij}^{F0F1})$ in the $F0 \rightarrow F1$ pathway (for the first 14 iterations of the network). For general layer design and parameter selection, refer to Chapter 5.

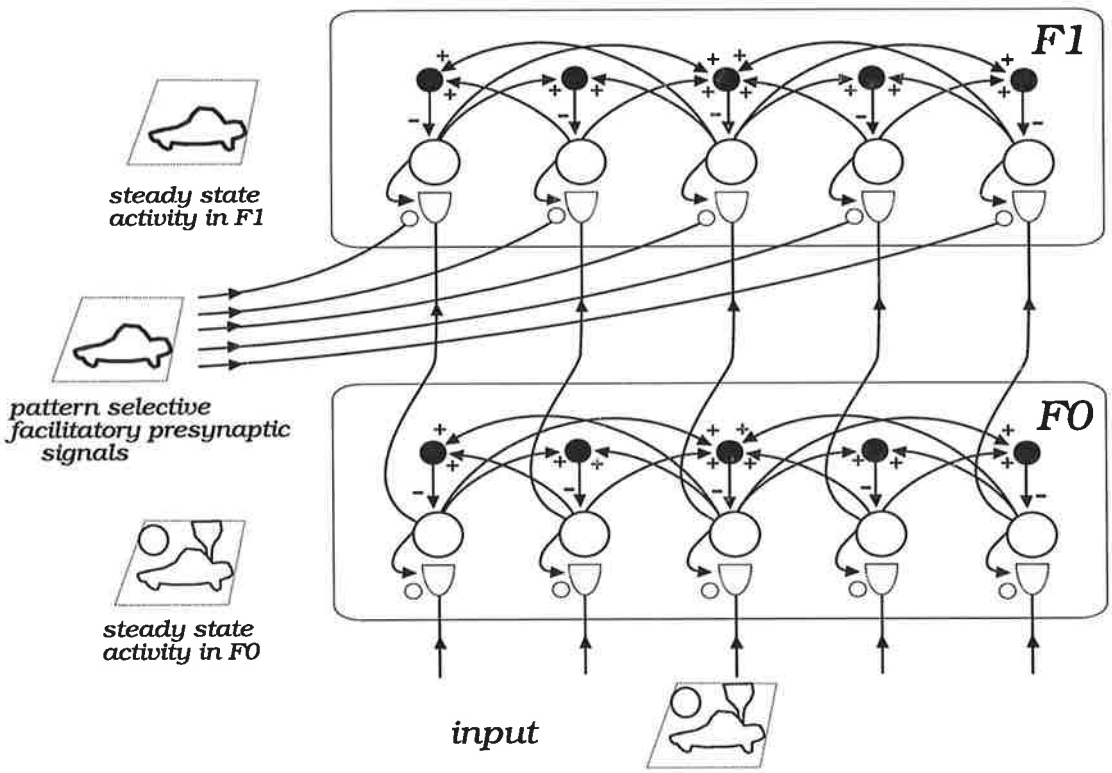


FIGURE 4.37. Neural circuit for pattern specific presynaptic facilitation of information transfer.

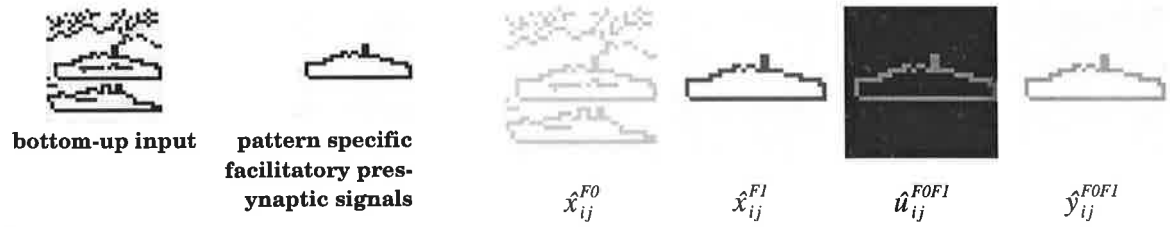


FIGURE 4.38. Steady state simulation results of pattern selective information transfer by presynaptic facilitation. The leftmost columns shows the bottom-up input to Field F0 and the pattern selective facilitatory presynaptic signals to Field F1. For layer design and parameter details refer to Chapter 5.

Simulation results in Figure 4.38 show that the pattern selective (or pattern specific) form of facilitatory presynaptic facilitation is very effective in transferring the desired bottom-up neural pattern of activity between two layers of shunting competitive neurons. Of particular interest is the effect on the input synapses of Field F1 (variables u_{ij}^{F0F1} and y_{ij}^{F0F1}) and the resultant steady state activity of Field F1 (x_{ij}^{F1}). The data shows that the resultant steady state neural pattern of activity across F1 contains only a specific (selected) portion of all the bottom-up inputs impinging F1 from F0.

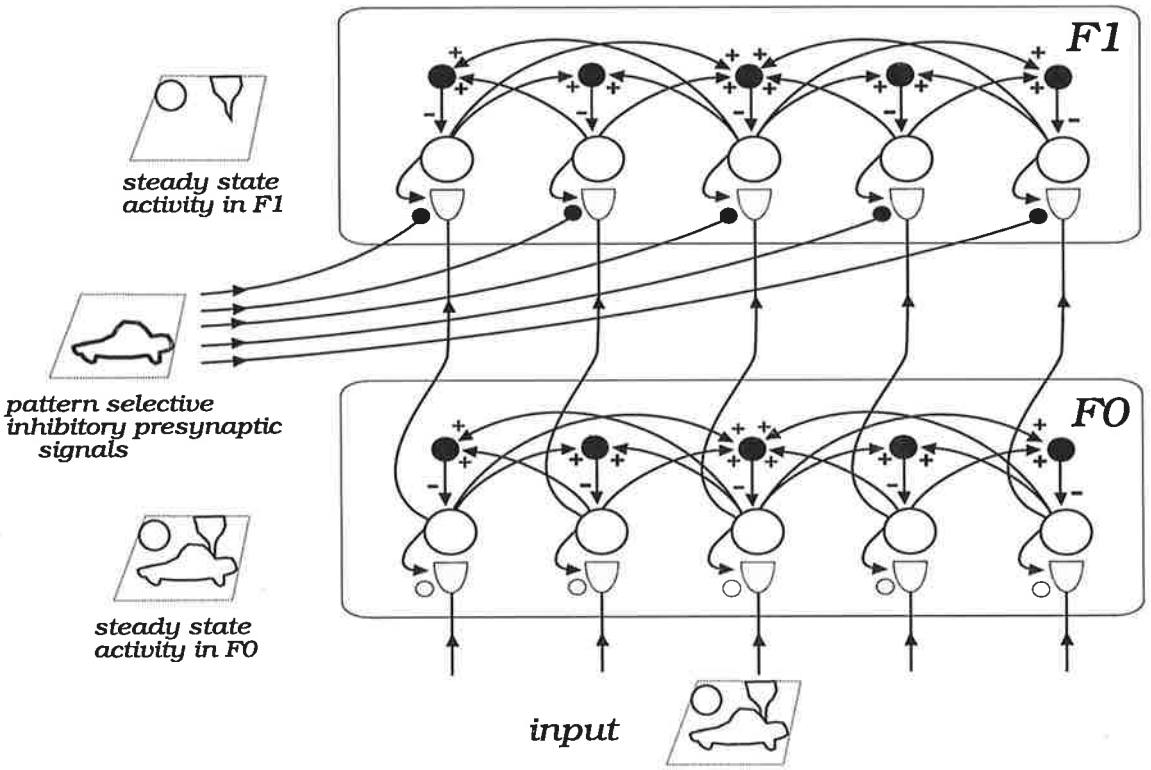


FIGURE 4.39. Neural circuit for pattern selective presynaptic inhibition of neural information transfer.

Figure 4.39 shows a complimentary neural circuit that inhibits a selected pattern from being transferred to Field F1, while Figure 4.40 shows the simulation results for the same bottom-up input as in the above example.

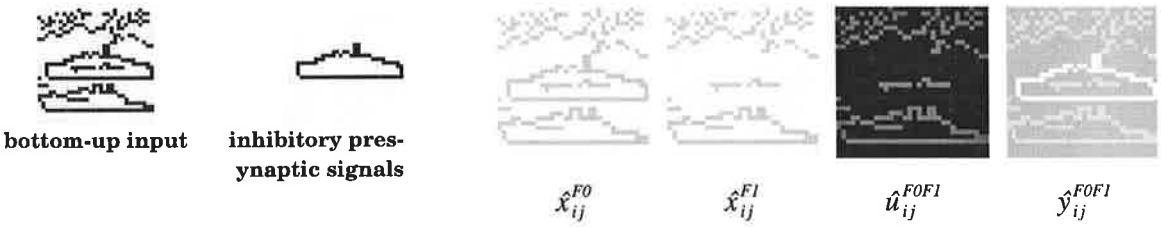


FIGURE 4.40. Simulation results of pattern selective inhibition of information transfer.

Note the absence of a portion of the bottom-up input that appears across F0 but not across F1 (i.e., the boundary of a ship that appears in the middle of the input array). The missing portion was eliminated from F1 by the inhibitory presynaptic signals that have effectively reduced the gain of signal transmission in shown locations. The above two examples demonstrate only the pattern-selective (pattern specific) nature of selective information transfer via presynaptic modulation. In addition to the pattern-selective transfer of 2-D neural signals, it is also possible for the modulatory gain control signals to be spatially-selective. Figures 4.41 and 4.43 show the neural circuits for spatially selective transfer of 2-D neural signals, while Figures 4.42 and 4.44 show their respective steady state simulation results.

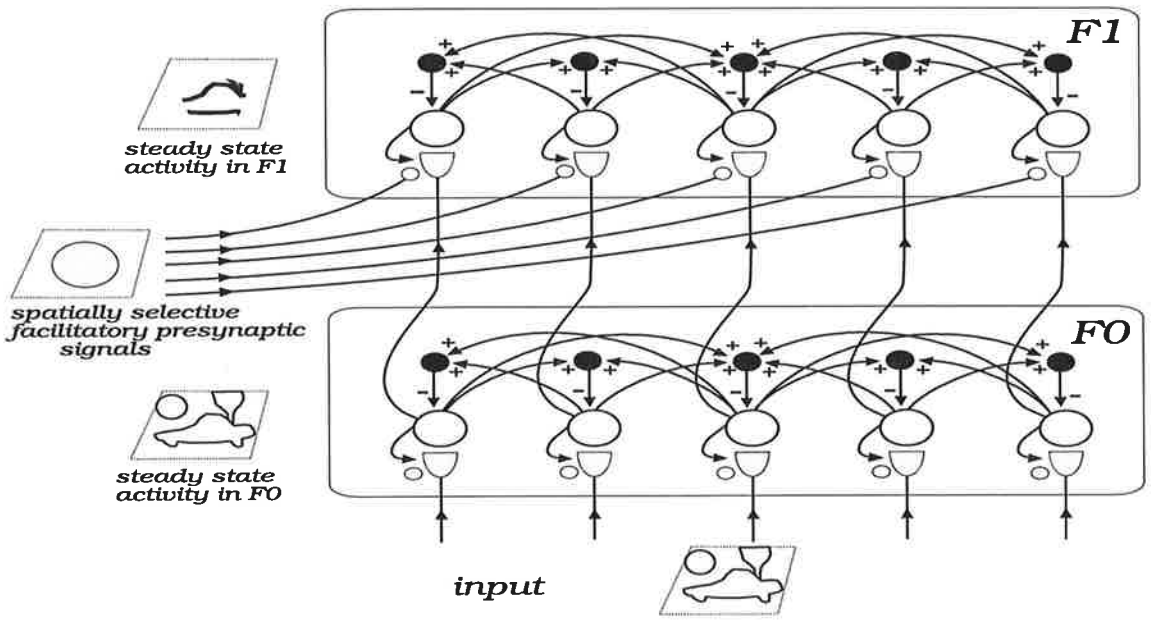


FIGURE 4.41. Neural circuit for spatially selective presynaptic facilitation of neural information transfer.

FIGURE 4.42. Simulation results of spatially selective information transfer by presynaptic inhibition. (For parameter design refer to Chapter 5).

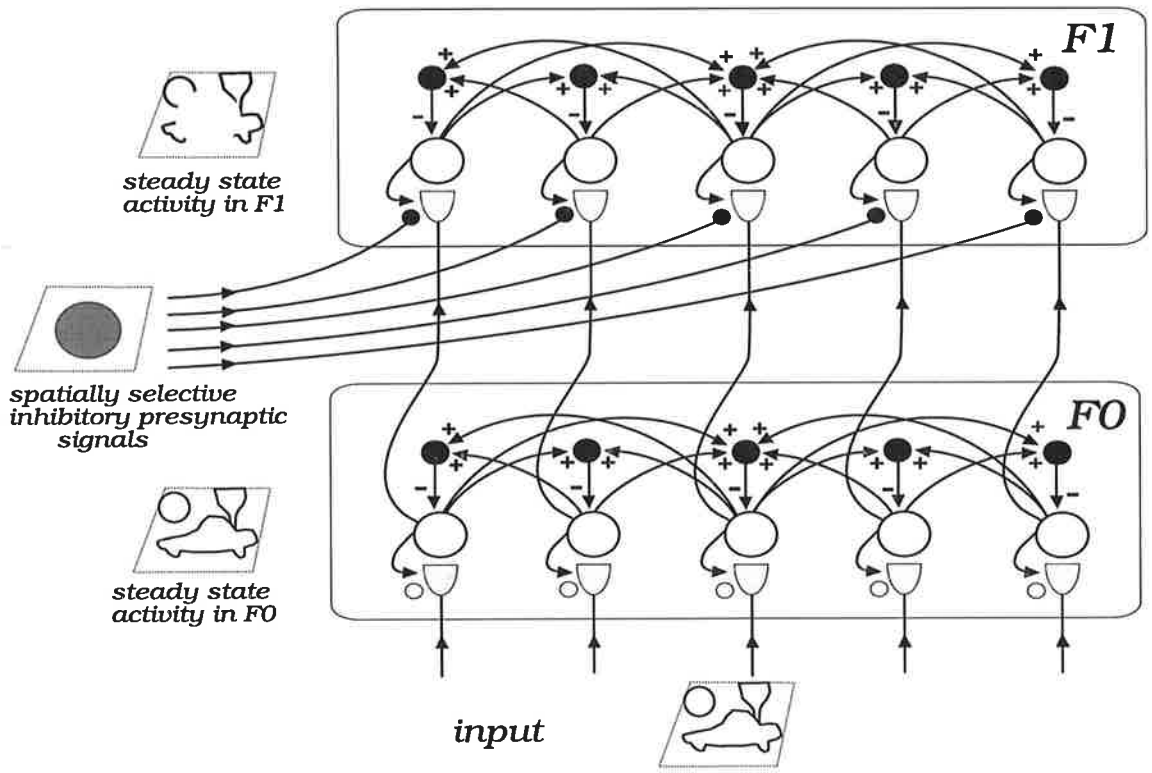


FIGURE 4.43. Neural circuit for spatially selective presynaptic inhibition of information transfer.

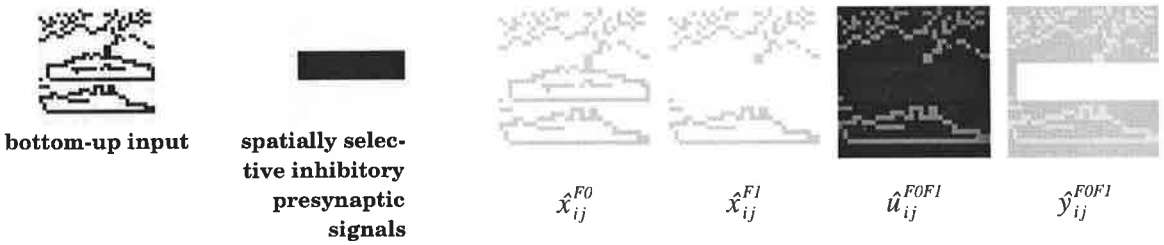


FIGURE 4.44. Simulation results of spatially selective information transfer by presynaptic inhibition. (For parameter design refer to Chapter 5).

In the above four examples we have considered only the case when Field *F0* provides excitatory synaptic inputs to Field *F1*. In Figure 4.45 we consider a more general case where Field *F1* may receive both excitatory and inhibitory synaptic inputs, each of which may be modulated by separate signals and each of which may be either pattern or spatially selective. To expose the dynamics of modulated inhibitory synapses, assume that Field *F1* also receives excitatory synaptic inputs from Field *F3*.

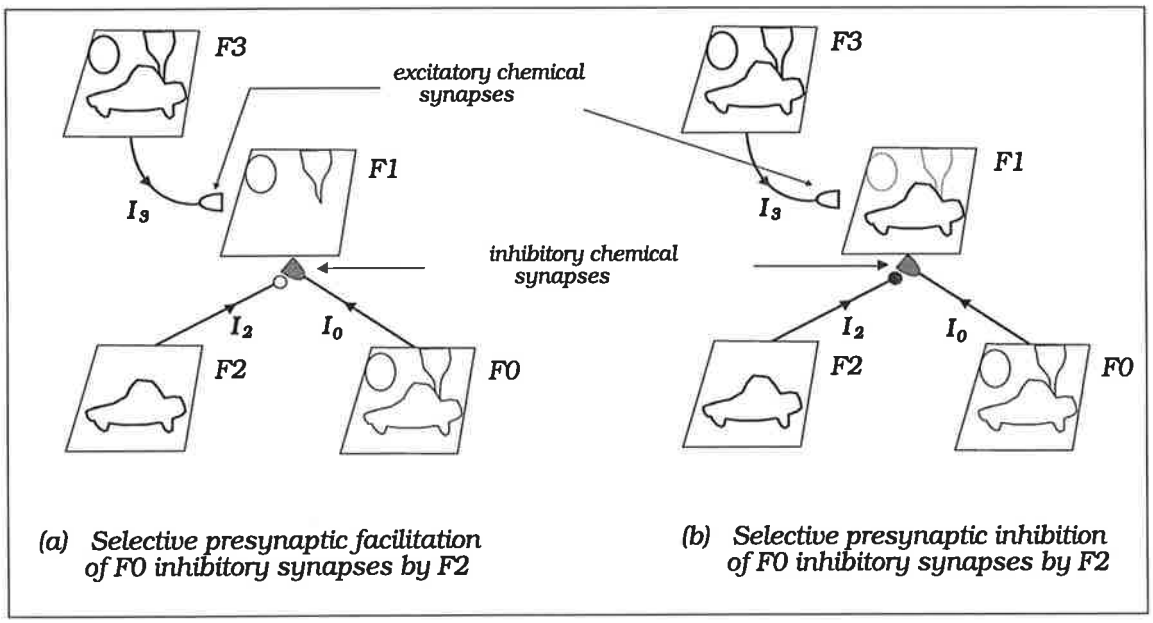


FIGURE 4.45. Schematic representing a neural circuit of presynaptic facilitatory and inhibitory interactions between transmitter gated 2-D shunting competitive neural layers.

If *F2* facilitates the $F0 \rightarrow F1$ pathways by pattern selective presynaptic signals, as in Fig. 4.45(a), then Field *F0* will selectively inhibit the cellular activity of Field *F1*. Therefore, only those active $F3 \rightarrow F1$ pathways that are not matched by the active $F0 \rightarrow F1$ pathways will be able to activate Field *F1*. However, if Field *F2* selectively inhibits the $F0 \rightarrow F1$ pathways, as shown in Fig. 4.45(b), then the steady state activity of Field *F1* will be enhanced in some locations and competitively suppressed in others. The level of enhancement (and suppression) is not only determined by the strength of signals from *F0*, *F2* and *F3*, but also by the tonic level of inhibition that Field *F0* can exert on *F1*. Figures 4.46 and 4.47 show the steady state simulation results for the two cases discussed.

FIGURE 4.46. Steady state simulation results of pattern selective facilitation of an inhibitory input synapse.

The above simulation results show that the steady state neural pattern of activity across Field F1 (see the column labelled by x_{ij}^{F1}) does not contain those portions of the excitatory synaptic input from Field F3 that was effectively removed by the facilitated inhibitory synapses from Field F0. Note that if an inhibitory synapse is to be facilitated (either by pattern or spatially specific modulatory signals) then its tonic level of inhibition should be very low. Similarly, if an inhibitory input synapse is to be inhibited presynaptically, then its tonic level of inhibition should be high. In general, where an inhibitory input synapses is to be presynaptically modulated by either facilitatory or inhibitory signals, the tonic level of inhibition provided by the inhibitory synapses should be moderate.

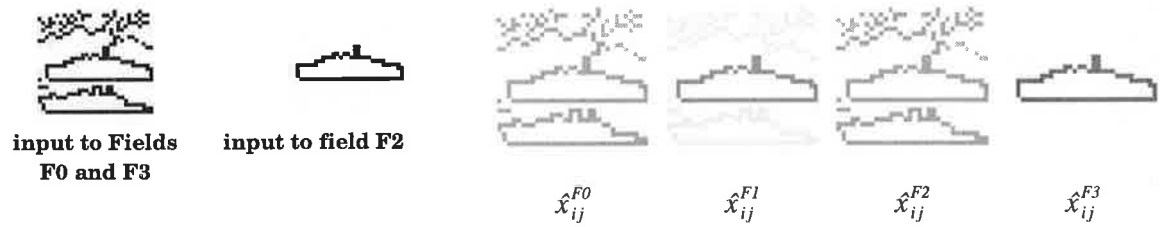


FIGURE 4.47. Steady state simulation results of pattern selective inhibition of an inhibitory input synapse.

Below we propose how the neural mechanisms of selective presynaptic facilitation and inhibition may be used to separate the spatial patterns of neural activities that are common to two neural Fields from the spatial patterns that is not shared. Consider the two dimensional scenario of Figure 4.48 which shows two input Fields F0 and F2 (that contain a common activity - the car boundary) and two output Fields $F1_a$ and $F1_b$. As shown, the common neural activity between F0 and F1 may be transferred to $F1_a$ if both $F0$ and $F2$ project excitatory synapses to $F1_a$, while each also being involved in the facilitation of the other. For the example shown, the neural activity that is in F0 but not in F2 may be transferred to the second output Field ($F1_b$) if Field F2 presynaptically inhibits the excitatory synapses of F0 at the input of F1.

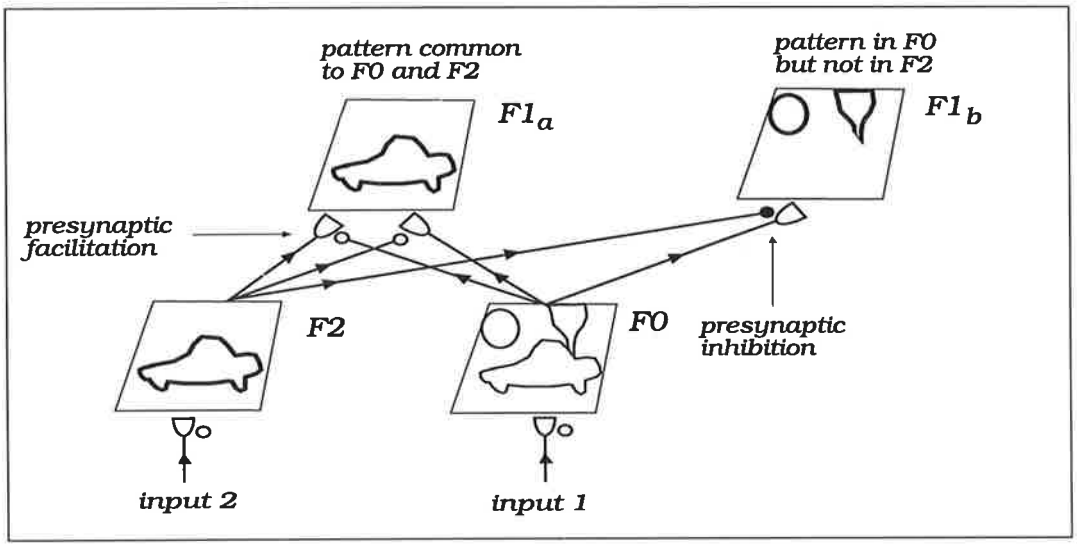


FIGURE 4.48. Proposed scheme for the separation of spatial patterns between two input fields by presynaptic facilitation and inhibition.

The simple circuit shown above is not sufficient when in addition to the common neural pattern of activity, Field F2 also contains an additional pattern of activity that does not appear across Field F0. The minimum neural circuit that can completely separate the neural activities is shown in Figure 4.49.

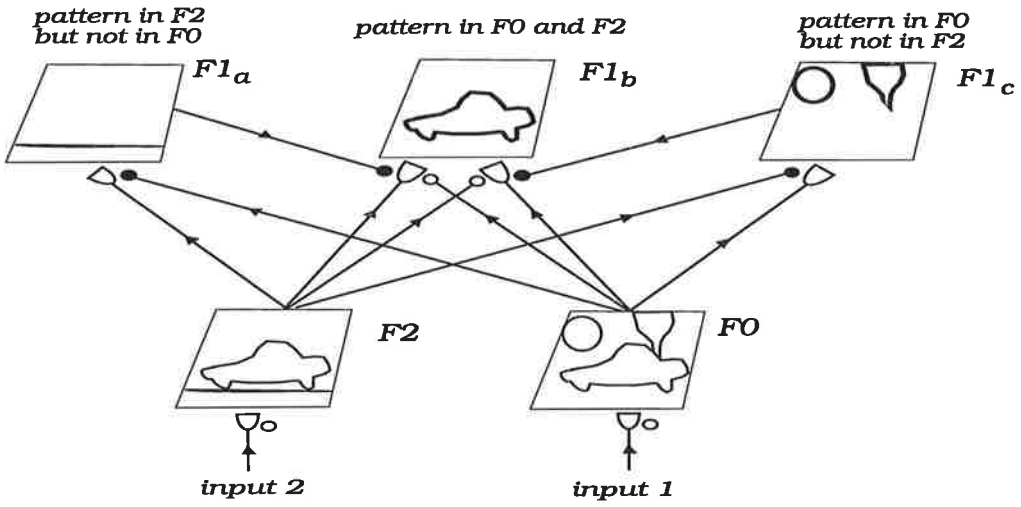


FIGURE 4.49. Neural circuit for the complete separation of 2-D neural patterns of activity between two shunting competitive neural layers.

The purpose of presynaptic inhibition of $F1_b$ by $F1_a$ and $F1_c$ is to ensure that Field $F1_b$ remains inactive when only one of its input Fields ($F0$ or $F2$) is active. Thus, there will be a transient activity in $F1_b$ as soon as one of its input Fields become activated.

The three neural Fields in Figure 4.49 ($F1_a, F1_b$ and $F1_c$) now contain a useful separation of the two input patterns. This breakdown of neural activities between two neural layers will form an important neural design principle when we propose the more advanced concepts and neural circuits in latter chapters of the thesis. As an example of how the above neural circuit can be extended to transfer the separated patterns into another layer, consider Figure 4.50.

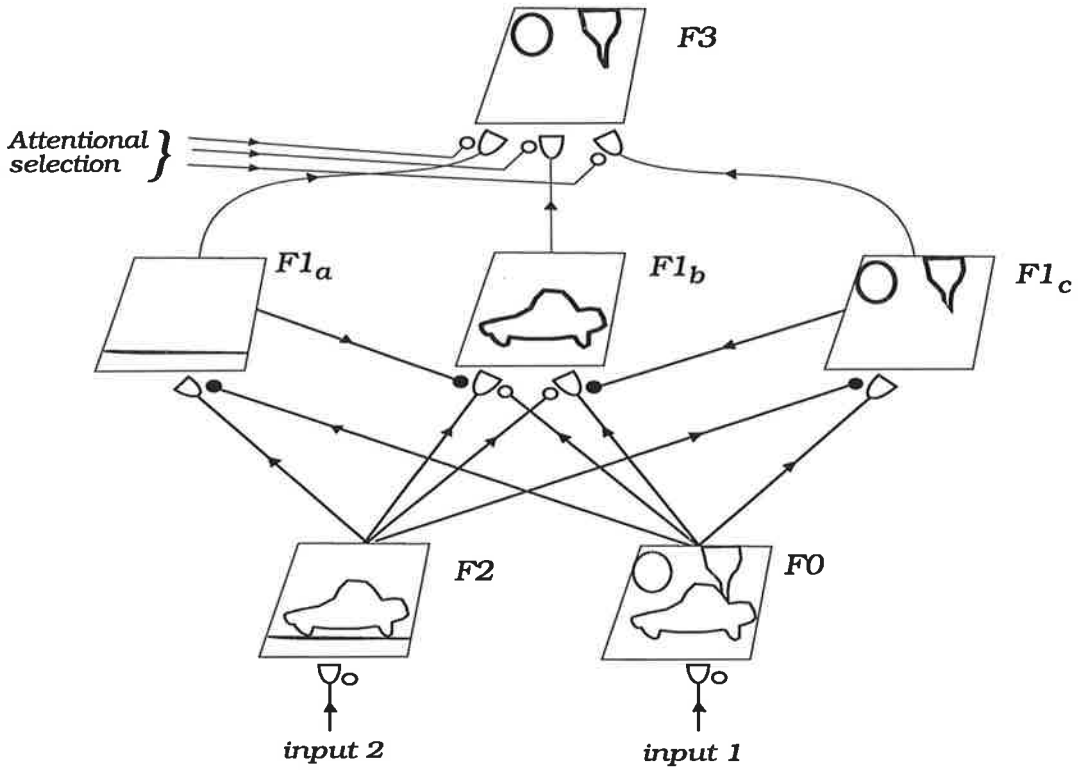


FIGURE 4.50. Neural circuit for the attentional selection of the separated spatial patterns.

As indicated in the above figure, the separated spatial patterns of activity may then be selectively transferred into another Field (F3) by a winner-take-all competitive decision system that may consist of a minimum of three neurons, each presynaptically facilitating one set of pathways into F3.

In the next section we extend the above ideas further by suggesting how the neural mechanisms of presynaptic facilitation and presynaptic inhibition may be used in selective resonance and top-down memory guided selective attention.

4.8 Top-down Memory Guided Selective Attention

In this section we formalise our neural design principles by considering the steady state 2-D neural patterns of activity shown in the three neural Fields of Figure 4.51. Let us suppose that Field F1 is more sensitive to neural signals from Field F2 than Field F0 (indicated by the thicker excitatory synaptic pathway $F2 \rightarrow F1$). Let us also assume that the two dimensional spatial input to Field F2 is a subset of the input to Field F0 (i.e., input 2 is a subset of input 1). Then the steady state 2-D neural pattern of activity across Field F1 (X_{ij}^{F1}) will contain this subset (the car boundary). Stated in another way, the car boundary, represented by neural activity X_{ij}^{F1}, is embedded in a complex bottom-up 2-D neural pattern of activity X_{ij}^{F0} that appears across Field F0. Thus, X_{ij}^{F1} will only partially match X_{ij}^{F0}.

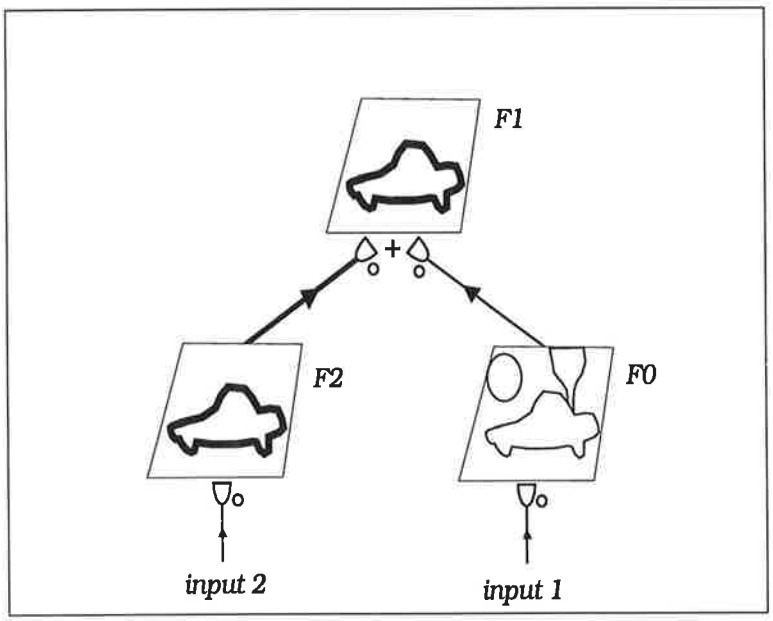


FIGURE 4.51. Example of steady state 2-D neural patterns of activity for two neural layers (F0 and F1) that is partially matched.

Suppose that it required to maximize the match between the spatial patterns of activity across Fields F0 and F1. In other words, we wish to address the following questions:

How can the 2-D neural pattern of activity X_{ij}^{F1} across one layer of transmitter gated shunting competitive neural cells be used to maximize the match with another layer from which it receives a complex bottom-up 2-D pattern of activity X_{ij}^{F0} within which X_{ij}^{F1} is embedded. How can the 2-D neural pattern of activity X_{ij}^{F1} selectively segment portions of the bottom-up input with which it can match or resonate? How is the selective resonance achieved?

Using the properties of transmitter gated 2-D shunting competitive neural layers and the properties of the neural mechanisms of presynaptic facilitation, we are now ready to postulate a solution to the above problem. We postulate that a facilitatory presynaptic feedback from the higher neural layer to the lower neural layer mediates the neural mechanism of selective attention. In the context of Figure 4.52, a top-down facilitatory presynaptic feedback (X_{ij}^{F1}) from Field F1 to Field F0 will ensure that F1 attends to the neural pattern X_{ij}^{F1} in the bottom-up pattern X_{ij}^{F0} (but only when F1 receives the relevant signals from F2). If the boundary of the car, represented by the 2-D neural pattern of activity, is perfectly aligned in F2 and F0 (in scale, position and orientation), then the net steady state will be a matching 2-D neural pattern of activity between Fields F0 and F1 (the problem of misalignment is addressed in subsequent chapters).

We call the neural mechanism by which Field F1 facilitates its own input from Field F0 as *top-down selective attention* and the resultant state as *selective resonance*. Figure 4.52 illustrates the basic neural mechanisms of top-down selective attention and selective resonance between two neural Fields F0 and F1, while Figure 4.53 illustrates the equivalent neural circuit.

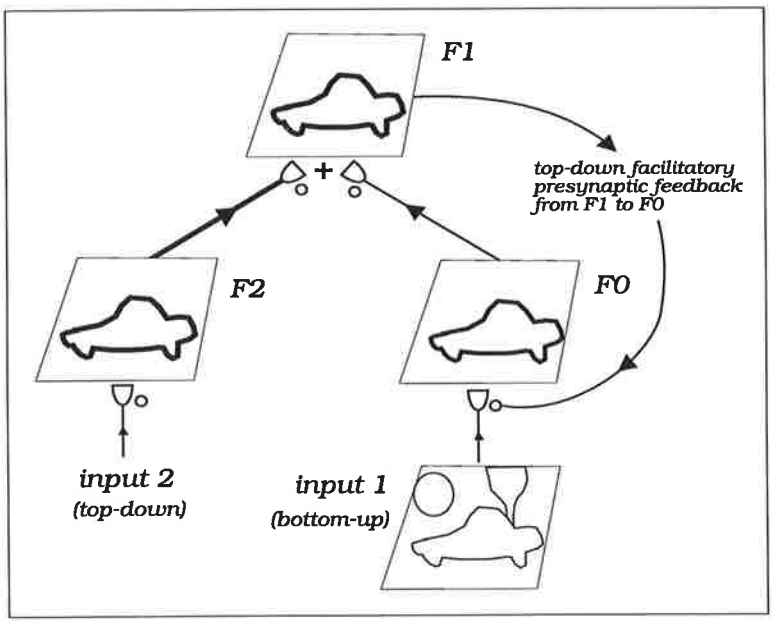


FIGURE 4.52. Selective attention and selective resonance via top-down facilitatory presynaptic feedback.

Top-down presynaptic facilitatory feedback from Field F1 to the input synapses of Field F0 regulates the bottom-up signal flow into Field F0. Since the gain of the facilitated synapses is amplified, the corresponding bottom-up inputs into F0 are also amplified as they enter the layer to provide increased excitation of their target cells. Competition in F0 quickly annihilates the neural activity of all other active but

non-facilitated cells in the layer. The resultant activity in F0 now begins to match the contents of F1. This process leads to a selective resonant state between Fields F0 and F1.

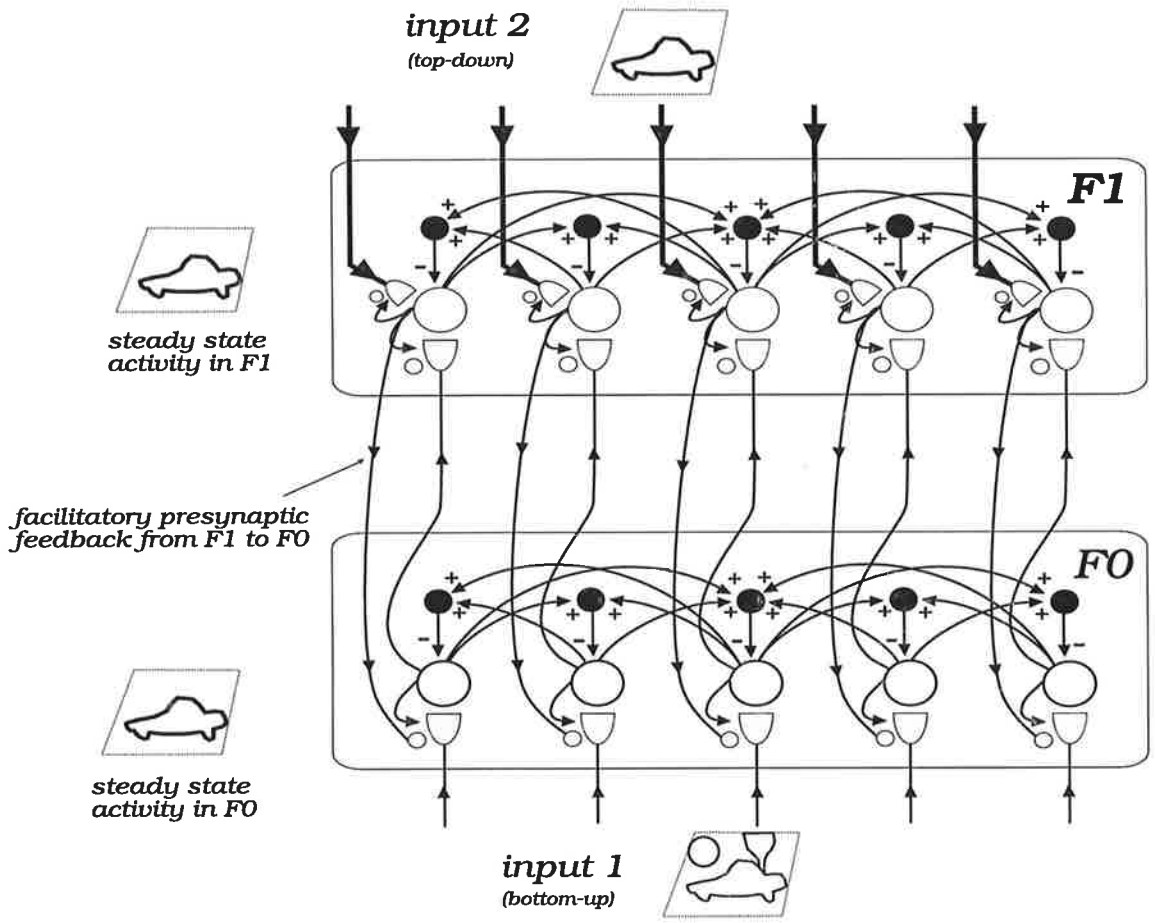


FIGURE 4.53. Neural circuit for the top-down selective attention and selective resonance via the top-down facilitatory presynaptic feedback.

Because of the bidirectional nature of interactions between the two neural layers, the circuit is called ***Feedforward Excitation-Feedback Presynaptic Facilitation.***

The simplest modification that needs to be made to the equations of each Field to model the FFE-FBPF neural circuit is to incorporate the top-down presynaptic facilitation into the synaptic dynamics of Field F0. Ignoring the effect on the storage rate of transmitter, the following equation then specifies the dynamics of the mobilized transmitter into the input synapses of Field F0.

Transmitter dynamics

$$\frac{dy_{ij}^{F0}}{dt} = [\beta_y + H_y^{F0} f(x_{ij}^{F1})](u_{ij}^{F0} - y_{ij}^{F0}) - J_{ij}[\rho_y + K_y f(x_{ij}^{F0})][y_{ij}^{F0} - Y^{F0}]^+ - \gamma_y y_{ij}^{F0} \tag{4.49}$$

which says that the transmitter mobilization rate at input synapses of Field *F0* is increased by the active top-down feedback signals $f(x_{ij}^{F1})$ from Field *F1*, where H_y^{F0} is the gain factor. This is counterbalanced by the release of the mobilized transmitter by the correlated firing of the bottom-up input signal and the postsynaptic feedback signal $f(x_{ij}^{F0})$ from Field *F0*. The equation for the storage transmitter in the input synapses of Field F0 is given by

$$\frac{du_{ij}^{F0}}{dt} = \alpha_u(z_{ij}^{F0} - u_{ij}^{F0}) - [\beta_u + K_u J_{ij} f(x_{ij}^{F0})] \tag{4.50}$$

In the simulation whose result is shown in Figure 4.54, Field *F1* was designed to be more sensitive to the top-down input (input 2), represented in the circuit of Figure 4.53 by the thicker top-down synaptic pathway. This is achieved by choosing the postsynaptic gains in Field *F1* to be much larger for the top-down synapses (see section 4.4). The purpose of choosing the larger top-down postsynaptic gain for Field *F1* is to transfer into *F1* a spatial 2-D neural pattern of activity that is also embedded in the complex bottom-up neural activity of Field *F0*.

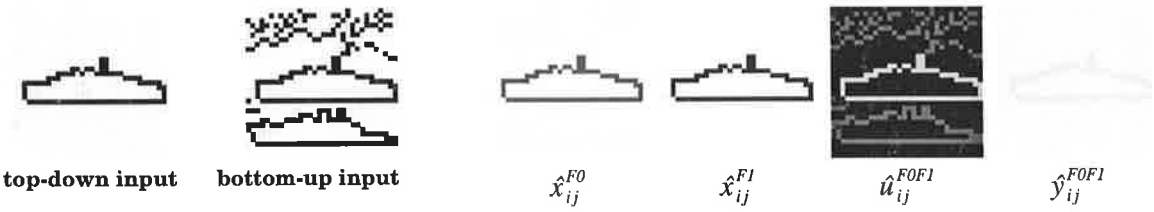


FIGURE 4.54. Steady state simulation results of selective resonance in the Feedforward Excitation-Feedback Presynaptic Facilitation (FFE-FBPF) neural circuit. Parameters for the individual layers are given in Chapter 5.

Above simulation results (achieved with the equilibrium point approximation of the cellular activity) demonstrate that although the two interacting neural layers are receiving inputs that are only partially matched, the top-down pattern selective facilitatory presynaptic feedback has forced one of the layers (Field *F0*) to shed-off the non-matching neural activity.

Other relevant simulations of the FFE-FBPF neural circuit are given in Chapter 5. In the next section we demonstrate the potential of the FFE-FBPF neural circuit in an application to object recognition in cluttered real-world images.

4.8.1 Recognition of an Object in Cluttered Visual Images

We now demonstrate the potential of the FFE-FBPF neural circuit in real-world visual applications by providing simulation results on the recognition of the 2D shape of a 3D object in cluttered images. The primary goal of the simulation described below is to demonstrate the unique nature and the power of the neural mechanism of top-down memory guided selective attention in complex visual images. The images used in the simulation are therefore of the type that we believe to be too difficult for conventional object segmentation algorithms and even some of the most advanced neural models of pre-attentive vision and object recognition, basically because of the difficulty in separating the target object from its background. The FFE-FBPF neural circuit in its current stage does not use distributed sampling of its input (i.e., its inputs are derived from single synapses) and because it ignores the orientation of object edges, it can still fail, particularly on textured images. Similarly, the current circuit cannot recognize an object when some of its boundary is not detected or is occluded. Nevertheless, this simulation is useful because it will reveal that the proposed neural mechanism of top-down memory guided object segmentation in cluttered visual backgrounds is an extremely efficient process. The simulation will also reveal further problems that need to be addressed.

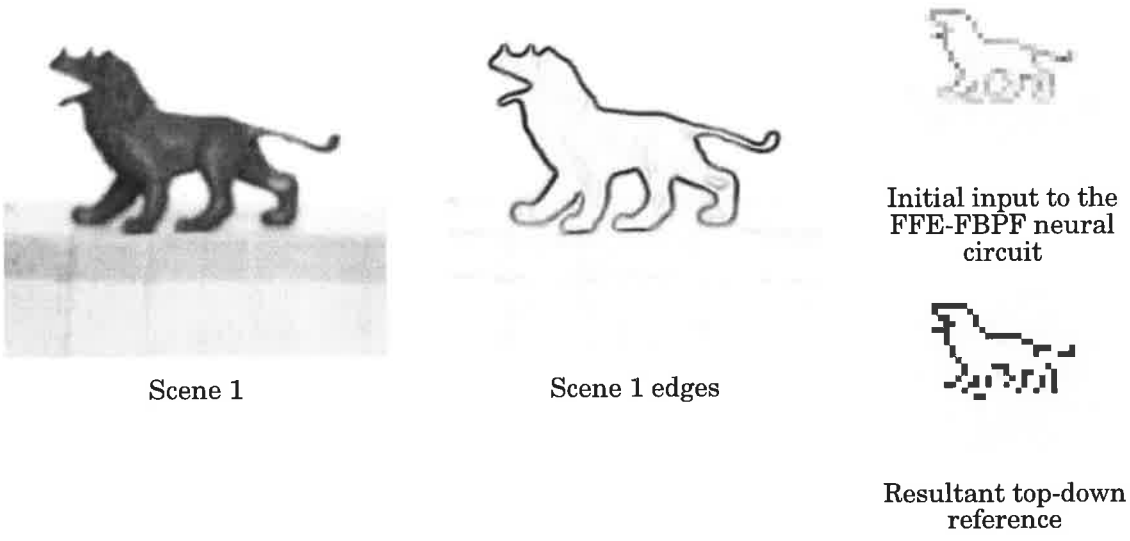


FIGURE 4.55. Target object to be recognized in cluttered visual images by the FFE-FBPF neural circuit.

Figure 4.55 shows the target object (a toy dinosaur) whose edges (all lumped into one array by a simple edge detector, regardless of their directions) are used as the initial input and hence the top-down reference into the FFE-FBPF neural circuit. The procedure for image rendering, generation of the top-down reference and the threshold settings is described below.

Image rendering, top-down reference and threshold settings

Original grey level images (8 bits, 256x256 pixels, captured by a Vidicon camera and then digitized) were preprocessed by a 3x3 Sobel edge operator to obtain object edges. The resultant edge processed images were then scaled to the size of the FFE-FBPF neural circuit (32x32 cells) by a simple averaging procedure (which therefore reduces the resolution by a factor of 8). The edge processed image of the target object (shown above) was initially fed to the circuit. The resultant steady state activity was then used as the top-down reference. This simple procedure by-passes the need to learn the shape, although the end result would have been equivalent had the learning of the top-down memory been enabled. Note that as a result of the competition in the circuit, the weaker edges in the image of the target object, such as the horizontal line and some parts of the target shape, did not survive and were absent in the top-down reference shape, as shown in Figure 4.55.

Recognition in the circuit is achieved when the match between the spatial patterns of activity across Fields F0 and F1 (currently measured by the cosine of the angle between the two multidimensional vectors) exceeds the pre-set threshold level of 0.95 and when the time-rate rate of change of the match is below the pre-set steady state threshold level of 0.001 (i.e., the computational decision is taken at the steady state). The threshold for recognition is determined by presenting two similar objects separately in clean background and then finding a threshold at which the two are discriminated (the two objects shown in scene 1 and scene 2). This threshold is then set and remains fixed thereafter. The steady state threshold is also determined from the same experiment and it too remains fixed thereafter.

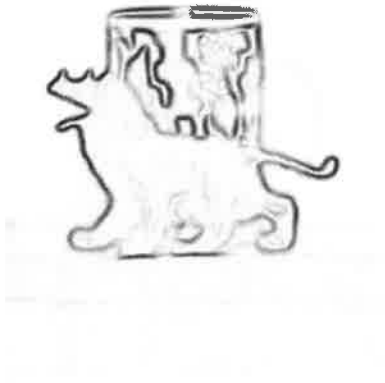
The recognition response of the circuit on the cluttered test images is shown graphically in Figures 4.57 and 4.58. In summary, the circuit failed on only one of the shown test images. The target object was not recognized in scene 9 at the required threshold because some of its input edges were either missing or were too weak to re-enter into the circuit. However, because of its simplicity, the circuit has a number of other problems, as discussed below.

Scene 2

Scene 2 edges

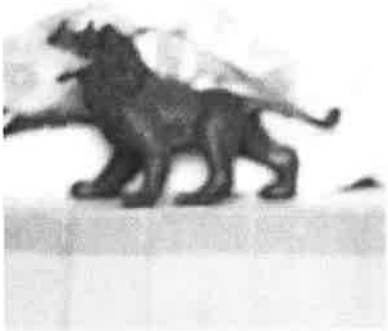
Input to FFE-FBPF circuit

Scene 3

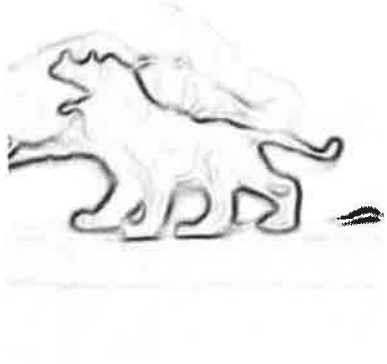


Scene 3 edges

Input to FFE-FBPF circuit



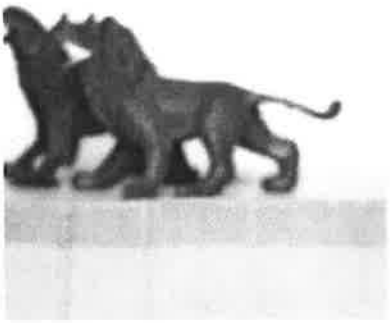
Scene 4



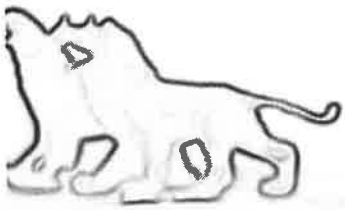
Scene 4 edges

Input to FFE-FBPF circuit

FIGURE 4.56. Cluttered visual images used to demonstrate the power of top-down memory guided object segmentation and recognition by the FFE-FBPF neural circuit.



Scene 5



Scene 5 edges

Input to FFE-FBPF circuit

Scene 6

Scene 6 edges

Input to FFE-FBPF circuit

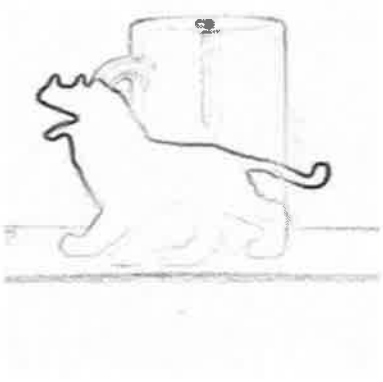
Scene 7

Scene 7 edges

Input to FFE-FBPF circuit

FIGURE 4.56. (Cont.) Cluttered visual images used to demonstrate the power of top-down memory guided object segmentation and recognition by the FFE-FBPF neural circuit.

Scene 8



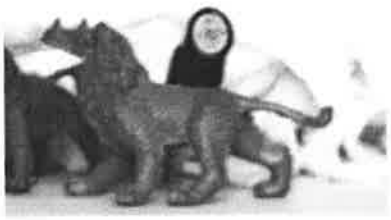
Scene 8 edges

Input to FFE-FBPF circuit

Scene 9

Scene 9 edges

Input to FFE-FBPF circuit



Scene 10

Scene 10 edges

Input to FFE-FBPF circuit

FIGURE 4.56. (Cont.) Cluttered visual images used to demonstrate the power of top-down memory guided object segmentation and recognition by the FFE-FBPF neural circuit.

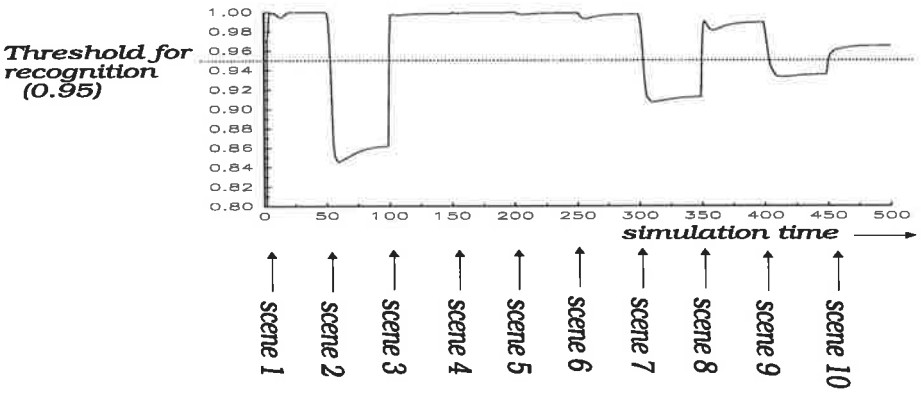


FIGURE 4.57. Measure of the match between the top-down reference across Field F1 (top-down memory) and the selectively attended object across Field F0, for various cluttered images.

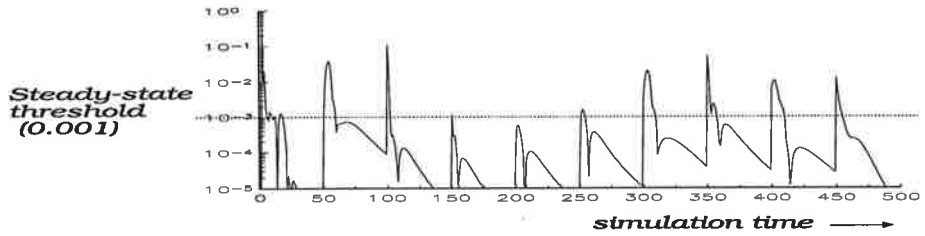


FIGURE 4.58. Time-rate of change of the match between Fields F0 and F1 on cluttered input images (log scale).

Note that if the top-down presynaptic facilitation (i.e., selective attention) was to be disabled, the circuit would completely fail to recognize the target in all the cluttered scenes. As an example of a typical state when a non-matching object is presented to the circuit, below we show the various variables in Field F0, at the time when a steady state is reached for the object shown in scene 2 of Figure 4.56.

Non-matching input (J_{ij}^{F0}) | Top-down reference ($\hat{x}_{ij}^{F1}$) | Activity of Field F0 ($\hat{x}_{ij}^{F0}$) | Stored transmitter, Field F0 ($\hat{u}_{ij}^{F0}$) | Mobilized transmitter, Field F0 ($\hat{y}_{ij}^{F0}$)

FIGURE 4.59. Example of a steady state activity in the FFE-FBPF neural circuit on a non-matching object.

The circuit was also tested on numerous other images, some of which were of extremely low contrast and high clutter, such as the image shown in Figure 4.60. Although the circuit responded correctly on such images in most of the cases (i.e., it gave a correct recognition response on virtually all the images that contained the target object), in a number of cases this was more due to coincidence than circuit robustness, as will be described below.

Target object in extreme clutter and low contrast

Low contrast edges of the cluttered scene

FIGURE 4.60. Example of an image with extreme clutter and low contrast on which the FFE-FBPF neural circuit succeeded.

Further testing on textured images has exposed the problem with the circuit. The problem, not unexpected, is due to the fact that the current FFE-FBPF neural circuit does not have sufficient processing flexibility nor does it use as much featural information as can be derived from an image in order to work robustly. Since we are using a very simple neural circuit whose edge inputs are obtained by the Sobel operator which lumps all the edges into one edge map (or layer), regardless of their direction, the circuit will respond with a recognition whenever spatial location of the active edge inputs matches the locations of the active top-down pathways. Hence, when a textured image is presented that contains edges in positions that correspond to the target boundary, the circuit in its present form will use whatever edges it can in order to match their spatial location with the location of the top-down edges. However, the existence of an active edge in the same location as the active top-down pathway does not necessarily mean that the two are matched unless their local directions are matched also. Hence the robustness of the circuit can be improved (at the expense of increased complexity and computations) by incorporating the mechanisms of top-down matching for various spatial orientations. Although the implementation and computer simulations of such an extensive circuit may pose a computational problem for the current serial processing computers, such an architecture is ideal for parallel processing and is probably implemented in the visual brain of primates. The simple scheme used in

our simulation whereby a simple edge detector (Sobel operator) lumps all the edges into one 2-D edge map, regardless of their directions, does not represent a robust solution nor does it represent the properties of the primary visual cortex. The Sobel edge operator was primarily used as a quick image pre-processing method to obtain the necessary edge inputs into the circuit so that we could demonstrate the concepts and the power of top-down selective attention in cluttered visual scenes. In Chapter 8 we will propose a possible extension to the circuit for orientation invariant recognition of 2-D shapes and objects.

4.8.2 Pulsating FFE-FBPF Neural Circuit

Up to now we have modelled the FFE-FBPF neural circuit by the equilibrium point behaviour (i.e., non-pulsating dynamics). Below we present two simulations for the case when the thresholds in the circuit are increased ($Y = 0.2, \Theta = 0.3$) until each neuron exhibits pulsating dynamics for a fixed input signal. The first example (Fig. 4.61) is an open loop simulation, while the second example (Fig. 4.62) is a closed loop simulation. In these simulations, input 2 (top-down input to Field *F1*) was introduced 12 iterations after the introduction of the bottom-up input to Field *F0* (i.e., roughly half way through the cycle of Field *F0*) which causes an initial 180^0 phase shift between oscillations in *F0* and *F1*.

Figure 4.61 shows that, for the given bottom-up and the top-down input patterns, the oscillating neural pattern of activity across the two Fields is not locked into synchronized oscillations. However, by closing the loop with the facilitatory presynaptic feedback, the two layers can be forced into resonance, as shown in Figure 4.62. In fact, the lower of the two layers (Field F0) is forced to resonate with the higher layer (Field F1).

It should be noted that while the two layers are almost always in resonance when the loop is closed, the 2-D neural pattern of activity that is resonating need not always be constant. Simulation data in Figure 4.62 shows that the resonating pattern on one cycle is replaced by a pattern that was selectively ignored on the previous cycle.

Comparison of the graphs of 4.63(a) and (c) (open and the closed loop response respectively) shows that while the two corresponding cells are firing asynchronously in the open loop system, they are quickly brought into synchronization (resonance) via the top-down facilitatory presynaptic feedback. These graphs also show that, during resonance, all amplitudes that are associated with the variables of the first layer (Field F0) are enhanced.

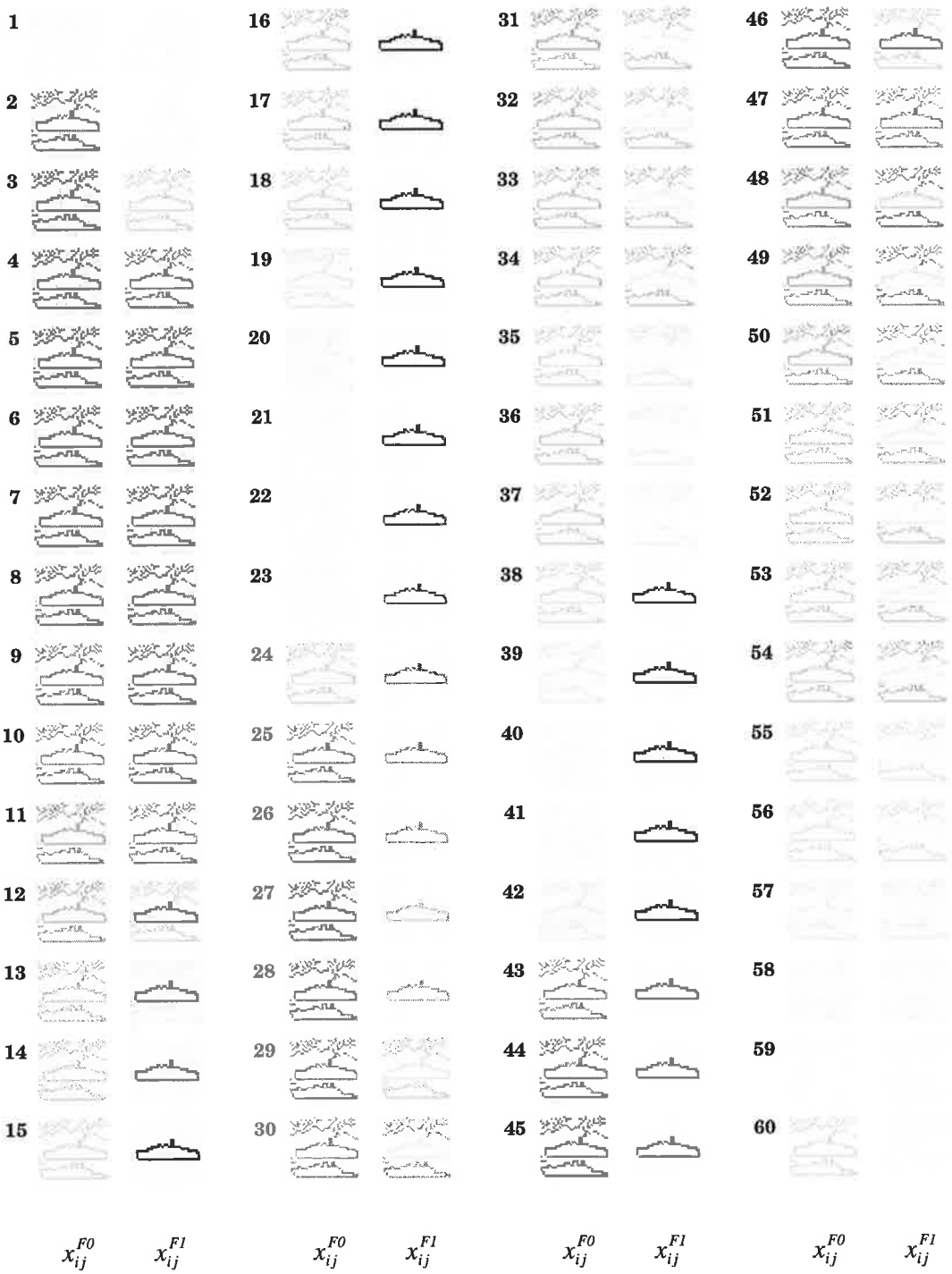


FIGURE 4.61. Asynchronous oscillations without the facilitatory top-down feedback. Time flows downwards (shown are iterations 1 - 60).

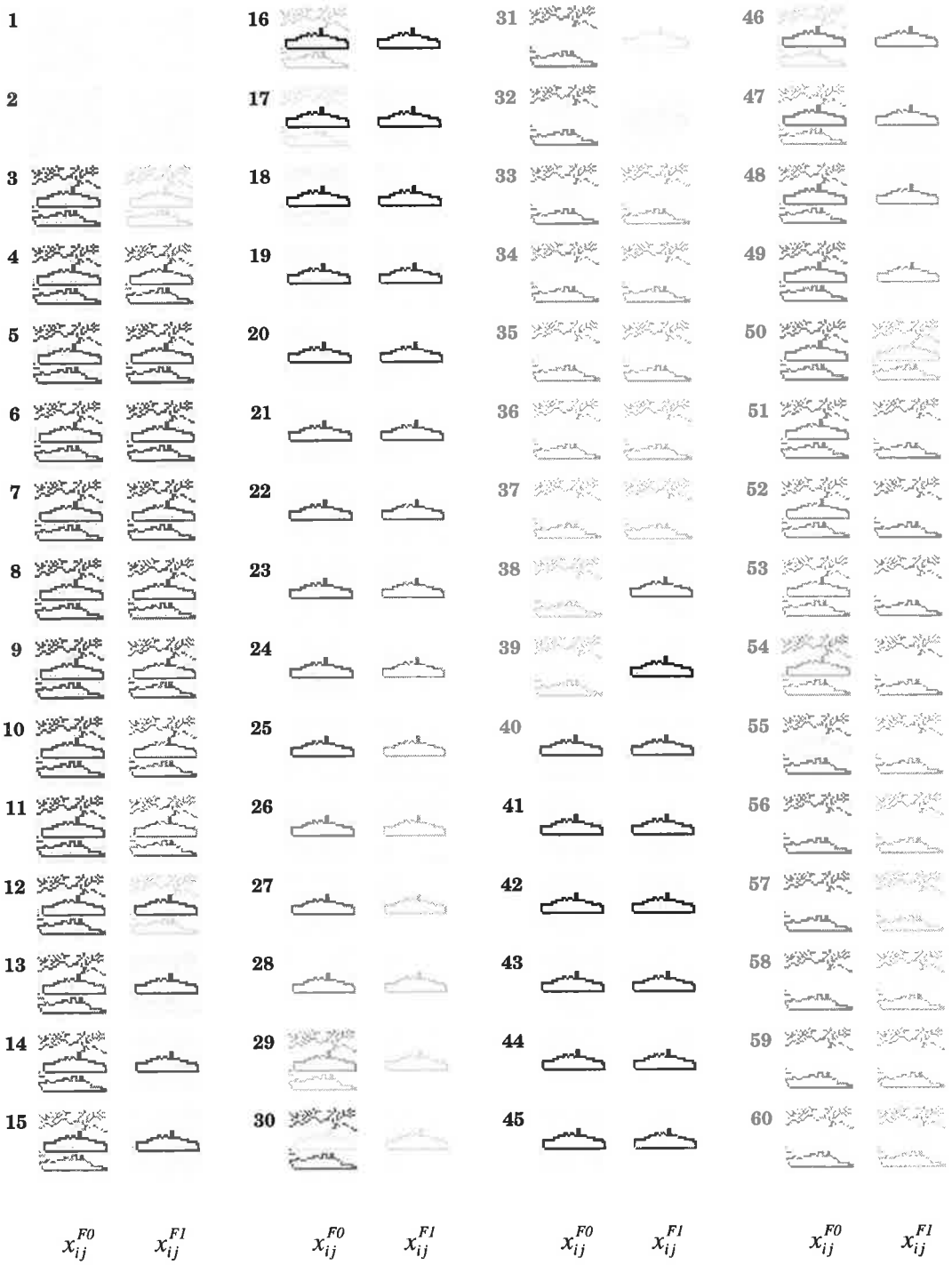


FIGURE 4.62 Selective synchronization (resonance) with the facilitatory top-down feedback. Time flows downwards (shown are iterations 1 - 60).

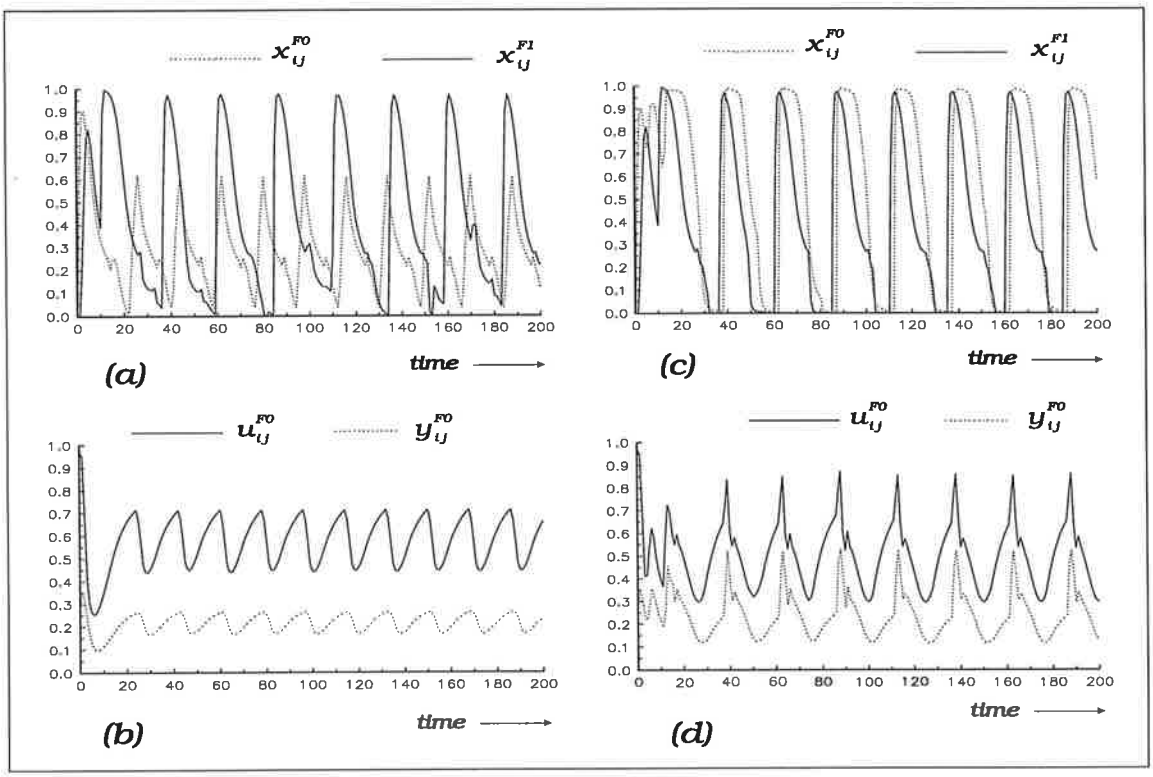


FIGURE 4.63. Cellular and synaptic activity for the corresponding locations in the open and the closed loop system of a pulsating FFE-FBPF neural circuit. Graphs (a) and (b) show the dynamics of the variables in the open loop system (without the facilitatory feedback from F1 to F0), while graphs (c) and (d) show the same variables when the loop is closed.

4.9 Selective In-attention

In contrast to the top-down facilitatory presynaptic feedback where the neural pattern of activity across the facilitated layer is forced to match the activity of another layer, top-down inhibitory presynaptic feedback, shown in Figure 4.64, does just the opposite. We call such process as *selective in-attention*. Figure 4.65 shows the steady state simulation results of selective in-attention for the same two inputs as used in the previous simulation.

As can be seen in the simulation results of Fig. 4.65, the top-down spatial pattern across Field F1 has, through its inhibitory presynaptic feedback to F0, removed the same spatial pattern from Field F0 (see the column labelled by x_{ij}^{F0}). The removed input cannot re-enter into the circuit at F0 as long as the relevant top-down pathways are active.

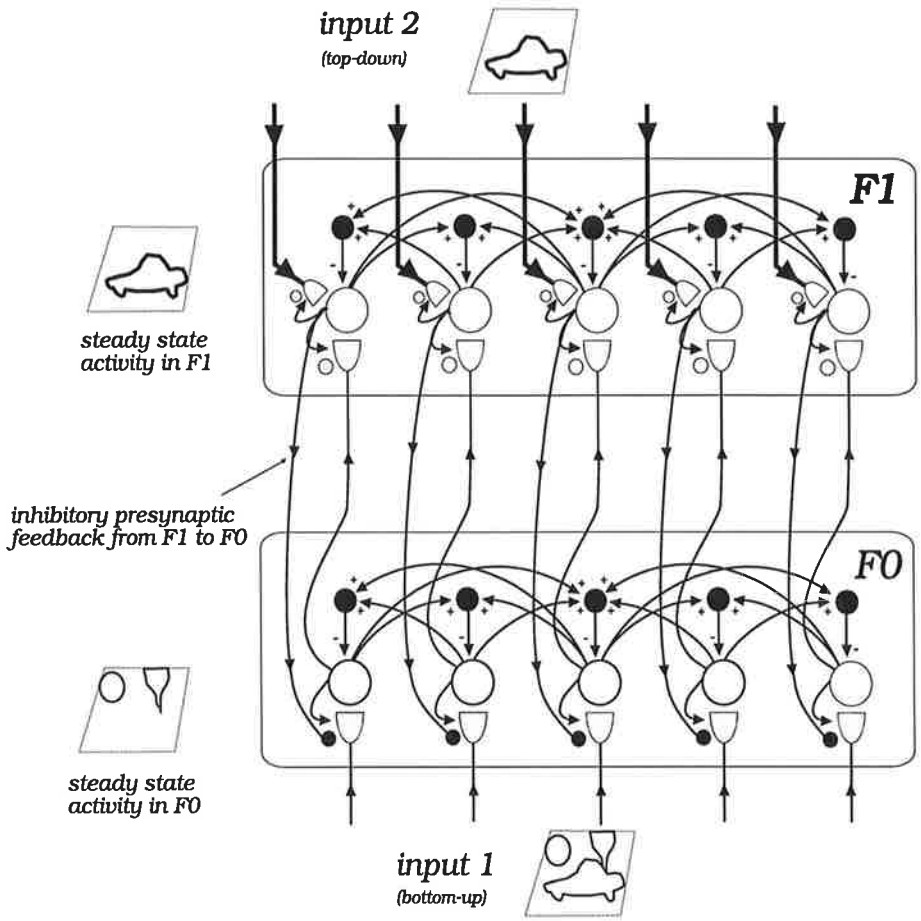


FIGURE 4.64. Neural circuit for selective in-attention by the top-down inhibitory presynaptic feedback.

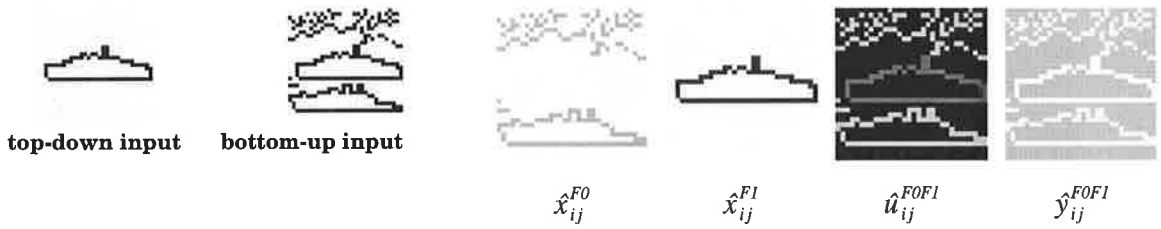


FIGURE 4.65. Simulation results of top-down presynaptic inhibition (selective in-attention).

There is no reason why the same top-down signals should not be used to selectively facilitate some bottom-up neural layers while being simultaneously involved in presynaptic inhibition of others. The purpose of selective in-attention in real-time self-organising neural networks is discussed in more detail in Chapters 7 and 8.

4.10 Modulation of Cellular Receptive Field Profiles

Up to now we have concentrated primarily on presynaptic modulation of neural layers whose cells were assumed to have a single input synapse. The experimental data from complex biological neural systems (e.g., from monkey's visual cortex, unlike that from the *Aplysia* circuit that is characterized by monosynaptic pathways) indicates that each cell samples its input via distributed input receptive fields (IRFs) that may consist of large number of input synapses (of the order of several thousand). Although the early electrophysiological recordings from the cat and the monkey visual cortex obtained during 1960's -1970's (Hubel and Wiesel, 1959, 1962, 1963, 1968, 1977), upon which most of the current neural models of cellular receptive field profiles are based, did not provide evidence for modulatory mechanisms, the most recent data from monkey's visual cortex provides striking evidence that cellular input receptive fields can be influenced by attentional factors (Moran and Desimone, 1985; Haenny and Schiller, 1988; Motter, 1993). The spatial profile of the cellular receptive fields in the monkey's visual cortex was found to be subject to various spatial and non-spatial attentional modulations that can dramatically alter their specificity and spatial size. Haenny and Schiller (1988) have for example reported that cellular selectivity of V4 cells can increase up to 1500%. The most recent data from V1 and V2 (Motter, 1993) suggests that even the very earliest feature detecting cells in the monkey's visual cortex, whose receptive fields are much smaller than those of V4, can be modulated. This data suggests that we need to extend our neuro-engineering design principles to explain how increased attention can fine tune cellular receptive fields. Some of the recent neural models of visual attention (Olshausen *et al.,* 1992, 1993; Lozo *et al.,* 1995, see Chapter 6) have been primarily concerned with the spatial aspects of visual attention and have not yet provided plausible neural mechanisms of how the input receptive field profile of a cell may be dynamically changed from a broadly tuned to a finely tuned spatial filter by an increased level of attention (or arousal).

In this section we provide simple modulatory mechanisms that can fine tune the cellular receptive fields from what are otherwise broadly tuned cells. First we provide a modulatory mechanism for simple unipolar and isotropic receptive fields and then extend the concepts to the more general case of bipolar and spatially oriented receptive fields. Computer simulations on grey level images demonstrate the effectiveness of the modulatory mechanisms. In Chapter 6 we will consider some possible sources of the modulatory signals and will show how the degree of attention itself may be automatically modulated in a simple memory based neural circuit that is able to detect the familiarity and the novelty of its inputs.

4.10.1 Modulation of Unipolar and Isotropic Receptive Fields

Let as now assume that each cell in a two dimensional PFE-SCNL samples its bottom-up inputs via a set of spatially distributed and dynamic excitatory synaptic pathways. A cell at location (i,j) in a 2-D array will thus sample its input from a local input region (or a cluster of cells in a preceding layer) through synaptic pathways whose strength is a two dimensional function of the distance between the input location (m,n) and the position of the sampling cell, (i,j). Let us also assume that this function specifies the level of transmitter production in each excitatory input synapse from spatial location (m,n) to cellular location (i,j), i.e., z^{F0}_{mnij}, and that it is constant in time. Then the simplest such function is a 2-D gaussian $G^{F0}_{EIRF}(m,n,i,j,\sigma_{ij})$, where the subscript *EIRF* refers to an **E**xcitatory **I**nput **R**eceptive **F**ield (as opposed to an **I**nhibitory **I**nput **R**eceptive **F**ield of IIRF, to be discussed in the next section), whose synapses provide cellular excitation.

$$\begin{aligned} z^{F0}_{mnij} &= G^{F0}_{EIRF}(m,n,i,j,\sigma_{ij}) \\ &= Z^{F0}_{ij} \exp\left(-\frac{(i-m)^2+(j-n)^2}{(\sigma_{ij})^2}\right) \end{aligned} \tag{4.51}$$

where the subscripts (mnij) refers to a bottom-up input pathway (m,n) into the cell at location (i,j); σ_{ij} defines the spatial extent of the EIRF. This EIRF can be facilitated by another 2-D spatial kernel $G^{F0}_{FMF}(m,n,i,j,\psi_{ij})$, referred to as the **F**acilitatory **M**odulating **F**ield or FMF (as opposed to the **I**nhibitory **M**odulating **F**ield or IMF), whose centroid is coincident with $G^{F0}_{EIRF}(m,n,i,j,\sigma_{ij})$, but whose spatial extent is narrower (i.e., $\psi_{ij} < \sigma_{ij}$). The spatial profile of the FMF may represent efferents from a cell (or cells) of another 2-D neural layer which may be at a higher hierarchical level, or it may represent efferents from neurons that regulate the general level of arousal, etc. The activity of the facilitatory neuron, represented by $f(x^{F1}_{ij})$, will therefore determine the strength of the FMF and hence the degree of fine tuning of the EIRF. Let F^{F0}_{mnij} be the magnitude of the FMF along the synaptic pathway (m,n) $\rightarrow$ (i,j). Then

$$F^{F0}_{mnij} = R^{F0}_{ij} \exp\left(-\frac{(i-m)^2+(j-n)^2}{(\psi_{ij})^2}\right) \tag{4.52}$$

where R^{F0}_{ij} is the maximum amplitude of the FMF (at cellular centroid), which in general may also be a function of the location in the 2-D array.

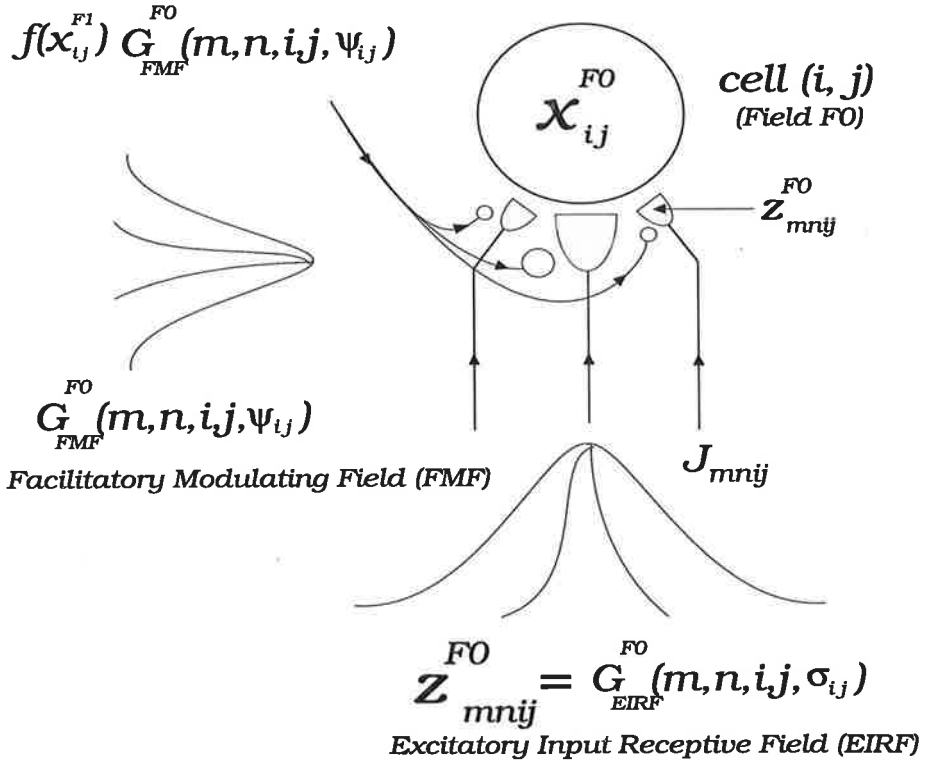


FIGURE 4.66. Facilitatory mechanism for the modulation of simple isotropic Excitatory Input Receptive Field (EIRF) profiles.

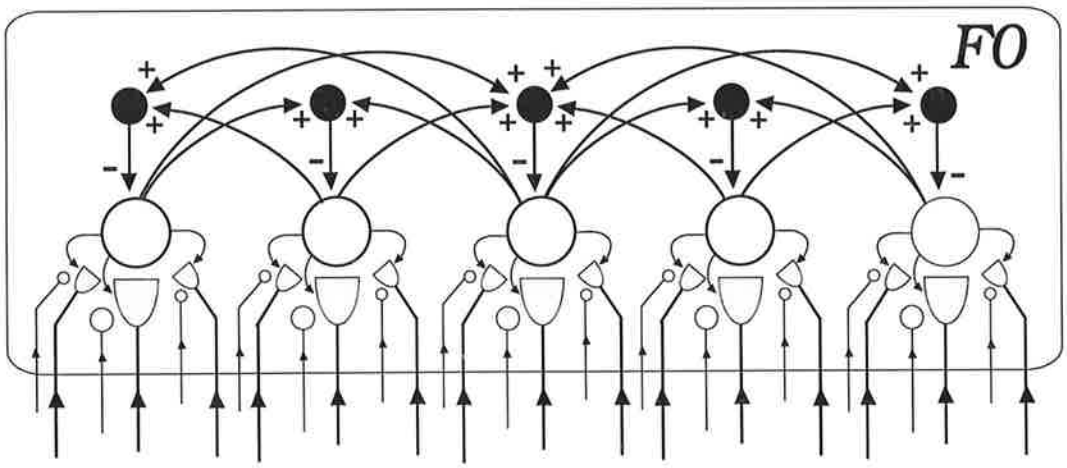


FIGURE 4.67. One-dimensional schematic of a 2-D PFE-SCNL with isotropic unipolar excitatory input receptive fields.

The following set of equations then define a two-dimensional PFE-SCNL whose cells sample their inputs via distributed and modulated excitatory synaptic pathways.

Postsynaptic cellular activity at equilibrium (x_{ij}^{F0})

$$x_{ij}^{F0} = \frac{BG \sum_{(m,n)} v_{mnij}^{F0}}{A + G \sum_{(m,n)} v_{mnij}^{F0} + \overline{G}\overline{v}_{ij}^{F0}} \tag{4.53}$$

Lateral inhibition ($\overline{v}_{ij}^{F0}$)

$$\frac{d\overline{v}_{ij}^{F0}}{dt} = -\overline{A}\overline{v}_{ij}^{F0} + \overline{B} \sum_{(p,q)\neq(i,j)} f(x_{pq}^{F0}) \tag{4.54}$$

Excitatory postsynaptic potential (v_{mnij}^{F0})

$$\frac{dv_{mnij}^{F0}}{dt} = -Dv_{mnij}^{F0} + J_{mn}^{F0}[\rho_v + K_v f(x_{ij}^{F0})][y_{mnij}^{F0} - Y^{F0}]^+ \tag{4.55}$$

Mobilized transmitter (y_{mnij}^{F0})

$$\frac{dy_{mnij}^{F0}}{dt} = [\beta_y + f(x_{ij}^{F1})G_{FMF}^{F0}(m,n,i,j,\psi_{ij})](u_{mnij}^{F0} - y_{mnij}^{F0})$$
$$- J_{mnij}^{F0}[\rho_y + K_y f(x_{ij}^{F0})][y_{mnij}^{F0} - Y^{F0}]^+ - \gamma_y y_{mnij}^{F0} \tag{4.56}$$

where J_{mnij}^{F0} is the input from the spatial location (m,n) to the cell (i,j). The term $f(x_{ij}^{F1})G_{FMF}^{F0}(m,n,i,j,\psi_{ij})$ in the above equation is the strength of the facilitatory modulating field that enhances the transmitter mobilization rate (and hence the gain) of the input synapses into the layer. This equation may thus be written as

$$\frac{dy_{mnij}^{F0}}{dt} = [\beta_y + f(x_{ij}^{F1})F_{mnij}](u_{mnij}^{F0} - y_{mnij}^{F0})$$
$$- J_{mnij}^{F0}[\rho_y + K_y f(x_{ij}^{F0})][y_{mnij}^{F0} - Y^{F0}]^+ - \gamma_y y_{mnij}^{F0} \tag{4.57}$$

Stored transmitter (u_{mnij}^{F0})

$$\frac{du_{mnij}^{F0}}{dt} = \alpha_u(z_{mnij}^{F0} - u_{mnij}^{F0}) - [\beta_u + J_{mnij}^{F0}K_u f(x_{ij}^{F0})](u_{mnij}^{F0} - y_{mnij}^{F0}) \tag{4.58}$$

which says that the stored transmitter is depleted by the correlated firing of the synaptic input signal J_{mnij}^{F0} and the postsynaptic feedback signal $f(x_{ij}^{F0})$.

A competitive neural layer whose cells sample their inputs via distributed receptive fields that can be modulated, thus altering their tuning characteristics, provides a building block for the neuro-engineering design of self-regulating artificial neural systems. We now qualitatively examine the processing characteristics of a modulated 2-D shunting competitive neural layer in an application to image filtering.

Image Filtering with Modulated Competitive Neural Layers

Below we provide computer simulations on several images for various levels of modulation. Note that the input image is 64x64 pixels in size while the simulated two dimensional PFE-SCNL consists of 28x28 cells, each cell sampling its inputs over a 7x7 image region. Although it would have been preferable to simulate a much larger layer, our simulations (being performed on a DOS based PC) are restricted by the computing platform. Hence the resolution of the filtered image is roughly half of the original. Nevertheless, these simulations clearly demonstrate that a change in a single facilitatory signal (such as what may be provided by a shift in the degree of general attention) can quickly alter the resolution of the sensory information that is to be passed to higher cortical layers.

In the following five examples, the original grey level images (256x256 pixels) of several indoor scenes were captured by a Vidicon camera, digitized and then stored. A 64x64 region was then selected from these images and used as the input to the PFE-SCNL. For each of the five test images, we have used the same but re-initialized PFE-SCNL. The only parameter that was varied during each of the simulation runs was the strength of the facilitatory signal $f(x_{ij}^{F1})$, which was constant across the whole layer and was slowly raised in steps of 0.1, starting from zero (unfacilitated case) up to 0.8. These increments were made after the layer reached a steady state (typically after 20 iterations of the equations using the first order Euler's approximation method). For each of the five examples, we show the original 256x256 image, the extracted 64x64 portion that was used in the simulation and the steady state response of the layer for 9 levels of facilitation.

Original gray level image (256x256 pixels).

Selected 64x64 region used as the input to the PFE-SCNL.

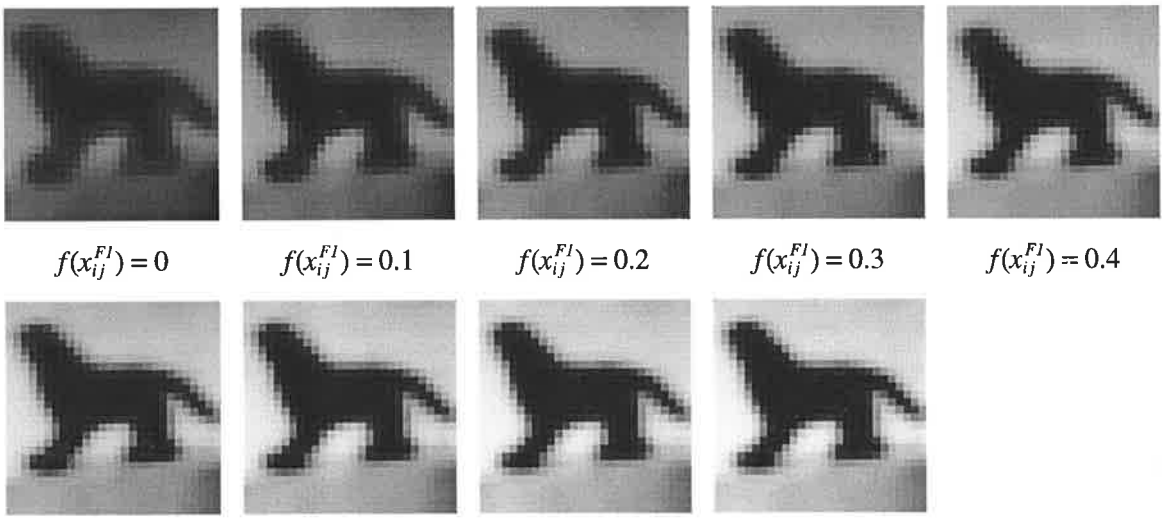
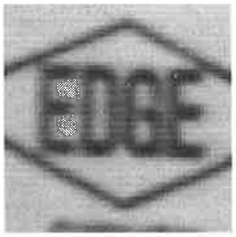


FIGURE 4.68. Filtering a 64x64 image by a 28x28 PFE-SCNL: Example 1. Shown are the steady state cellular levels for various magnitudes of the facilitatory signal $f(x_{ij}^{Fl})$. Layer parameters: $n = 786$ (28x28 cells); $B = 1.0$; $A = 1.0$; $\overline{A} = 0.1$; $\overline{B} = 0.1/786$; $\Theta = 0.1$; $G = 1000$; $\overline{G} = 50000$; $D = 0.5$; $\rho_v = 0.5$; $K_v = 5$; $\alpha_u = 0.05$; $\beta_u = \beta_y = 0.01$; $\rho_y = 0.05$; $K_y = 0.5$; $\gamma_y = 0.5$; $K_u = 0.5$, $\Upsilon = 0$.

Original 256x256 image depicting text on a cardboard box.



Selected 64x64 region used as the input to the PFE-SCNL.

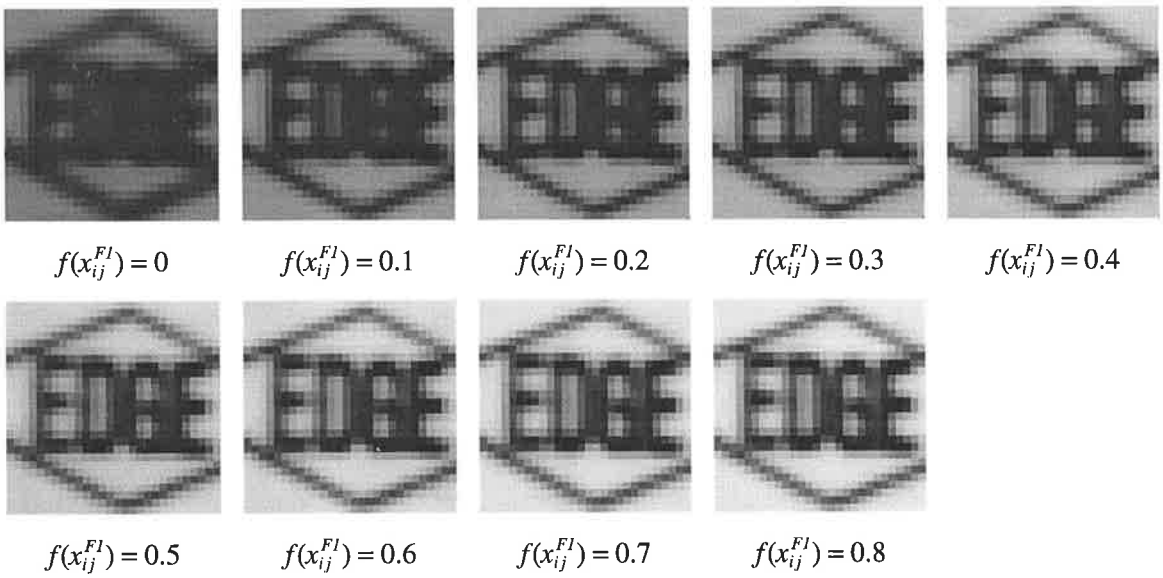
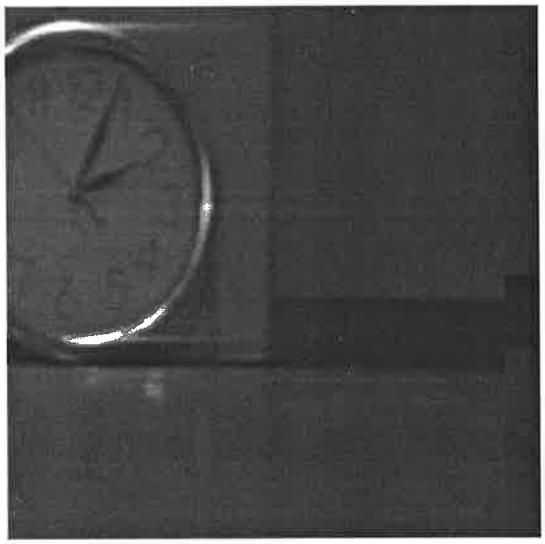


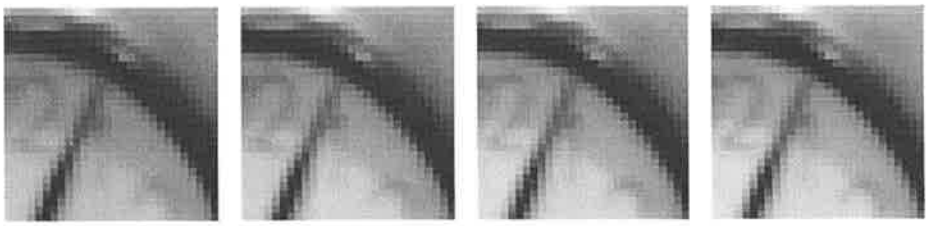
FIGURE 4.69. Filtering a 64x64 image by a 28x28 PFE-SCNL: Example 2. Parameters as for Figure 4.68.



Original 256x256 low contrast image of a clock face.

Selected 64x64 region used as the input to the PFE-SCNL.

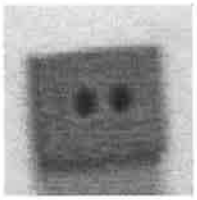
$f(x_{ij}^{Fl}) = 0$ $f(x_{ij}^{Fl}) = 0.1$ $f(x_{ij}^{Fl}) = 0.2$ $f(x_{ij}^{Fl}) = 0.3$ $f(x_{ij}^{Fl}) = 0.4$



$f(x_{ij}^{Fl}) = 0.5$ $f(x_{ij}^{Fl}) = 0.6$ $f(x_{ij}^{Fl}) = 0.7$ $f(x_{ij}^{Fl}) = 0.8$

FIGURE 4.70. Filtering a 64x64 image by a 28x28 PFE-SCNL: Example 3. Parameters as for Figure 4.68.

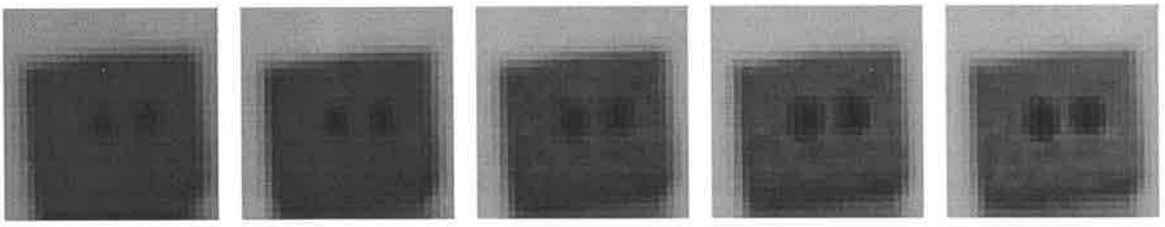
Original 256x256 image of an office scene showing a light switch and a partially occluded power point.



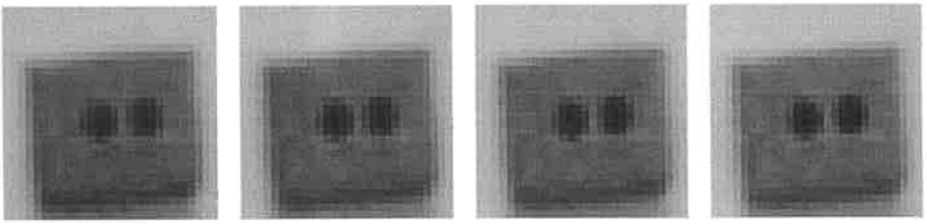
(a) Light switch

(b) Power point

Selected 64x64 regions used as the input to the PFE-SCNL.

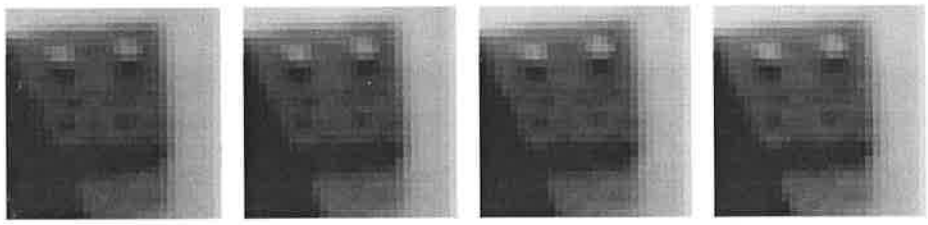


$f(x_{ij}^{F1}) = 0$ $f(x_{ij}^{F1}) = 0.1$ $f(x_{ij}^{F1}) = 0.2$ $f(x_{ij}^{F1}) = 0.3$ $f(x_{ij}^{F1}) = 0.4$



$f(x_{ij}^{F1}) = 0.5$ $f(x_{ij}^{F1}) = 0.6$ $f(x_{ij}^{F1}) = 0.7$ $f(x_{ij}^{F1}) = 0.8$

$f(x_{ij}^{F1}) = 0$ $f(x_{ij}^{F1}) = 0.1$ $f(x_{ij}^{F1}) = 0.2$ $f(x_{ij}^{F1}) = 0.3$ $f(x_{ij}^{F1}) = 0.4$



$f(x_{ij}^{F1}) = 0.5$ $f(x_{ij}^{F1}) = 0.6$ $f(x_{ij}^{F1}) = 0.7$ $f(x_{ij}^{F1}) = 0.8$

FIGURE 4.71. Filtering a 64x64 image by a 28x28 PFE-SCNL: Example 4. Parameters as for Figure 4.68.

In the next example we use an edge processed image as the input to the layer. The shown image was processed by a 3x3 Sobel edge operator and was then scaled to 32x32 elements, which therefore reduces the resolution by a factor of 8.

Original 256x256 image region.

Edge processed image of the original (256x256).

Scaled version of the edge processed image used as the input to the layer (32x32).

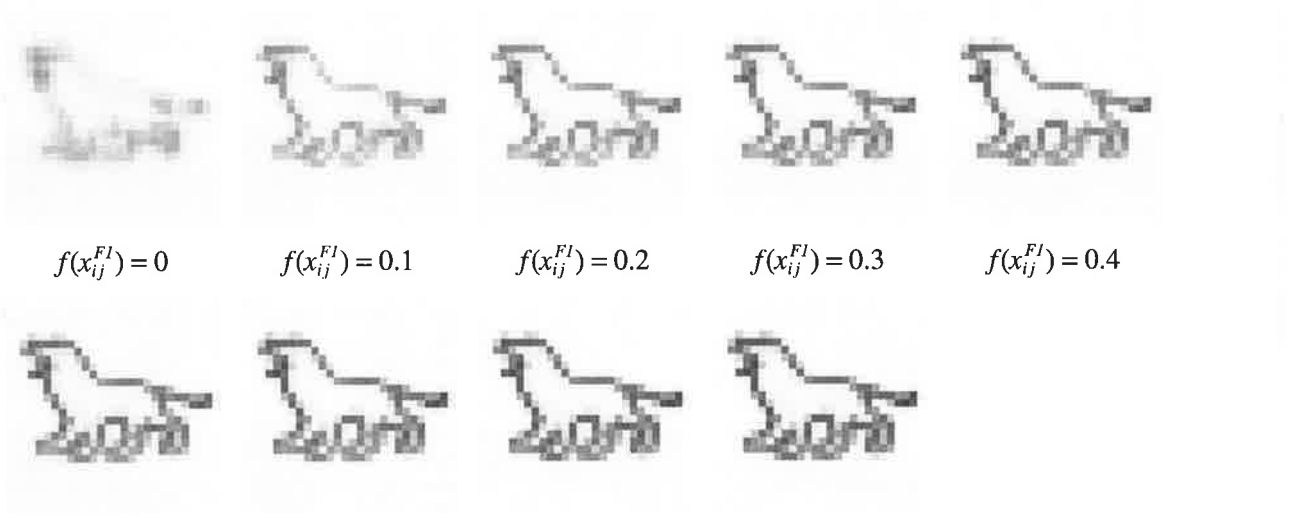


FIGURE 4.72. Filtering a 32x32 edge processed image by a 28x28 PFE-SCNL. Parameters as for Figure 4.68.

In addition to demonstrating the effect of a neuronal mechanism of facilitation via facilitatory modulating fields on the contrast and the resolution of the processed images, these simulations also demonstrate the robustness of the designed layer on a wide variety of input images. In the next section we show how more complex types of input receptive fields may be modulated to produce new types of PFE-SCNLs for use in edge and contrast detection, as well as in the detection of more complex input features.

4.10.2 Modulation of Bipolar Receptive Fields

Unlike the simple unipolar and isotropic receptive fields of the previous section, bipolar receptive fields are characterized by excitatory and inhibitory synaptic inputs which therefore model more complex feature detecting cells of the visual cortex. Typical examples of bipolar receptive fields that have found numerous applications in image processing and computer vision are the two dimensional Gabor functions (Daugman, 1985) and the standard Laplacian of a Gaussian or its approximation by the difference of two gaussians (i.e., the DoG filter of Marr and Hildreth, 1980), where the 2-D gaussian of the inhibitory region has a larger spatial extent. Since they use inhibitory and excitatory regions that may be elongated in various directions, bipolar receptive fields provide a very rich variety of cell types. Below we will consider three special cases of modulated bipolar receptive field profiles (RFs): (i) isotropic; (ii) oriented and even; and (iii) oriented and odd bipolar RFs.

(i) Isotropic Bipolar Receptive Fields

Figure 4.73 represents the simplest case of a cell that samples its input via distributed inhibitory and excitatory synapses. The cell thus samples its excitatory and inhibitory inputs via spatially distributed and isotropic input receptive fields. The spatial extent of the inhibitory input receptive field (IIRF), defined by a two-dimensional gaussian $G_{IIRF}(\sigma^-)$ is generally larger than the spatial extent of the excitatory input receptive field (EIRF), defined by $G_{EIRF}(\sigma^+)$.

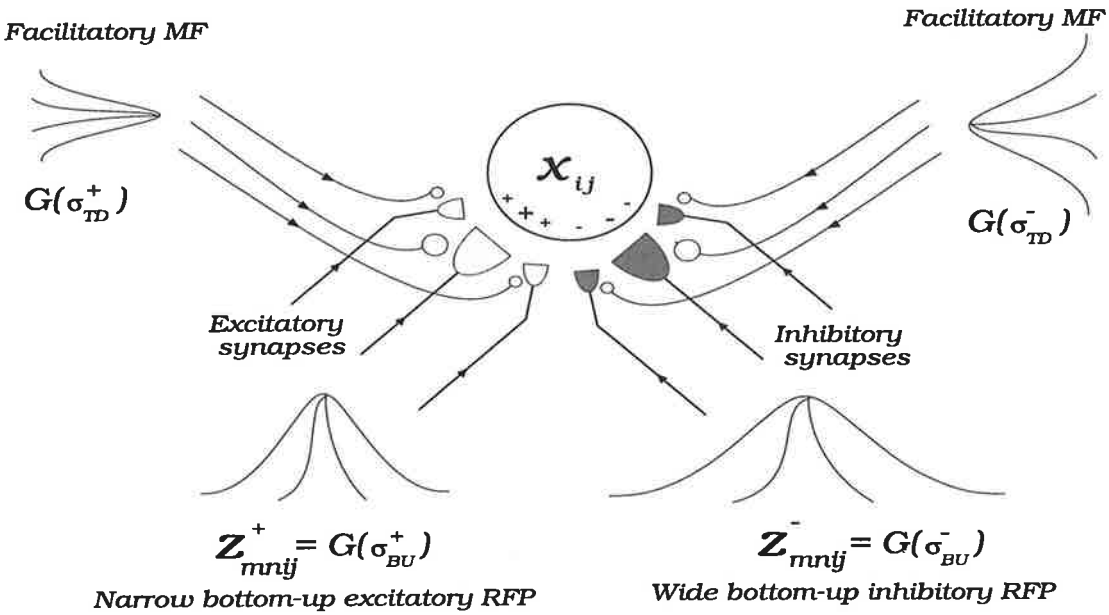


FIGURE 4.73. One dimensional schematic of a neuronal mechanism for the modulation of isotropic bipolar 2-D input receptive field profiles.

The receptive field of such a cell may thus be represented by the difference of the two gaussians, i.e., $DoG(\sigma_{ij}^{+}, \sigma_{ij}^{-}) = G_{EIRF}(\sigma_{ij}^{+}) - G_{IIRF}(\sigma_{ij}^{-})$, where the gaussian of each field specifies the transmitter production level for each synapse as a function of its distance from the cellular centroid. The subscripts (i,j) in σ_{ij}^{+} and σ_{ij}^{-} are used to indicate that in general the spatial extent of the gaussians may also be a function of the cell's position in the input array. The transmitter production levels for the two input RFs at cellular location (i,j) may be written as:

$$z_{mnij}^{+} = Z_{ij}^{+} \exp\left(-\frac{(i-m)^2 + (j-n)^2}{(\sigma_{ij}^{+})^2} \right) \tag{4.59}$$

$$z_{mnij}^{-} = Z_{ij}^{-} \exp\left(-\frac{(i-m)^2 + (j-n)^2}{(\sigma_{ij}^{-})^2} \right) \tag{4.60}$$

where (i, j) is the cell's centroid (i.e., the location of the cell in the 2-D neural layer), while Z_{ij}^{+} and Z_{ij}^{-} define the maximum amplitudes of the gaussians (these may in general be also dependant on the location of the cell within a 2-D neural layer, but are presently assumed to be constant across the layer). If we choose $Z_{ij}^{+} > Z_{ij}^{-}$ and $\sigma_{ij}^{+} < \sigma_{ij}^{-}$, then we are specifying the conventional DoG filter.

As shown in the above figure, modulation of isotropic bipolar receptive fields, such as the DoG, may be achieved by two independent facilitatory modulating fields (FMFs), one for each of the input receptive fields, whose spatial extent is smaller than the IRFs they modulate. Let F_{mnij}^{+} and F_{mnij}^{-}, given by the following expressions, represent the facilitatory modulating fields for the excitatory input RF and the inhibitory input RF.

$$F_{mnij}^{+} = R_{ij}^{+} \exp\left(-\frac{(i-m)^2 + (j-n)^2}{(\psi_{ij}^{+})^2} \right) \tag{4.61}$$

$$F_{mnij}^{-} = R_{ij}^{-} \exp\left(-\frac{(i-m)^2 + (j-n)^2}{(\psi_{ij}^{-})^2} \right) \tag{4.62}$$

where R_{ij}^{+} and R_{ij}^{-} specify their maximum amplitude (which may be a function of the location in the 2-D array), while ψ_{ij}^{+} and ψ_{ij}^{-} specify their spatial extents respectively. These may also be a function of location in the array. For effective modulation of the input RFs, it is preferable to have $(\psi_{ij}^{+}, \psi_{ij}^{-}) < (\sigma_{ij}^{+}, \sigma_{ij}^{-})$, i.e., the facilitatory fields are generally considered to be smaller than the corresponding input receptive fields. The

following equations specify the synaptic dynamics for each input RF, where positive and negative superscripts on synaptic variables and constants are now used to differentiate between the excitatory and inhibitory synapses.

Mobilized transmitter in the excitatory input synaptic pathways (y^+_{mnij})

$$\frac{dy^+_{mnij}}{dt} = [\beta^+_y + \lambda^+ F^+_{mnij} f(x^{FI}_{ij})]\,(u^+_{mnij} - y^+_{mnij})$$

$$- \frac{J_{mnij}[\rho^+_y + K^+_y f(x_{ij})]\,[y^+_{mnij} - Y]^+}{1 + q^+ W^+_{ij}} - \gamma^+_y\, y^+_{mnij} \tag{4.63}$$

where $W^+_{ij} = \sum_m \sum_n w^+_{mnij}$ represent the total excitatory transmitter that is released onto the postsynaptic cell and serves to normalize the excitatory postsynaptic potential (q^+ is the normalizing gain). This normalization procedure approximates the lateral synaptic competition that has been proposed by Grossberg (1984, 1988) to be involved in self-regulated transmitter release via the lateral diffusion of the released transmitter and subsequent reuptake by synaptic autoreceptors. Thus each synapse will have its rate of transmitter release inhibited by all the activated synapses which ensures that the resultant EPSP is approximately constant regardless of the number of active synapses that are participating in cellular excitation.

The amount of excitatory transmitter that is instantaneously released at each synapse by the action of the input signal is given by

$$w^+_{mnij} = J_{mnij}[\rho^+_y + K^+_y f(x_{ij})]\,[y^+_{mnij} - Y^+]^+ \tag{4.64}$$

Stored transmitter in the excitatory input synaptic pathways (u^+_{mnij})

$$\frac{du^+_{mnij}}{dt} = \alpha^+_u (z^+_{mnij} - u^+_{mnij}) - [\beta^+_u + J_{mnij} K^+_u f(x_{ij})]\,(u^+_{mnij} - y^+_{mnij}) \tag{4.65}$$

Excitatory postsynaptic potential (v^+_{mnij})

$$\frac{dv^+_{mnij}}{dt} = -D^+ v^+_{mnij} + \frac{J_{mn}[\rho^+_v + K^+_v f(x_{ij})]\,[y^+_{mnij} - Y]^+}{1 + q^+ W^+_{ij}} \tag{4.66}$$

Next three equations specify the dynamics in the inhibitory synapses.

Mobilized transmitter in the inhibitory input synaptic pathways (y^-_{mnij})

$$\frac{dy^-_{mnij}}{dt} = [\beta^-_y + \lambda^- F^-_{mnij} f(x^{F1}_{ij})]\,(u^-_{mnij} - y^-_{mnij})$$

$$- \frac{J_{mnij}[\rho^-_y + K^-_y f(x_{ij})]\,[y^-_{mnij} - Y]^+}{1 + q^- W^-_{ij}} - \gamma^-_y\, y^-_{mnij} \tag{4.67}$$

where $W^-_{ij} = \sum_m \sum_n w^-_{mnij}$ now represents the total inhibitory transmitter that is released onto the postsynaptic cell and serves to normalize the inhibitory postsynaptic potential.

Stored transmitter in the inhibitory input synaptic pathways (u^-_{mnij})

$$\frac{du^-_{mnij}}{dt} = \alpha^-_u(z^-_{mnij} - u^-_{mnij}) - [\beta^-_u + J_{mnij} K^-_u f(x_{ij})]\,(u^-_{mnij} - y^-_{mnij}) \tag{4.68}$$

Inhibitory postsynaptic potential (v^-_{mnij})

$$\frac{dv^-_{mnij}}{dt} = -D^- v^-_{mnij} + \frac{J_{mnij}[\rho^-_v + K^-_v f(x_{ij})]\,[y^-_{mnij} - Y]^+}{1 + q^- W^-_{ij}} \tag{4.69}$$

Figure 4.74 shows the response of the facilitated isotropic bipolar receptive field to increased levels of facilitation for a constant input signal. Plotted is the net postsynaptic potential along a line passing through the central region of the cell. The complete 3-D RF is shown in Figure 4.75. The 3-D plot represents the net synaptic potential of a 65x65 input receptive field for even illumination across the whole input, as a function of the facilitatory magnitude $f(x^{F1}_{ij})$.

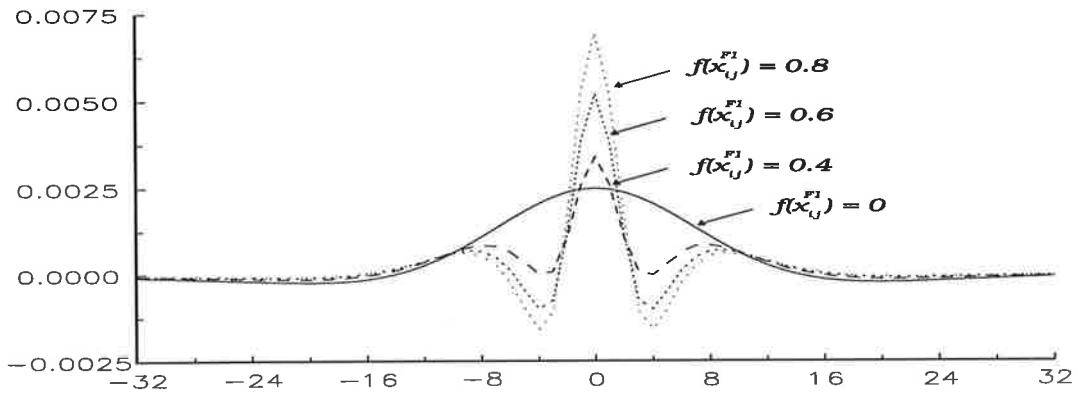


FIGURE 4.74. Net postsynaptic potential along the central axis of a facilitated isotropic bipolar input receptive field (spatial extent 65x65).

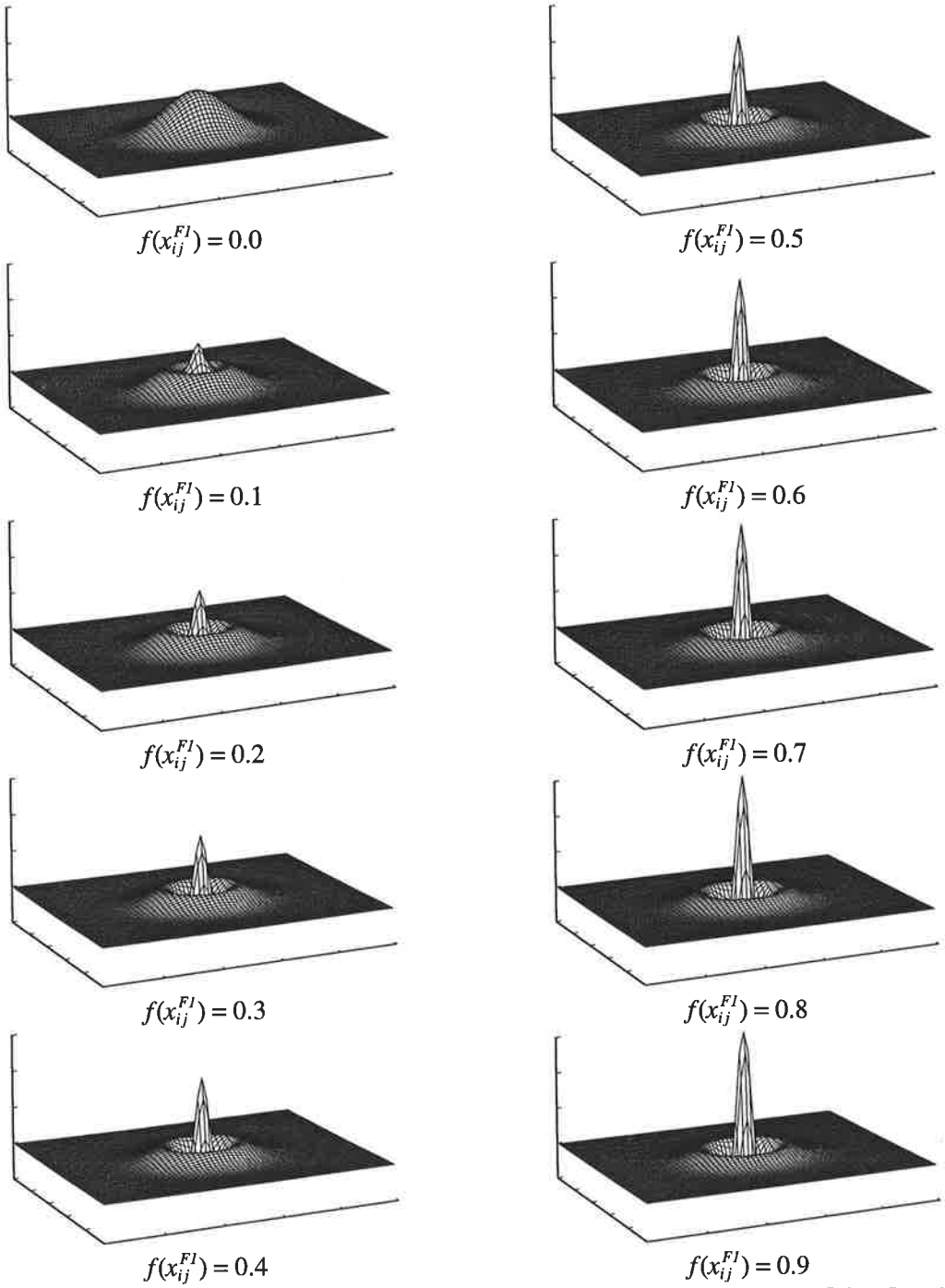


FIGURE 4.75. Three dimensional view of a facilitated isotropic bipolar input receptive field. Parameters: $Z_{ij}^{+} = 1$; $Z_{ij}^{-} = 0.5$; $(\sigma_{ij}^{+})^2 = 100$; $(\sigma_{ij}^{-})^2 = 500$; $(\psi_{ij}^{+})^2 = 5$; $(\psi_{ij}^{-})^2 = 25$; $R_{ij}^{+} = 1$; $R_{ij}^{-} = 0.5$; $\lambda^{+} = \lambda^{-} = 0.05$; $q^{+} = q^{-} = 1$; $\rho_y^{+} = \rho_y^{-} = 0.05$; $\gamma_y^{+} = \gamma_y^{-} = 0.5$; $K_y^{+} = K_y^{-} = 0.5$; $D^{+} = D^{-} = 0.5$; $\rho_v^{+} = \rho_v^{-} = 0.5$; $K_v^{+} = K_v^{-} = 5$; $\alpha_y^{+} = \alpha_u^{-} = 0.05$; $\beta_y^{+} = 0.05$; $\beta_y^{-} = 0.005$ $\beta_u^{+} = \beta_u^{-} = 0.01$; $K_u^{+} = K_u^{-} = 0.1$; cellular threshold $\Theta = 0.1$; threshold for transmitter release $Y = 0$; all inputs $J_{mnij} = 1$. All differential equations are iterated using Euler's first approximation method, with $\Delta t = 1$.

Note that with the exception of: (i) the spatial extent and the amplitude of the gaussians; and (ii) the tonic level of transmitter mobilization ($\beta_y^+ = 10\beta_y^-$), all the parameters for the two types of synapses are equal. The purpose of choosing a higher level of transmitter mobilization in the excitatory synapses was to provide net cellular excitation for a bottom-up input of even illumination across the cellular receptive field ($J_{mnij} = 1, \quad \forall \quad m,n$).

In earlier sections of this chapter (and in Appendix A) we have proposed that the postsynaptic cellular activity in a layer of shunting competitive neurons with bipolar synaptic inputs and a cellular range restricted to (0, B), can be written as:

$$\frac{dx_{ij}}{dt} = -Ax_{ij} + (B - x_{ij})[G(V_{ij}^+ - V_{ij}^-)]^+ - x_{ij}\left([G(V_{ij}^- - V_{ij}^+)]^+ + \overline{G}\overline{v}_{ij}\right) \tag{4.70}$$

where we have said that $[G(V_i^+ - V_i^-)]^+ = \max(G(V_i^+ - V_i^-), 0)$ is the net excitatory postsynaptic potential that drives the cell towards its positive saturation level of B while $[G(V_{ij}^- - V_{ij}^+)]^+ + \overline{G}\overline{v}_{ij}$ is the net inhibitory postsynaptic potential that drives the cell towards zero; G is the gain of excitatory and inhibitory synaptic inputs; $\overline{G}$ is the gain of the lateral competitive feedback inhibition; $V_{ij}^+ = \sum_{(m,n)} v_{mnij}^+$ is the total transmitter gated excitatory postsynaptic potential (EPSP) acting on the cell at location (i,j); $V_{ij}^- = \sum_{(m,n)} v_{mnij}^-$ is the total transmitter gated inhibitory postsynaptic potential (IPSP) acting on the same cell. This equation can be approximated at its equilibrium by

$$x_{ij} = \frac{BG[V_{ij}^+ - V_{ij}^-]^+}{A + G([V_{ij}^+ - V_{ij}^-]^+ + [V_{ij}^- - V_{ij}^+]^+) + \overline{G}\overline{v}_{ij}} \tag{4.71}$$

A Presynaptically Facilitated Excitatory Shunting Competitive Neural Layer (PFE-SCNL) with isotropic bipolar IRFs may be used to detect contrast difference (i.e., object edges) in an input image. For example, if the parameters as chosen such that for even illumination across the input RF the total excitatory input is always matched by the total inhibitory input (regardless of the level of facilitation), then the cell will respond only when there is a contrast difference across its input RF. The output of such a layer of competitive neurons for an input grey level image will be an edge processed image, but whose resolution will depend on the degree of presynaptic facilitation. At low levels of facilitation, the processed output will have coarse edge representation which would become progressively finer as the level of facilitation is increased. Thus, for an input image of an object, the edge processed output along the object's boundary at low level

of facilitation will not contain the fine boundary detail (small curves and corners). The resultant resolution can then be progressively increased by increased levels of facilitation to reveal more accurate boundary details.

Because the bipolar isotropic receptive fields do not distinguish between edges of different orientations, a neural layer with such isotropic properties has limited potential. It therefore becomes necessary to consider the design of a layer whose cells are tuned to various spatial orientations. Although we are currently not concerned with the actual number of such layers that may be needed when designing a machine vision and object recognition system, it does become obvious that even a modest number of between 8-16 directions (as is typically done in some of the current neural network models) places an extremely heavy computational burden on present day workstations. While the need for edge and contrast detection at a number of different orientations increases the complexity and the size of a neural circuit, the fact that each layer may be used for edge detection over various spatial scales may help in keeping the overall size of the circuit somewhat smaller because it reduces the need to perform the same type of computation at a large number of different spatial scales. We are not suggesting that there is no need for simultaneous edge and contrast detection at various spatial scales, but that perhaps a very small number, each being able to be tuned to a variety of spatial scales, may in general be sufficient.

(ii) Oriented Even Bipolar Receptive Fields

An oriented even bipolar receptive field may be also be formed from a combination of 2-D gaussians as in the previous case of isotropic fields, with the exception that the gaussians need to be elongated in a certain direction. The simplest way of generating an elongated 2-D gaussian is by a product of two perpendicular one-dimensional gaussians whose spatial extents differ. For a general case of an elongated RF in direction θ (where θ is taken with respect to a reference, say a horizontal axis), the receptive field of a cell can be written as

$$G(x,y,\theta) = G_0 e^{-\left[\frac{x\cos(\theta)+y\sin(\theta)}{\sigma_x}\right]^2} e^{-\left[\frac{-x\sin(\theta)+y\cos(\theta)}{\sigma_y}\right]^2} \tag{4.72}$$

where θ is defined by σ_x and σ_y (which specify the elongicity of the two 1-D gaussians). For example, a vertically elongated even bipolar input receptive field, denoted by $IRF\left(x,y,even,\frac{\pi}{2}\right)$ can be represented by the following expression:

$$IRF\left(x,y,even,\frac{\pi}{2}\right)=EIRF\left(x,y,even,\frac{\pi}{2}\right)-IIRF\left(x,y,even,\frac{\pi}{2}\right)$$

$$=G_0^+e^{-\left[\frac{y}{\sigma_x^+}\right]^2}e^{-\left[\frac{x}{\sigma_y^+}\right]^2}-G_0^-e^{-\left[\frac{y}{\sigma_x^-}\right]^2}e^{-\left[\frac{x}{\sigma_y^-}\right]^2} \tag{4.73}$$

where G_0^+ and G_0^- define the maximum amplitudes of the gaussians, while (σ_x^+,σ_y^+) and (σ_x^-,σ_y^-) define their spatial extent and elongicity. We can thus represent the transmitter production levels for the excitatory and the inhibitory IRFs by $EIRF(x,y,even,\theta)$ and $IIRF(x,y,even,\theta)$ respectively. Hence, a cell at location (i,j) will have its transmitter production levels specified by the following two equations:

$$z_{mnij}^+=EIRF(x,y,even,\theta)$$

$$=Z_{ij}^+e^{-\left[\frac{x\cos(\theta)+y\sin(\theta)}{\sigma_x^+}\right]^2}e^{-\left[\frac{-x\sin(\theta)+y\cos(\theta)}{\sigma_y^+}\right]^2} \tag{4.74}$$

$$z_{mnij}^-=IIRF^-(x,y,even,\theta)$$

$$=Z_{ij}^-e^{-\left[\frac{x\cos(\theta)+y\sin(\theta)}{\sigma_x^-}\right]^2}e^{-\left[\frac{-x\sin(\theta)+y\cos(\theta)}{\sigma_y^-}\right]^2} \tag{4.75}$$

Facilitation of even bipolar receptive fields may be achieved by two independent facilitatory fields, FMF^+ and FMF^- (one for each of the IRFs), that are more finely tuned than the IRFs they facilitate and whose centroids are also coincident with the cellular centroid. For example, the facilitatory MFs ($FMF^+(x,y,even,\theta)$ and $FMF^-(x,y,even,\theta)$) for the two IRFs ($EIRF(x,y,even,\theta)$ and $IIRF(x,y,even,\theta)$, respectively) can thus be written as

$$F_{mnij}^+(\theta)=FMF^+(x,y,even,\theta) \quad = \quad R_{ij}^+e^{-\left[\frac{x\cos(\theta)+y\sin(\theta)}{\psi_x^+}\right]^2}e^{-\left[\frac{-x\sin(\theta)+y\cos(\theta)}{\psi_y^+}\right]^2} \tag{4.76}$$

and

$$F_{mnij}^-(\theta)=FMF^-(x,y,even,\theta) \quad = \quad R_{ij}^-e^{-\left[\frac{x\cos(\theta)+y\sin(\theta)}{\psi_x^-}\right]^2}e^{-\left[\frac{-x\sin(\theta)+y\cos(\theta)}{\psi_y^-}\right]^2} \tag{4.77}$$

where (ψ_x^+,ψ_y^+) and (ψ_x^-,ψ_y^-) define their spatial extent and elongicity (assumed smaller than for IRFs), while θ specifies their orientation (assumed to be same as for IRFs).

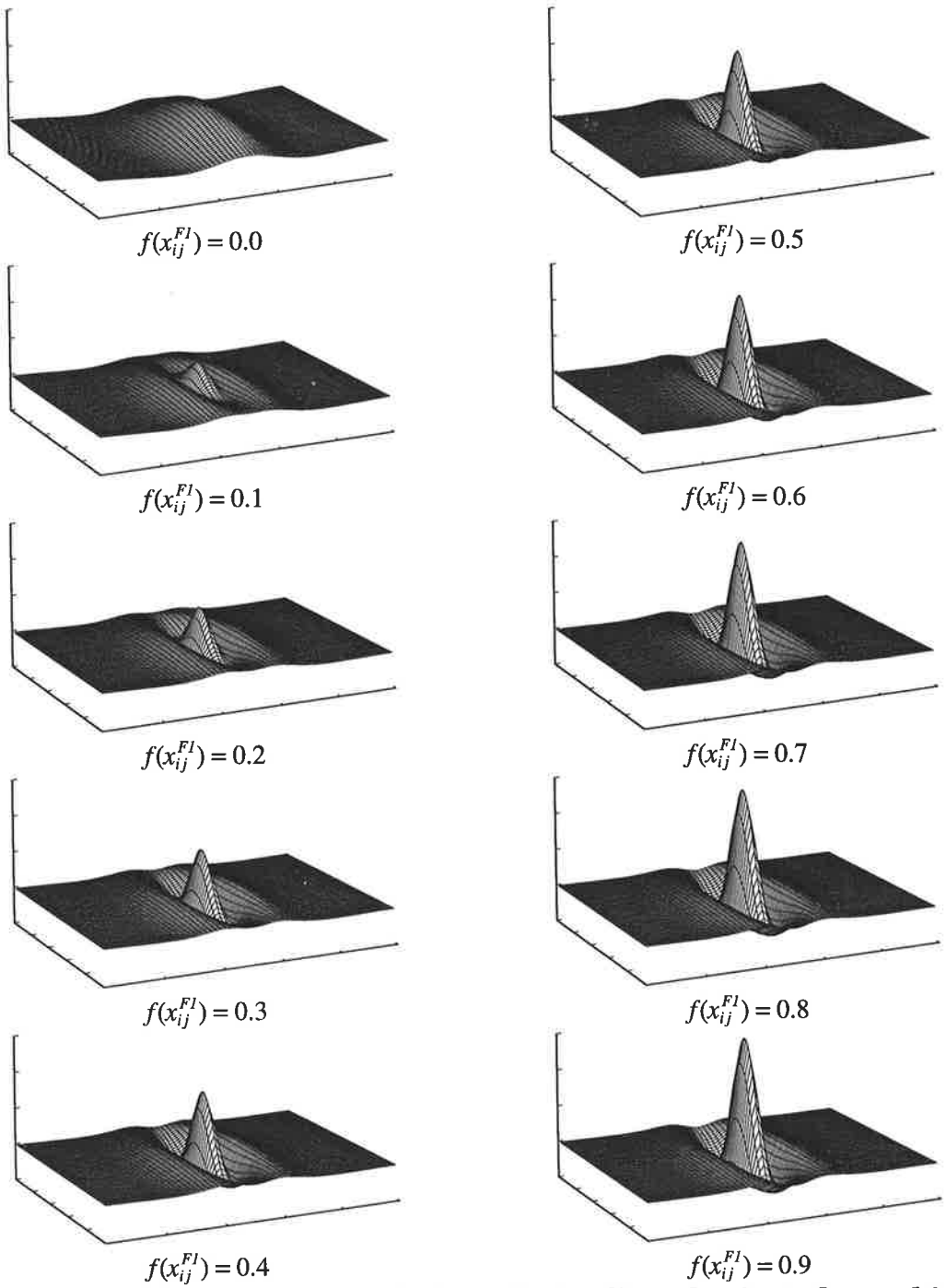


FIGURE 4.76. Three dimensional view of a facilitated oriented even bipolar input receptive field. $Z_{ij}^{-} = 0.5$; $(\sigma_x^{+})^2 = 100$; $(\sigma_y^{+})^2 = 1000$;$(\sigma_y^{-})^2 = 500$; $(\sigma_y^{-})^2 = 5000$; $(\psi_x^{+})^2 = 5$; $(\psi_y^{+})^2 = 100$; $(\psi_x^{-})^2 = 25$; $(\psi_y^{-})^2 = 500$; $R_{ij}^{+} = 1$; $R_{ij}^{-} = 0.5$; $\lambda^{+} = \lambda^{-} = 0.05$; $q^{+} = q^{-} = 1$; $\rho_y^{+} = \rho_y^{-} = 0.05$; $\gamma_y^{+} = \gamma_y^{-} = 0.5$; $K_y^{+} = K_y^{-} = 0.5$; $D^{+} = D^{-} = 0.5$; $\rho_v^{+} = \rho_v^{-} = 0.5$; $K_v^{+} = K_v^{-} = 5$; $\alpha_u^{+} = \alpha_u^{-} = 0.05$; $\beta_y^{+} = 0.05$; $\beta_y^{-} = 0.005$ $\beta_u^{+} = \beta_u^{-} = 0.01$; $K_u^{+} = K_u^{-} = 0.1$; cellular threshold $\Theta = 0.1$; all inputs $J_{mnij} = 1$. All differential equations are iterated using Euler's first approximation method, with $\Delta t = 1$.

When the oriented even bipolar receptive field is used in the input stage of a PFE-SCNL, cells within the layer become selective to the strength of its inputs but along a certain direction. Although not widely used in traditional image processing systems (with the exception of the current applications of 2-D Gabor functions), these bipolar receptive fields, having an elongated excitatory central zone, may be used to detect illuminations of an image in a particular orientation (i.e, 'bar' detection). Their role in biological visual systems is not yet clear.

Figure 4.76 shows the response of the modulated and even bipolar receptive field to various levels of facilitation. The input is held constant across the whole input field, while the strength of the facilitation is slowly raised (in steps of 0.1) every 20 iterations of the synaptic variables (i.e., after a steady state is reached for each level of facilitation).

(iii) Oriented Odd Bipolar Receptive Fields

Oriented odd bipolar input receptive fields may also be generated by a combination of elongated gaussians, but unlike the case of even receptive fields, the inhibitory and the excitatory input RFs need to be displaced in the opposite directions from the cellular centroid (and perpendicular to their major axis), as shown in Figure 4.77.

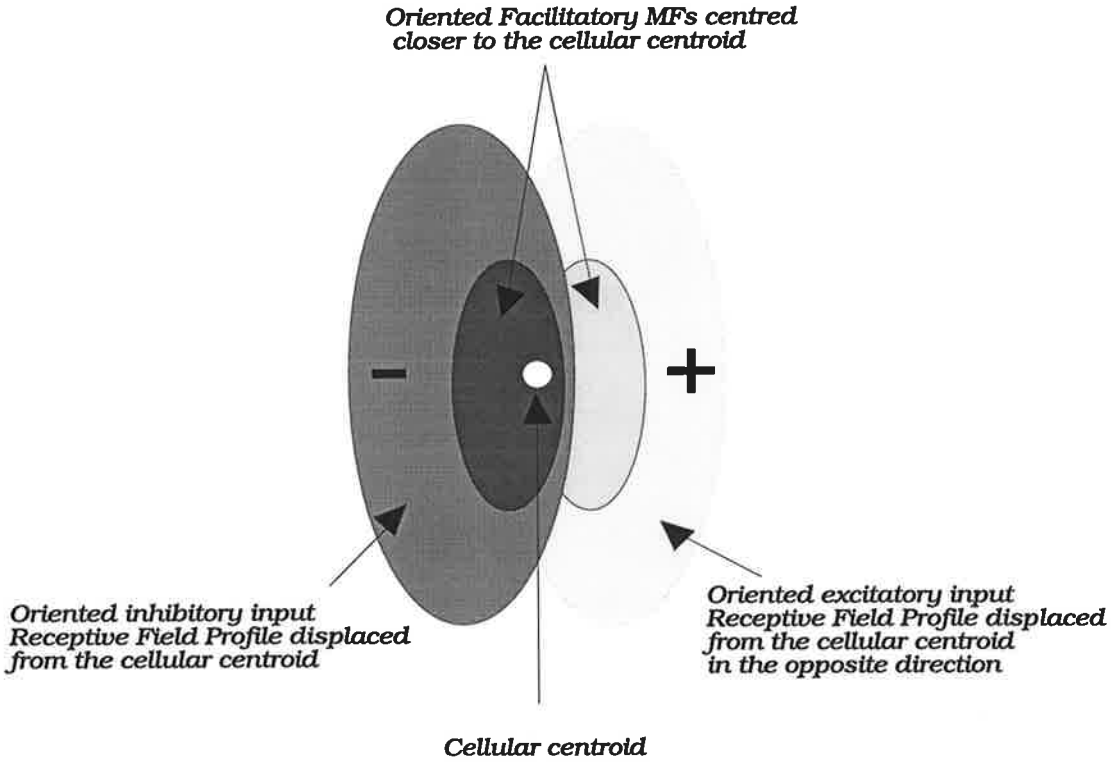


FIGURE 4.77. Two dimensional schematic of a neuronal mechanism for the modulation of oriented odd bipolar receptive fields with offset input and offset facilitatory receptive fields.

Modulation of such an input RF may be done by two elongated facilitatory fields whose spatial extents are smaller than the input RFs they facilitate and whose centroids are also closer to the cellular centroid, as shown in the figure. The transmitter production level for the two input RFs that are oriented in direction θ can thus be written in general terms as follows:

$$z^{+}_{mnij}(\theta) = EIRF(x - x_0, y - y_0, odd, \theta)$$

$$= Z^{+}_{ij}(\theta) e^{-\left[\frac{(x-x_0)\cos(\theta)+(y-y_0)\sin(\theta)}{\sigma^{+}_{x}}\right]^2} e^{-\left[\frac{-(x-x_0)\sin(\theta)+(y-y_0)\cos(\theta)}{\sigma^{+}_{y}}\right]^2} \tag{4.78}$$

$$z^{-}_{mnij}(\theta) = IIRF(x + x_0, y + y_0, odd, \theta)$$

$$= Z^{-}_{ij}(\theta) e^{-\left[\frac{(x+x_0)\cos(\theta)+(y+y_0)\sin(\theta)}{\sigma^{-}_{x}}\right]^2} e^{-\left[\frac{-(x+x_0)\sin(\theta)+(y+y_0)\cos(\theta)}{\sigma^{-}_{y}}\right]^2} \tag{4.79}$$

where (x_0, y_0) is the cellular centroid.

When $\theta = 0$ and when the two input receptive fields are equally displaced from the cellular centroid, but in opposite directions, these equations can be written as

$$z^{+}_{mnij} = Z^{+}_{0} \exp\left(-\left[\frac{(i + x^{+}_{0} - m)}{\sigma^{+}_{x}}\right]^2\right) \exp\left(-\left[\frac{(j + y^{+}_{0} - n)}{\sigma^{+}_{y}}\right]^2\right) \tag{4.80}$$

$$z^{-}_{mnij} = Z^{-}_{0} \exp\left(-\left[\frac{(i - x^{-}_{0} - m)}{\sigma^{-}_{x}}\right]^2\right) \exp\left(-\left[\frac{(j - y^{-}_{0} - n)}{\sigma^{-}_{y}}\right]^2\right) \tag{4.81}$$

Oriented odd bipolar IRFs may be facilitated by two RFs that are also elongated (with their major axis of elongation being aligned with IRFs), but whose spatial extent is smaller and whose centroids are closer to the cellular centroid.

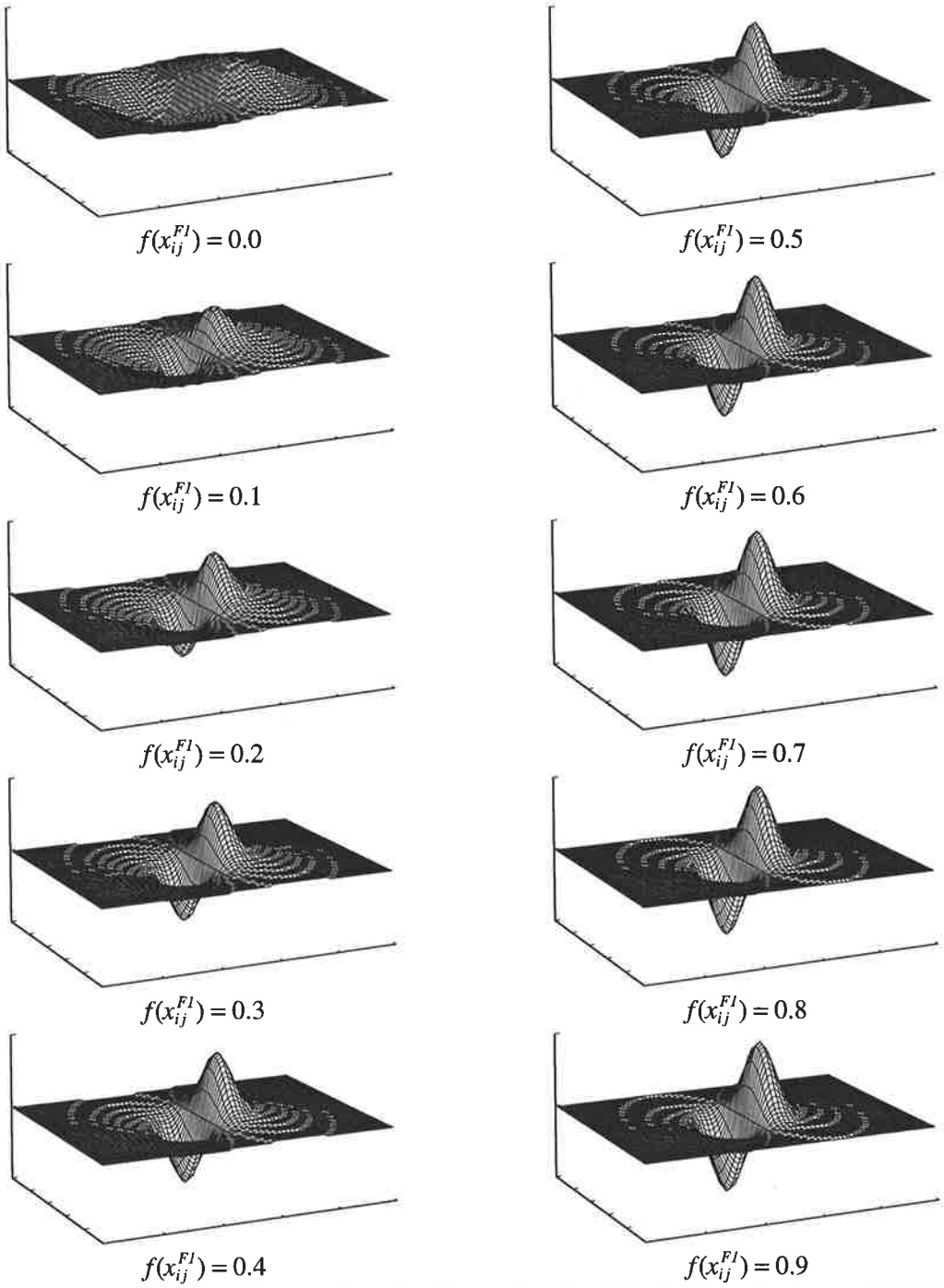


FIGURE 4.78. Three dimensional view of a facilitated oriented odd bipolar input receptive field. $Z_{ij}^{-} = 0.5$; $(\sigma_x^{+})^2 = 100$; $(\sigma_y^{+})^2 = 1000$;$(\sigma_y^{-})^2 = 500$; $(\sigma_y^{-})^2 = 5000$; $(\psi_x^{+})^2 = 5$; $(\psi_y^{+})^2 = 100$; $(\psi_x^{-})^2 = 25$; $(\psi_y^{-})^2 = 500$; $R_{ij}^{+} = 1$; $R_{ij}^{-} = 0.5$; $\lambda^{+} = \lambda^{-} = 0.05$; $q^{+} = q^{-} = 100$; $\rho_y^{+} = \rho_y^{-} = 0.05$; $\gamma_y^{+} = \gamma_y^{-} = 0.5$; $K_y^{+} = K_y^{-} = 0.5$; $D^{+} = D^{-} = 0.5$; $\rho_v^{+} = \rho_v^{-} = 0.5$; $K_v^{+} = K_v^{-} = 5$; $\alpha_u^{+} = \alpha_u^{-} = 0.05$; $\beta_y^{+} = \beta_y^{-} = 0.01$; $\beta_u^{+} = \beta_u^{-} = 0.05$; $K_u^{+} = K_u^{-} = 0.1$; cellular threshold $\Theta = 0.1$; all inputs $J_{mnij} = 1$. All differential equations are iterated using Euler's first approximation method, with $\Delta t = 1$. The two input receptive fields are displaced from the cellular centroid by 10 synaptic pathways.

Unlike the case of even receptive fields, modulation of elongated odd bipolar receptive fields provide a layer of competitive cells the capability to be tuned to contrast differences in its input along a certain direction. If such a layer was used to process a gray level image, then the response of the layer will represent all edges that are oriented in the same direction as the cellular RFs. However, at low levels of facilitation, the receptive fields (being reasonably large with crude orientational selectivity) will similarly give a crude edge representation of the input image. Fine detailed edge structure will thus not be seen at the output of the layer. The layer will in general respond to image contrasts over broad input region and in directions whose angular displacement from cell's optimum orientation is reasonably large. Fine tuning the cellular receptive field will have two effects. Since increased facilitation shifts the peaks of the excitatory and the inhibitory maxima closer to the cellular centroid, contrast differences will be detected over smaller spatial extent. Simultaneous with this reduction of the spatial extent is the narrowing of the two input RFs. The net effect is that the cell becomes progressively tuned to higher spatial frequency and a narrower range of contrast directions.

It is thus possible to design more complex modulated RFs simply by a suitable arrangement of the individual inhibitory and excitatory zones and their facilitatory modulating fields.

We have shown how presynaptic facilitation may be used to dynamically fine-tune various types of cellular input receptive fields. However, for this dynamic mechanism to be useful, a self-organising artificial neural system (capable of unsupervised learning) must be able to decide autonomously whether to fine-tune its cellular receptive fields, by how much and when. Self initiated regulation of cellular receptive fields, attention, general arousal, etc., may be beneficial to an unsupervised learning system, particularly if it has to deal with complex and cluttered visual inputs where fixed receptive fields may limit its robustness. For an artificial neural system to be able to regulate its receptive fields appropriately, it must have some rudimentary form of 'consciousness' or 'awareness'. This is unlike the case for the currently popular artificial neural networks, whose inability of self-regulation requires that in most cases human consciousness decides what the network is to learn, when, for how long, at what rate, etc. Once trained, these networks become rigid thereafter and are therefore not able to cope effectively with more complex and cluttered inputs than the ones that were used during training. In order to enable a self-organising artificial neural system to self-regulate its dynamics and cellular receptive field profiles, we must first discover the relevant neuro-engineering design principles. The simple neural circuit of *Aplysia*, which has inspired most of the neural circuits and mechanisms that were proposed so far in this thesis, is too simple a system to exhibit the type of new mechanisms that are needed. One can hardly call the facilitation of the sensory to motor neuron synapse by

the facilitatory interneurons from the tail sensory neurons as being due to a high level memory based decision from elsewhere within the circuit. We must therefore find a neuro-engineered solution to the problem of self-regulation and investigate its potential on problems that cannot be handled robustly with neural circuits that do not posses such capability. The scientific study of the neuronal basis of awareness and consciousness has just recently begun and forms the frontier of current neuroscientific and theoretical investigations (Crick and Koch, 1990a, 1990b, 1992; Koch and Crick, 1994). In Chapter 6 we will propose a self-regulating mechanism for a simple neural circuit that uses its established memory to decide whether its input is familiar or novel.

Throughout this section we have concentrated only on the facilitatory mechanisms for the modulation of cellular receptive field profiles and have also assumed that the facilitation was non-specific (i.e., each cell in the layer received the same level of facilitation). We have also considered a simple case of facilitation where the facilitatory MFs were similar to their IRFs. It is, however, also possible for FMFs and IRFs to be very different. In Chapters 6, 7 and 8 we will propose more complex neural circuits that use the types of synapses discussed above. In addition, we will propose a neural mechanism for self-regulation as well as new types of modulatory fields that may be very different to their IRFs and each being able to convey specific 2-D spatial information, such as a complete boundary of an object.

4.11 Recognition of Distorted 2-D Shapes

In section 4.8 we have proposed a two layered neural circuit called Feedforward Excitation-Feedback Presynaptic Facilitation (FFE-FBPF) whose cells had only one input synapse. We now propose a simple extension to the circuit that provides it with additional processing capability, such as the recognition of distorted 2-D shapes. Similar capability has already been provided to ART-2 and Fuzzy ARTMAP (Seibert and Waxman, 1992; Gove et al., 1995) by sampling the bottom-up inputs via 2D receptive fields (i.e., coarse coding of the input). However, the scheme that we are proposing below is based on modulated 2D receptive fields of section 4.10. First we propose a useful neuroengineering design principle called "***Principle of biological implausibility***" which says that if an artificial neural network has layers whose cells sample their inputs via a single synaptic pathway, then this will in some way limit the capability of the network. This design principle (whether biologically plausible or not) is a useful design guide since it allows one to quickly spot those parts of the network design that may need further development.

The extended circuit, shown in Figure 4.79, samples its bottom-up input via excitatory input receptive fields whose synapses are presynaptically modulated by top-down facilitatory fields that are more finely tuned, the purpose of which was discussed in the previous section. Each cell in the first PFE-SCNL or Field F0 at location (i,j) samples its bottom-up input via a set of dynamic synaptic pathways whose level of transmitter production is defined by a 2-D gaussian. Each cell in Field F1 at location (i,j) backprojects a 2-D facilitatory field to a cell in the corresponding location of Field F0. As shown in Figure 4.80, the top-down reference into Field F1 and the distributed sampling by cells in F0 enables the FFE-FBPF neural circuit to test for the presence of the target shape in a wider range of locations than is possible by single input synapses. If the target shape is located within the facilitated bottom-up input receptive fields, it will be recognized even if its boundary does not fall within the central region of the input receptive fields. This enables the circuit to recognize distorted 2-D shapes.

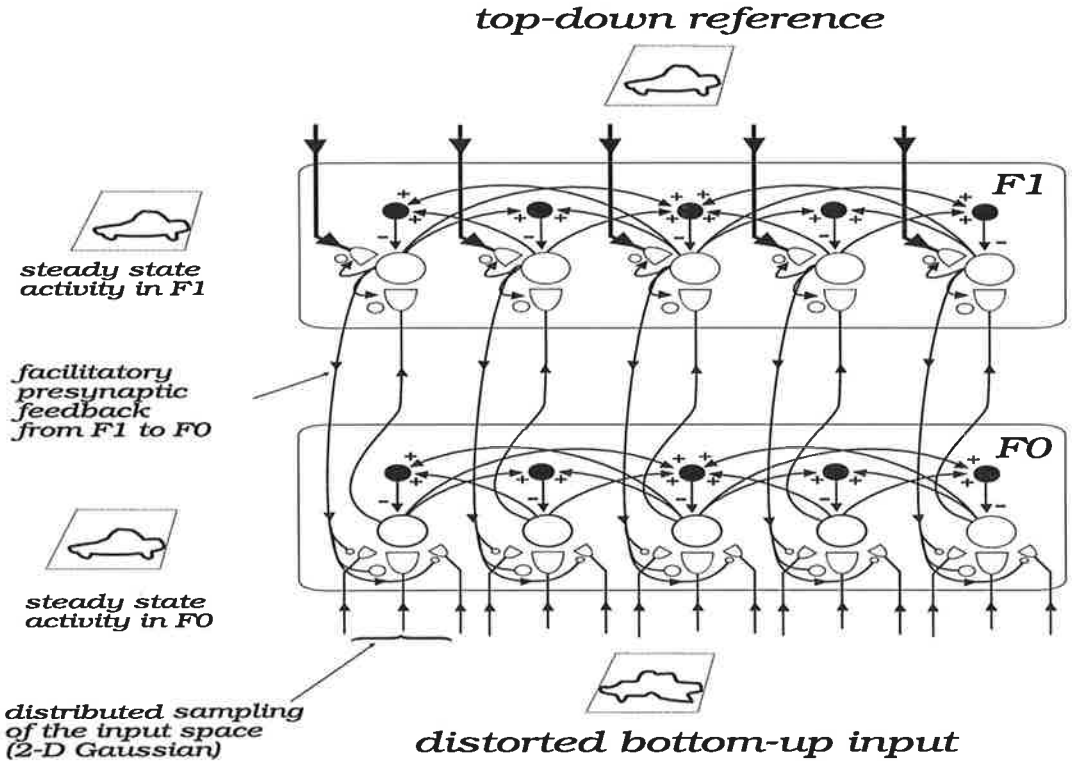


FIGURE 4.79. Extended Feedforward Excitation-Feedback Presynaptic Facilitation (FFE-FBPF) neural circuit for the recognition of distorted 2-D shapes.

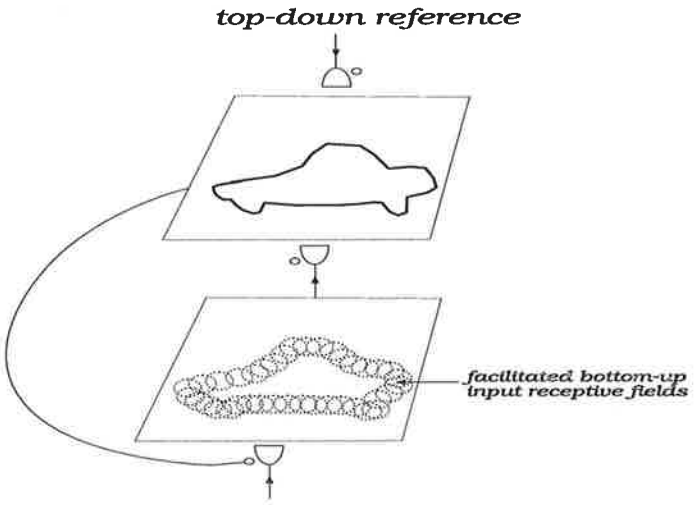


FIGURE 4.80. Distributed sampling enables the FFE-FBPF circuit to test for the presence of the relevant 2-D shape within the facilitated input receptive fields.

To demonstrate the capability of the proposed neural circuit we provide computer simulations for three cases: (i) with a top-down reference shape into Field F1; (ii) without the top-down reference; and (iii) with a top-down reference that is generated by first presenting the target shape to the circuit and then using the resultant steady state spatial pattern as the top-down reference. For each of the simulations, the input array is 40x40 elements, while the two PFE-SCNLs in the circuit are both 34x34 cells in size. All network variables were re-initialized upon each new input presentation.

Each input shown in the left column of Figure 4.81 was presented to the extended FFE-FBPF circuit for 50 iterations. The resultant steady state spatial patterns across Fields F0 and F1 are shown in the second and the third column respectively.

TABLE 4.1: Receptive and Facilitatory Fields in FFE-FBPF circuit

0.018316	0.082085	0.135335	0.082085	0.018316
0.082085	0.367879	0.606531	0.367879	0.082085
0.135335	0.606531	1.000000	0.606531	0.135335
0.082085	0.367879	0.606531	0.367879	0.082085
0.018316	0.082085	0.135335	0.082085	0.018316

Receptive field

0.000335	0.006738	0.018316	0.006738	0.000335
0.006738	0.135335	0.367879	0.135335	0.006738
0.018316	0.367879	1.000000	0.367879	0.018316
0.006738	0.135335	0.367879	0.135335	0.006738
0.000335	0.006738	0.018316	0.006738	0.000335

Facilitatory field

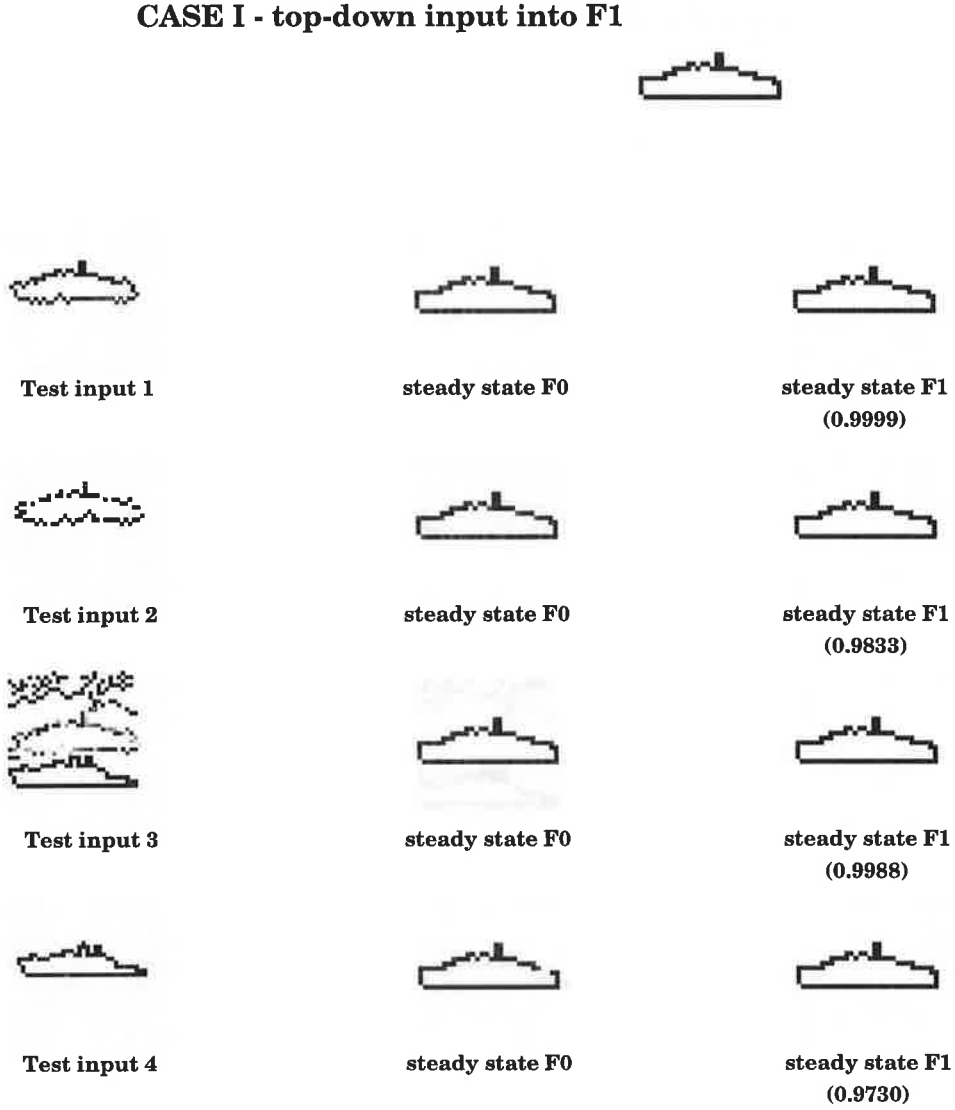


FIGURE 4.81. Simulation of the extended FFE-FBPF neural circuit on distorted 2-D shapes. Case I: with a top-down reference. The data is shown in reverse contrast (black = 1.0, white = 0).

The circuit parameters are as follows (refer to equations in section 4.8 and Chapter 5): $A^{F0} = B^{F0} = A^{F1} = B^{F1} = 1$; $\overline{A}^{F0} = \overline{A}^{F1} = 0.1$; threshold for transmitter release $Y = 0.001$; post-synaptic threshold $\Theta = 0.1$; $\overline{B}^{F0} = \overline{B}^{F1} = 0.1/n$ where n is the number of neurons in the layer ($n = 1156$); $\alpha_u^{F0} = \alpha_u^{F1} = 0.5$; $\beta_u^{F0} = \beta_y^{F0} = 0.005$; $\beta_u^{F1} = \beta_y^{F1} = 0.05$; $\gamma_y^{F0} = \gamma_y^{F1} = 0.5$; $D^{F0} = D^{F1} = 0.5$; gain of top-down presynaptic facilitation $H = 0.99$; $K_u^{F0} = K_u^{F1} = 0.5$; $\rho_v^{F0} = \rho_v^{F1} = 0.5$; $\rho_y^{F0} = \rho_y^{F1} = 0.05$; $K_y^{F0} = K_y^{F1} = 0.5$; $K_v^{F0} = 10$; $K_v^{F1} = 5$; postsynaptic gains are $G^{F0} = G^{F1} = 1000$; gain of lateral feedback inhibition $\overline{G}^{F0} = \overline{G}^{F1} = 5000$. The spatial extent of the input receptive field and the facilitatory field into F0 is 5x5 synapses, the latter being more finely tuned. Table 4.1 shows the elements of the two 2-D gaussians that define each of the fields ($\sigma^2 = 2$ for the input receptive field, $\psi^2 = 1$ for the facilitatory field).

The bracketed numbers in the rightmost column indicate the steady state match between the spatial patterns across Fields F0 and F1 (the required match for recognition is ≥ 0.9800; steady state is assumed when the rate of change of the match falls below 0.0001). Note that all distorted versions of the reference shape were recognized because of the additional bottom-up synapses through which the relevant parts of the input may be routed. The shape embedded in test input 4 did not cause a match above the required threshold because some parts of its boundary fell beyond the facilitated receptive fields. The circuit is also tolerant to small variations in the position and the size of the input. The degree of tolerance is controlled by the spatial extent of the input receptive fields. The size of the input receptive fields in the shown simulation is 5x5 (see Table 4.1 for the actual values). Although a broader input receptive field would provide a higher degree of tolerance to distortions and positional displacements, it would also lead to greater number of discrimination errors between different shapes that share many common features.

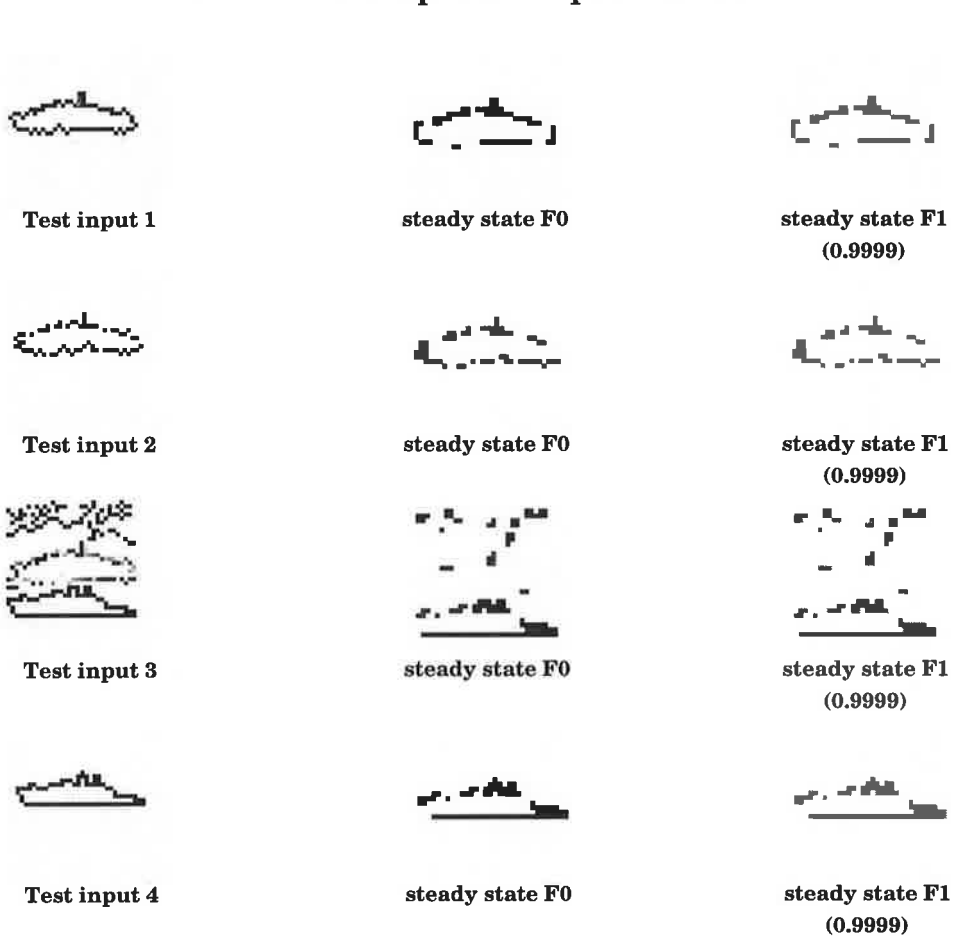


FIGURE 4.82. Simulation results of the extended FFE-FBPF neural circuit on distorted 2-D shapes. Case II: without the top-down reference.

Figure 4.82 shows that in the absence of a top-down reference, the steady state depends solely on the strength of the bottom-up inputs. As expected, weak bottom-up inputs do not enter into the circuit because of the competitive effects in each neural layer. If weak bottom-up inputs are facilitated by a top-down reference (as in the case of input 3, Figure 4.82), then weak and relevant inputs can re-enter into the circuit at the expense of strong and non-relevant inputs. The above simulation data also shows that some diagonal elements of the various inputs did not enter into the circuit. The reason for this is given below.

CASE III - top-down reference generated by the circuit

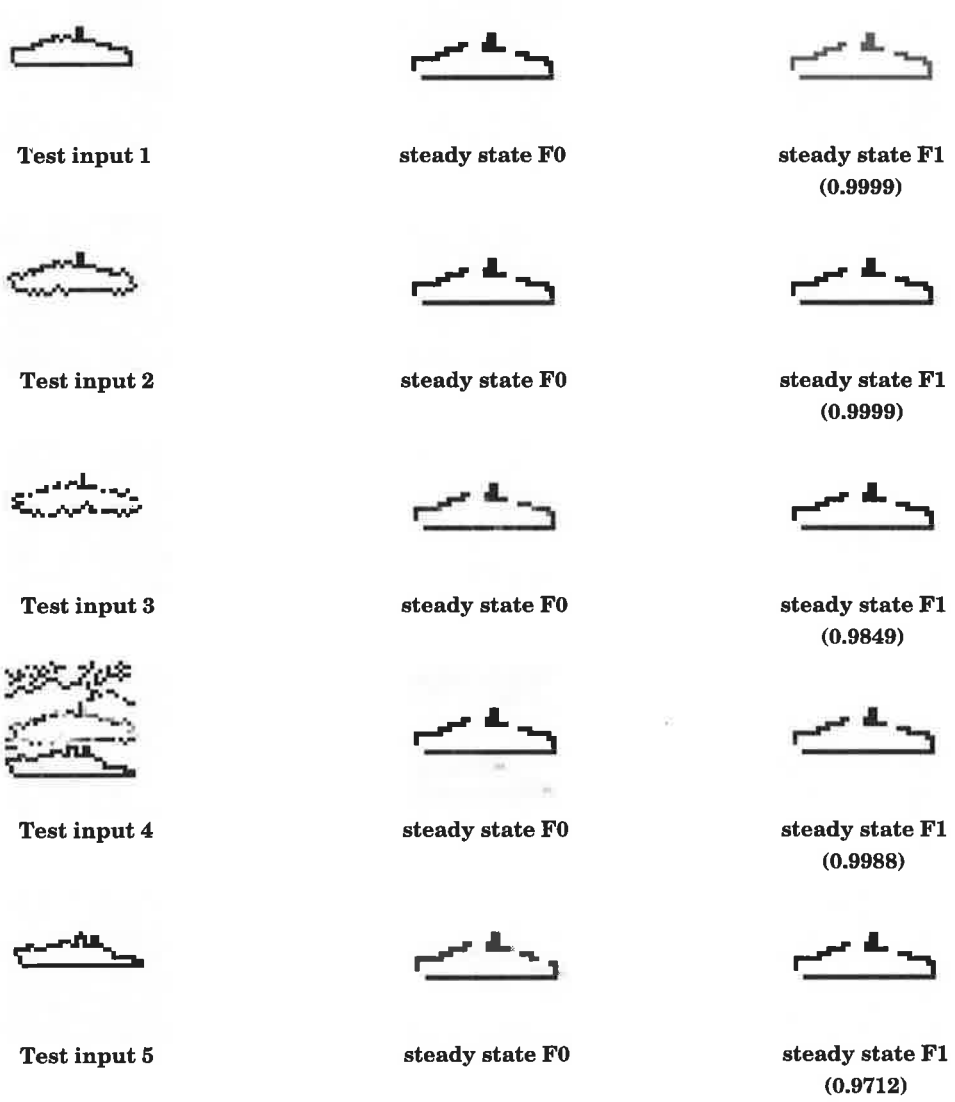


FIGURE 4.83. Simulation of the extended FFE-FBPF neural circuit on distorted 2-D shapes. Case III: top-down reference generated by the circuit.

In the next simulation (Case III), we first present the target shape as the bottom-up input into the circuit and then use the resultant steady state spatial pattern across F1 as the top-down reference. The results are shown in Figure 4.83. Note that the generated top-down reference has some missing parts. The reason for this can be found in Table 4.1. The spatial kernel that defines the input receptive fields, being generated by using the euclidian distances in the 2-D gaussian function, is biased towards synaptic pathways that are offset either vertically or horizontally from the cellular centroid. Vertical and horizontal parts of the input shape therefore provide a larger bottom-up excitation of their target cells than is provided by the diagonally oriented line segments. Cells that are excited by diagonally oriented parts of the input shape do not receive sufficiently high bottom-up excitation to overcome the inhibitory inputs from the more activated cells. Nevertheless, the circuit was still able to recognize the distorted versions of the original shape, as indicated by the degree of match between F0 and F1.

We have thus shown that the solution to distorted 2-D shape recognition can be neuro-engineered into a 2-D neural circuit by providing additional bottom-up input pathways (via a distributed input receptive field), through which the various portions of the input may enter into the circuit. Note that the proposed circuit interactions also simultaneously solve a number of other related problems: recognition in clutter, tolerance to small displacements of the input, tolerance to small variations in the size and the orientation of the input shape, as well as the recognition of the shape when some of its local parts are missing.

From the above simulation data we conclude that in order for the circuit to register diagonal elements it needs to be split into multiple spatial orientations. In Chapter 8 we will consider how the circuit may be extended to cater for size and orientation invariant 2-D shape recognition.

4.12 Conclusions

In this chapter we have proposed novel neural layers, a number of new design principles and mechanisms for the design of 2-D neural circuits. These design principles and mechanisms are based on the properties of shunting competitive neural layers that have not been previously utilised in neural network designs. By realising that the amplification of some inputs into a layer of shunting competitive neurons can competitively suppress the activity of other neurons in the layer whose inputs are left unaltered, we have been able to derive a new type of a competitive neural layer, Presynaptically Modulated Shunting Competitive Neural Layer (PM-SCNL). PM-SCNL uses models of chemical synapses whose transmission gain may be facilitated (or inhibited) by a variety of methods that are either non-specific or are specific for a

particular pattern or spatial location. We have proposed how these new layers may be used in selective information processing. In particular, we have derived a novel neural circuit, Feedforward Excitation-Feedback Presynaptic Facilitation (FFE-FBPF), which uses a top-down reference to selectively modulate the bottom-up signal transmission into the circuit. The potential of the FFE-FBPF neural circuit was demonstrated in an application to object recognition in cluttered visual images.

The principle of synaptic gain control was extended to the general case of cells with multiple input synapses to show how the filtering characteristics of the cellular receptive fields (and thus a layer of neurons) may be altered and fine-tuned by a single parameter. However, we have not yet provided a self-regulating mechanism that should enable an artificial neural circuit (or a network) to autonomously decide when and by how much it should alter its level of attention. Similarly we have not considered how a particular input may be attended and learned when it does not have a top-down reference or is not aligned with the top-down reference and appears in a cluttered background of other objects.

Since the primary aim of this chapter (which also includes Appendix A) was to derive a set of useful and flexible neuro-engineering design principles, mechanisms and building blocks for more complex neural circuits, we have limited our discussion and computer simulations to simple cases in order to reveal the dynamics of selective information processing. Numerous computer simulations have not only demonstrated the potential of the proposed design principles and mechanisms, but have also revealed a host of further problems that need to be solved. For example, the simulation results of the FFE-FBPF neural circuit in an application to object recognition in cluttered visual images suggest that we need to design a parallel set of circuits, such that each circuit is devoted to the processing of edges of a particular spatial orientation (or spatial extent). Although the need for parallel neural layers and circuits may have been anticipated from prior knowledge of the primate visual neurophysiology, it is not yet known how the parallel circuits of the primate visual system are integrated at higher levels to give rise to coherent visual perception of an object's shape. It is debatable whether experimental neurophysiology will ever be able to provide an answer to this "feature integration" problem of visual perception. However, by careful attention to good neuro-engineering design principles and design logic, it is highly likely that the solution to the feature integration problem, as well as the solutions to the other problems of biological vision (some of which were mentioned in the previous paragraph), will reveal itself when one attempts to neuro-engineer an artificial visual object recognition system for cluttered images and scenes.

The rest of the thesis builds onto the concepts proposed in this chapter in order to provide neuro-engineered solutions to the above mentioned and related problems of cognitive biological vision. In the next chapter we provide a mathematical procedure for the design of presynaptically modulated competitive neural layers.

Chapter 5

Mathematical Analysis, Parameter Design and Simulations

5.1 Introduction and Overview

In this chapter we provide a mathematical analysis and a parameter design procedure for facilitated shunting competitive neural layers to ensure stability, wide dynamic range and invariance to the size of the layer. We analyse the stability of a layer under the worst possible condition, i.e., when all the cells are driven by the same input (zero input contrast). If the layer converges to a stable (non-oscillatory) state above the cellular threshold Θ for zero input contrast patterns whose input intensity is J_M, then it will also converge to a stable state for all input patterns whose average intensity J_A is above J_M/n, where n is the number of neurons in the layer. We thus analyse the stability of the layer for a uniform input J_M in the open loop condition (i.e., without the postsynaptic feedback).

5.2 Determination of Steady State

We first find a steady state solution for the case when there is no postsynaptic feedback acting on the synaptic dynamics. We then find the critical value for the inhibitory gain ($\overline{G}_c$) that leads to oscillatory behaviour. After finding $\overline{G}_c$, we then choose $\overline{G} < \overline{G}_c$ (for the case where all cells are equally activated by a maximum possible input J_M) such that it gives as a desired steady state. Since each layer has a large number of parameters (about 15), our design procedure is heavily based on our extensive computer simulations of various layers and their interactions. Note that all neural layers, circuits

and network architectures proposed in the thesis are based on the same set of five basic equations that are listed in this chapter (equations (5.12), (5.13), (5.17), (5.21) and (5.22)).

We thus start by finding a steady state solution to the dynamics of an isolated chemical synapse. After obtaining the steady state solution we then relate the expression to a desired steady state.

Synaptic variables

u_i is the stored transmitter;

y_i is the mobilized transmitter;

z_i is the transmitter production level (assumed fixed or very slow compared to the other two).

Write the synaptic equations in the following form

$$\dot{u}_i = a(z_i - u_i) - c(u_i - y_i) \tag{5.1}$$

$$\dot{y}_i = d(u_i - y_i) - e\,y_i \tag{5.2}$$

In order to solve the above equations, let us initially assume that there is no postsynaptic feedback signal acting on either variable. The effect of pre-postsynaptic interactions will subsumed into the parameter design.

Initial conditions: $u_i(0) = z_i$

$y_i(0) = 0$

Rewrite (5.1) and (5.2) as

$$[D_t + A_1]u_i - c\,y_i = a z_i \tag{5.3}$$

$$-u_i + \frac{1}{d}[D_t + A_2]y_i = 0 \tag{5.4}$$

where $A_1 = a + c$, $A_2 = d + e$, $D_t = \frac{d}{dt}$ operator.

Apply the operator $[D_t + A_1]$ to (5.3) and add to (5.4). Then

$$[D_t^2 + (A_1 + A_2)D_t + (A_1A_2 - cd)]y_i = z_i ad \tag{5.5}$$

which can be written as

$$\ddot{y}_i + (A_1 + A_2)\dot{y}_i + (A_1A_2 - cd)y_i = z_i ad \tag{5.6}$$

This system has a general solution

$$y_i(t) = \frac{z_i ad}{\lambda_1\lambda_2} + c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t} \tag{5.7}$$

where c_1 and c_2 are constants; λ_1 and λ_2 are the roots of

$$\lambda^2 + (a + c + d + e)\lambda + (ad + ae + ce) = 0 \tag{5.8}$$

Using the initial conditions and the method of variation of parameters, it can be shown that

$$y_i(t) = \frac{z_i ad}{\lambda_1\lambda_2} + \frac{z_i d(a + \lambda_1)}{\lambda_1(\lambda_1 - \lambda_2)} e^{\lambda_1 t} + \frac{z_i d(a + \lambda_2)}{\lambda_2(\lambda_2 - \lambda_1)} e^{\lambda_2 t} \tag{5.9}$$

The roots of (5.8) are related by

$$\lambda_1\lambda_2 = ad + ae + ce \tag{5.10}$$

Note that if the roots of (5.8) both have negative real parts (i.e., $Re(\lambda_1) < 0$ and $Re(\lambda_2) < 0$) then the exponential terms in (5.9) will in the limit $t \to \infty$ decay to zero. Therefore the steady state solution of (5.2) is given by

$$\hat{y}_i = \frac{z_i ad}{\lambda_1\lambda_2}$$

$$= \frac{z_i ad}{ad + ae + ce} \tag{5.11}$$

where $\hat{y}_i$ denotes the steady state solution to equation (5.2), where we have assumed that there is no postsynaptic feedback acting on the synapse. Let there be n cells in a layer of shunting competitive neurons whose equilibrium activity is given by (5.12) and whose lateral inhibition is given by (5.13).

$$\hat{x}_i = \frac{BG\hat{v}_i}{A + G\hat{v}_i + \overline{G}\hat{\overline{v}}_i} \tag{5.12}$$

$$\frac{d\overline{v}_i}{dt} = -\overline{A}\overline{v}_i + \frac{1}{n}\overline{B}\sum_{j \neq i} f(x_j) \tag{5.13}$$

The thresholding function in (5.13) is given by $f(x_j) = \max(x_j - \Theta, 0)$ where Θ is the threshold value below which the cell is not contributing to lateral inhibition of the other cells in the layer. In order to design the parameters of the layer such that its cells do not enter into persistent oscillations, we want the equilibrium value of each cell to be above the threshold Θ when each cell is driven by the maximum input J_M.

Let this equilibrium value be

$$\hat{x}_i = \Theta + \delta \tag{5.14}$$

where δ is the amount by which the cell's activity at equilibrium exceeds the threshold. The steady state solution of (5.13) is then given by

$$\hat{v}_i = \delta\left(\frac{\overline{B}}{\overline{A}}\right)\left(\frac{n-1}{n}\right) \tag{5.15}$$

Substituting (5.14) and (5.15) into (5.12) and rearranging gives

$$\hat{v}_i = \left[\frac{\Theta + \delta}{B - (\Theta + \delta)}\right]\left[\left(\frac{\overline{B}}{\overline{A}}\right)\left(\frac{\overline{G}}{G}\right)\left(\frac{n-1}{n}\right)\delta + \frac{A}{G}\right] \tag{5.16}$$

Now, the steady state EPSP, $\hat{v}_i$, can also be derived from

$$\begin{aligned}\frac{dv_i}{dt} &= -Dv_i + J_M[y_i - Y]^+(\rho_v + K_w f(x_i)) \\ &= -Dv_i + J_M[y_i - Y]^+(\rho_v + K_v\delta)\end{aligned} \tag{5.17}$$

where we have said that $f(\hat{x}_i) = \hat{x}_i - \Theta = \delta$; Y is the threshold for transmitter release. At the desired steady state ($\hat{x}_i = \Theta + \delta$), this becomes

$$\hat{v}_i = \frac{J_M}{D}[\hat{y}_i - Y]^+(\rho_v + K_v\delta) \tag{5.18}$$

where $\hat{y}_i$ is the steady state level of the mobilized transmitter. Equating (5.16) and (5.18) gives

$$[\hat{y}_i - Y]^+ = \frac{[D(\Theta + \delta)]\left[\left(\frac{\overline{G}}{G}\right)\left(\frac{\overline{B}}{\overline{A}}\right)\left(\frac{n-1}{n}\right)\delta + \frac{A}{G}\right]}{J_M[B - (\Theta + \delta)]\,[\rho_v + K_v\delta]} \tag{5.19}$$

For simplicity, let us assume that Y = 0. We know that $\hat{y}_i$ in equation (5.11) leads to a steady state when there is no postsynaptic feedback. Hence in order to design the parameters that will force the layer of shunting competitive neurons to reach a steady state at $\hat{x}_i = \Theta + \delta$, we can let

$$\hat{y} = \frac{z_i a d}{ad + ae + ce} = \frac{[D(\Theta+\delta)]\left[\left(\frac{\overline{G}}{G}\right)\left(\frac{\overline{B}}{\overline{A}}\right)\left(\frac{n-1}{n}\right)\delta + \frac{A}{G}\right]}{J_M[B-(\Theta+\delta)]\,[\rho_v + K_v\delta]} \tag{5.20}$$

Thus, a PFE-SCNL whose parameters are related by (5.20) and whose cells are receiving the maximum input J_M, will reach a stable state at $\hat{x}_i = \Theta + \delta$.

Now let us rewrite equations (5.1) and (5.2) in the following form:

Stored Transmitter

$$\dot{u}_i = \alpha_u(z_i - u_i) - (\beta_u + K_u J_M \delta)(u_i - y_i) \tag{5.21}$$

Mobilized transmitter

$$\dot{y}_i = (\beta_y + F_i)(u_i - y_i) - J_M(\rho_y + K_y\delta)y_i - \gamma_y y_i \tag{5.22}$$

Comparison of (5.21) and (5.22) with (5.1) and (5.2) respectively, shows that the parameters a, c, d and e can be written in terms of the desired equilibrium state $\hat{x}_i = \Theta + \delta$ (with all synaptic inputs at J_M). Since we want the tonic level of transmitter mobilization (β_y) to be able to drive the synapse so that it satisfies (5.20), we can let $F_i = 0, \forall i$. Then

$$a = \alpha_u \tag{5.23}$$

$$c = \beta_u + K_u J_M \delta \tag{5.24}$$

$$d = \beta_y \tag{5.25}$$

$$e = J_M(\rho_y + K_y\delta) + \gamma_y \tag{5.26}$$

Substituting (5.23)-(5.26) into (5.20) gives

$$\frac{z_i\,\alpha_u\,\beta_y}{\alpha_u\,\beta_y + \alpha_u\,[J_M(\rho_y + K_y\delta) + \gamma_y] + [\beta_u + K_u J_M\delta]\,[J_M(\rho_y + K_y\delta) + \gamma_y]}$$

$$= \frac{\overline{G}D(\Theta+\delta)\left(\frac{\overline{B}}{\overline{A}}\right)\left(\frac{n-1}{n}\right)\delta + AD(\Theta+\delta)}{GJ_M[B-(\Theta+\delta)]\,[\rho_v + K_v\delta]} \tag{5.27}$$

Note that the above equality holds only for the case where every neuron receives the same input. Now rearrange (5.27) and express $\overline{G}$ in terms of all the other variables. First write

$$Q = \frac{\overline{G}M + N}{P} \tag{5.28}$$

where

$$Q = \hat{y} = \frac{z_i \alpha_u \beta_y}{\alpha_u \beta_y + \alpha_u [J_M(\rho_y + K_y\delta) + \gamma] + [\beta_u + K_u J_M \delta][J_M(\rho_y + K_y\delta) + \gamma_y]} \tag{5.29}$$

$$M = D(\Theta + \delta)\left(\frac{\overline{B}}{\overline{A}}\right)\left(\frac{n-1}{n}\right)\delta \tag{5.30}$$

$$N = AD(\Theta + \delta) \tag{5.31}$$

$$P = GJ_M[B - (\Theta + \delta)][\rho_v + K_v\delta] \tag{5.32}$$

Then we can express $\overline{G}$ (the inhibitory gain) in terms of all the other parameters as

$$\overline{G} = \frac{PQ - N}{M} = \frac{GJ_M[B - (\Theta + \delta)][\rho_v + K_v\delta]Q - AD(\Theta + \delta)}{D\delta(\Theta + \delta)\left(\frac{\overline{B}}{\overline{A}}\right)\left(\frac{n-1}{n}\right)} \tag{5.33}$$

5.3 Determination of Critical Values

Since our whole design procedure relies on our ability to find a robust value of δ that is not too large, we first need to determine the critical values $(\overline{G}_c, \overline{v}_{ic}, \delta_c)$ that are to be avoided since they may lead to oscillations. That is, we need to know the critical value of δ. Let δ_c be the critical value at which the layer will oscillate persistently. Below we describe a method for obtaining δ_c.

By realising that for a sufficiently large inhibitory gain (large $\overline{G}$ in equation (5.12)), the layer will persistently oscillate regardless whether the postsynaptic feedback is enabled or not, it should be clear that δ_c can be determined by ignoring the postsynaptic feedback. This simple realisation enables us to determine δ_c because the synaptic variables (u_i and y_i in equations (5.21) and (5.22) respectively) as well as the EPSP variable (v_i in equation (5.17)) will reach a stable state, while $x_i(t)$ in (5.12) and the inhibitory potential $\overline{v}_i(t)$ (in (5.13)) will oscillate ($x_i(t)$ will oscillate about $\Theta + \delta_c/2$, while $\overline{v}_i(t)$ will oscillate about its average that is related to δ_c).

Write (5.12) as

$$x_i(t) = \frac{BG\hat{v}_i}{A + G\hat{v}_i + \overline{G}\overline{v}_i(t)} \tag{5.34}$$

where $\hat{v}$ denotes the steady state EPSP. There exists a critical value of $\overline{G}_c\overline{v}_{i_c}$ that drives the layer into persistent oscillations. If this oscillation is a limit-cycle (between Θ and $\Theta + \delta_c$), then the inhibitory potential will oscillate about its average with an amplitude $\overline{A}\overline{v}_{i_c}$. Below we provide a method for determining the critical δ_c at which the variable $x_i(t)$ will oscillate between Θ and δ_c.

First determine the critical value of the product $\overline{G}_c\overline{v}_{i_c}$ from

$$\Theta = \frac{BG\hat{v}_i}{A + G\hat{v}_i + \overline{G}_c\overline{v}_{i_c}} \tag{5.35}$$

where $\overline{v}_{c_c}$ is the critical value of the inhibitory potential. Rearranging (5.35) gives

$$\overline{G}_c\overline{v}_{i_c} = \frac{G\hat{v}_i(B - \Theta) - A\Theta}{\Theta} \tag{5.36}$$

Now realise that the critical value of δ_c (i.e., the maximum value of $x_i(t)$) is given when the inhibitory potential is at its minimum value of $\overline{v}_{i_c}(1 - \overline{A})$. Then

$$\delta_c = \frac{BG\hat{v}_i}{A + G\hat{v}_i + \overline{G}_c\overline{v}_{i_c}(1 - \overline{A})} - \frac{BG\hat{v}_i}{A + G\hat{v}_i + \overline{G}_c\overline{v}_{i_c}} \tag{5.37}$$

which can be rearranged as

$$\delta_c = \Theta\left(\frac{BG\hat{v}_i}{G\hat{v}_i[B - \overline{A}(B - \Theta)] + \overline{A}A\Theta} - 1\right) \tag{5.38}$$

Now we can determine $\overline{v}_{i_c}$ (the upper peak of $\overline{v}_i(t)$) from

$$\frac{(\overline{v}_{i_c} - \overline{A}\overline{v}_{i_c}) + \overline{v}_{i_c}}{2} = \frac{\delta_c}{2}\left(\frac{\overline{B}}{\overline{A}}\right)\left(\frac{n-1}{n}\right) \tag{5.39}$$

i.e., at the critical values, the average of the inhibitory potential is related to the peak of the oscillation in $x_i(t)$. Therefore the critical $\overline{v}_{i_c}$ is given by

$$\overline{v}_{i_c} = \frac{\delta_c\left(\frac{\overline{B}}{\overline{A}}\right)\left(\frac{n-1}{n}\right)}{(2-\overline{A})} \tag{5.40}$$

The critical inhibitory gain ($\overline{G}_c$) can then be determined by substituting $\overline{v}_{i_c}$ into (5.36). This gives

$$\overline{G}_c = \frac{(2-\overline{A})\,[G\hat{v}(B-\Theta)-A\theta]}{\Theta\delta_c\left(\frac{\overline{B}}{\overline{A}}\right)\left(\frac{n-1}{n}\right)} \tag{5.41}$$

Note that this critical gain represents the gain of the lateral inhibition at which the layer will oscillate with a maximum amplitude (i.e., between Θ and δ_c, with an amplitude of $\delta_c/2$). However, there exists a lower critical gain $\overline{G}_c^L$ above which the layer will also converge to a stable oscillatory state but with a smaller amplitude, and below which it will converge to an absolute stable state. Therefore to ensure layer stability we need to be able to determine a suitable value of $\overline{G}$ which is smaller than $\overline{G}_c^L$. Before we specify how a suitable value of $\overline{G}$ may be determined, we first provide a procedure for the selection of the various parameters and demonstrate the oscillatory behaviour of the layer via several computer simulations. We then show that a layer will converge to an absolute stable state if $\overline{G}$ is chosen such that it maintains the cellular activity $x_i(t)$ at or above $\Theta+\frac{\delta_c}{2}$.

5.4 Intuition in the Parameter Selection

With such a larger number of parameters embedded in (5.33), it is obvious that most of them should be chosen on the basis of good intuition since not all combinations that satisfy (5.33) are going to be useful. Here we present an intuitive argument on how to choose most of these parameters.

Stored transmitter

The first thing to realise is that if the presynaptic facilitation is to be effective in the regulation of the bottom-up synapses for general cases where the attended bottom-up input spatial pattern is not of constant amplitude across the input layer, then by far the main contribution to the depletion of the stored transmitter should be due to the correlated firing of the input signal and the postsynaptic feedback signal. The implication of this is that if the facilitatory presynaptic 2-D spatial pattern is uniform in strength but the corresponding input pattern is not, then the presynaptic facilitation should increase the synaptic gains for these pathways by a larger amount

than for those pathways whose bottom-up inputs are much stronger. Therefore, we want the depletion of the stored transmitter to be mainly controlled by the term $K_u J_M \delta$ in equation (5.21). This means that we want $K_u \gg \beta_u$.

Now we need to choose the rate of transmitter storage (α_u). Since we want the dynamics of the synapse to be slower than the dynamics of EPSP, we need $\alpha_u < D$ (see equations (5.17) and (5.21)). Let us choose $\alpha_u = D/10$.

Mobilized transmitter

Our next intuition comes from the understanding of how the non-relevant inputs may be suppressed by lateral competition in the layer. If the attended bottom-up input pattern is presynaptically facilitated, then the amplitude of the facilitated cells must be able to increase in order for them to suppress the non-facilitated cells. For this suppression to be effective, the initial cellular activity must be low. Otherwise, the inhibition from term $(B - x_i)$ (see equation 4.21, Chapter 4) will prevent the facilitated cells from increasing their activity to a level that can suppress the non-facilitated cells to below the threshold. Hence, we want the tonic level of transmitter mobilization to be low but sufficient to enable an input to excite the layer above the threshold Θ. The implication of this is that $F_i \gg \beta_y$. Choose $\beta_y = 0.01$. This gives us plenty of room to provide a large gain for presynaptic facilitation of the synapse.

We are now in a position to choose the other parameters, leaving one aside ($\overline{G}$ - the gain of lateral inhibition, which we will determine from (5.41) after we calculate δ_c).

5.5 Example of a Design

In this section we provide an example of the application of the parameter design procedure. Computer simulation results of the designed layer and the Feedforward Excitation-Feedback Presynaptic Facilitation (FFE-FBPF) neural circuit will be presented in section 5.7.

Parameters that define the global property of the PFE-SCNL (the number of neurons, the upper saturation level, the postsynaptic threshold, etc.) can be set according to the desired characteristics. We can thus specify some of the parameters (A, B, Θ, J_M) according to a particular requirement. To demonstrate the design procedure, let us consider the following requirements:

Let the cellular activity be bounded in the range (0,1). This implies B = 1. We can also let $A = 1$ (the cellular dynamics is much faster than the synaptic dynamics). Let $\Theta = B/10 = 0.1$. Let the maximum bottom-up input be $J_M = B = 1.0$. Choose the layer to have 1024 neurons (32x32 array). Let the transmitter production level be $z_i = 1.0, \forall i$. Choose the EPSP to be slower than the cellular dynamics but faster than the synaptic dynamics. Let $D = 0.5$. Let the contribution to the EPSP from the correlated firing of the input signal and the postsynaptic feedback signal be ten times that of the input signal alone, i.e., let $K_v = 10\rho_v$. Choose $K_v = 10D = 5$, therefore $\rho_v = 0.5$. Let the EPSP gain (G) be 1000 . This is large enough to excite the layer above the threshold for small inputs (typically G is set according to the required dynamic range, larger G implies a wider dynamic range). Let the mobilized transmitter decay at a rate $\gamma_y = 0.5$. Let the combined action of the postsynaptic feedback and the input signal deplete the mobilized transmitter at ten times the rate of the input signal alone, i.e., $K_y = 10\rho_y$. Thus for a layer that is approximated by the first order Euler's iterative method with a time step $\Delta t = 1$, this gives:

$$A = B = 1.0; \quad G = 1000; \quad \overline{A} = \frac{A}{10} = 0.1; \quad \overline{B} = \frac{B}{10} = 0.1; \quad n = 1024; \quad J_m = 1.0;$$

$$\Theta = 0.1; \quad z_i = 1.0; \quad K_v = 5; \quad D = 0.5; \quad \rho_y = 0.05; \quad K_y = 0.5; \quad \gamma_y = 0.5; \quad \rho_v = 0.5;$$

$$K_u = 0.99; \quad F_i = 0.99; \quad \alpha_u = 0.05; \quad \beta_u = \beta_y = 0.01.$$

Steady state mobilized transmitter

Substitute the chosen values into (5.29). Note that for zero-postsynaptic feedback, this equation simplifies to

$$Q = \hat{y}_i = \frac{z_i \alpha_u \beta_y}{\alpha_u \beta_y + \alpha_u [J_M \rho_y + \gamma] + \beta_u [J_M \rho_y + \gamma]}$$

$$= 0.01492537 \tag{5.42}$$

Estimation of δ_c, $\overline{G}_c$, δ and $\overline{G}$

Plugging the calculated value of Q and the other selected parameters into (5.41) gives $\delta_c = 0.0098092613$. Plugging this value into (5.41) gives $\overline{G}_c = 25850.215$. To ensure layer stability we need $\overline{G} < \overline{G}_c$. In section 5.6 we specify how a suitable $\overline{G}$ may be determined to ensure convergence to a non-oscillatory steady state.

In the following two simulations we use the critical $\overline{G}_c$ for two cases: (i) without the postsynaptic feedback (Figure 5.1); and (ii) with the postsynaptic feedback (Figure 5.2).

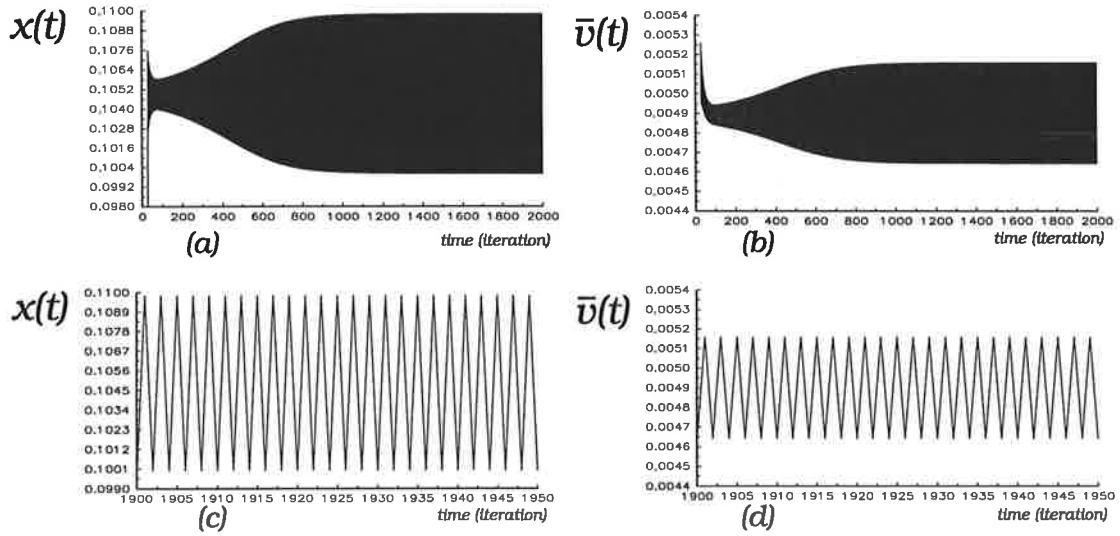


FIGURE 5.1. Simulation of the PFE-SCNL without the postsynaptic feedback, oscillating at the critical values ($\delta_c = 0.0098092613$, $\overline{G}_c = 25850.2$). Shown is the dynamics of one cell in the layer: (a) Oscillatory behaviour of x_i observed over 2000 time steps; (b) oscillatory behaviour of $\overline{v}_i$ observed over 2000 time steps; (c) Oscillatory behaviour of x_i observed over 50 time steps; (d) Oscillatory behaviour of $\overline{v}_i$ observed over 50 time steps. The three synaptic variables reached their steady state values after 100 iterations: $\hat{v}_i = 0.014925$; $\hat{u}_i = 0.83582$; $\hat{y}_i = 0.014295$.

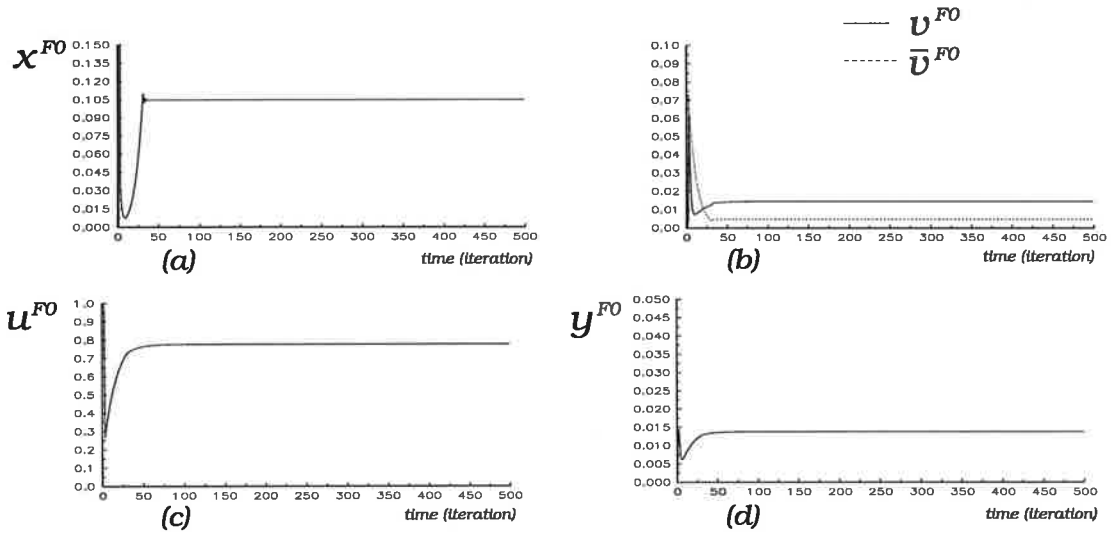


FIGURE 5.2. Simulation of the PFE-SCNL at the critical values and with the postsynaptic feedback. Steady state values are: $\hat{x}_i = 0.104745$; $\hat{v}_i = 0.014453$; $\overline{\hat{v}}_i = 0.004740$; $\hat{u}_i = 0.77970$; $\hat{y}_i = 0.013798$.

Figure 5.2 shows that postsynaptic feedback provides a stabilizing effect on the layer that would otherwise oscillate. Comparison of the synaptic steady state values ($\hat{v}_i$, $\hat{u}_i$ and $\hat{y}_i$) for the above two simulations shows that in the latter case all variables have a lower equilibrium point. The reason for this is that the combined action of the input signal and the postsynaptic feedback signal simultaneously depletes the levels of both the mobilized and the stored transmitter in the synapses. Note however that the postsynaptic feedback has stabilized the layer.

5.6 Absolute Stability Requirement

In this section we show that if $\overline{G}$ is chosen such that it maintains the cellular activity of an open-loop layer (i.e., the layer with no postsynaptic feedback) above $\Theta + \delta_c/2$, then the open loop layer will converge to a stable non-oscillatory state. By assuming that the synaptic variables have reached their steady state, we can analyse the behaviour of the layer by forcing it with a suitable initial input condition.

Let us consider the following initial conditions:

Synaptic variables at steady state: $v(0) = \hat{v}$, $y(0) = \hat{y}$; $u(0) = \hat{u}$;

$x(0) = 0$;

$\overline{v}(0) = \delta_0$

We can now force the layer towards the specified equilibrium point δ_0 by driving it with a inhibitory gain $\overline{G}$ that can be determined from (5.43). Note that this equation is derived from (5.34) by letting the equilibrium point be $x(\infty) = \hat{x} = \delta_0$. If δ_0 is below some critical value δ_c^L, then the layer will converge to a stable oscillatory dynamics, otherwise it will converge to an absolute stable state (non-oscillatory). We aim to show that if $\delta_0 \geq \delta_c/2$, where δ_c is given by (5.38), then the layer will converge to a non-oscillatory state for zero contrast in the input spatial pattern whose average is J_M/n (where n is the number of neurons in the layer).

$$\overline{G} = \frac{BG\hat{v} - (A + G\hat{v})(\Theta + \delta_0)}{\delta_0(\Theta + \delta_0)\left(\frac{\overline{B}}{\overline{A}}\right)\left(\frac{n-1}{n}\right)} \tag{5.43}$$

Since the initial condition for the cellular variable $x(0) = 0$, then $x(1)$ will jump above Θ after the first iteration. The attained value after the first iteration will be

$$x(1) = \frac{BG\hat{v}}{A + G\hat{v} + \overline{G}\delta_0} \tag{5.44}$$

If $(\overline{B}/\overline{A})(n - 1/n) \leq 1$ and for $\delta t = 1$, then for the layer to converge to an oscillatory steady state, the upper peak of the oscillation will converge above δ_0, while the lower peak will converge below δ_0. Now the necessary condition for convergence is that the upper peak of the oscillation ($x^u(t)$) decays in magnitude and approaches δ_0 from above, while the lower peak should approach δ_0 from below. To illustrate this condition, Figure 5.3 shows a simulation for two different initial values of δ_0 (all other parameters are as in section 5.5).

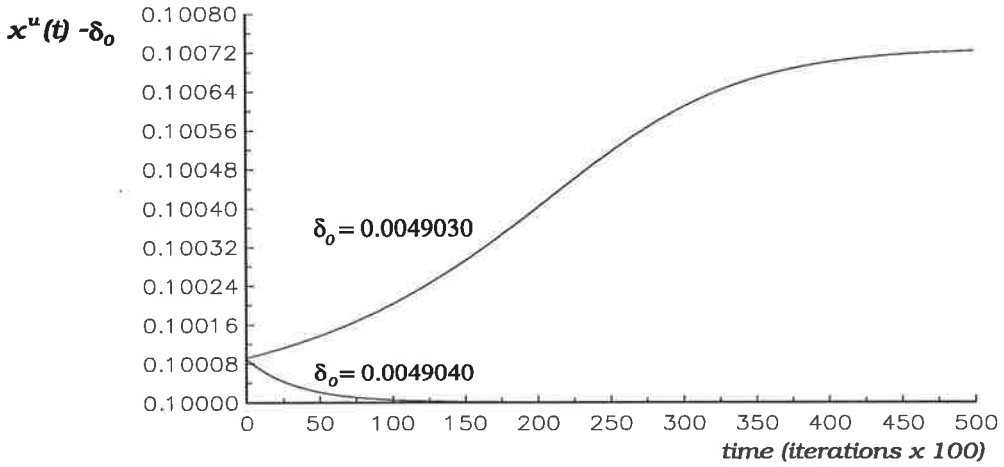


FIGURE 5.3. Departure of the cellular variable $x^u(t)$ from the forced steady state value for two input conditions. The vertical axis represents the difference between the peak value of $x(t)$ and the forced equilibrium value of δ_0. The graphs shows that the layer cannot be forced into a stable non-oscillatory state when δ_0 is below some critical value. That is, there exists some critical value below which the layer will enter into persistent oscillations.

That is, for $x(t)$ to converge to a stable non-oscillatory state, then two iterations after it is given an input, the value of $x(3)$ must be below $x(1)$. The following three equations represent the evolution of the variables $x(t)$ and $\overline{v}(t)$ during the first few iterations.

$$\overline{v}(1) = (1 - \overline{A})\delta_0 + \overline{B}\left(\frac{n-1}{n}\right)[x(1) - \Theta]^+$$

$$= (1 - \overline{A})\delta_0 + \overline{B}\left(\frac{n-1}{n}\right)\left[\frac{BG\hat{v}}{A + G\hat{v} + \overline{G}\delta_0} - \Theta\right]^+ \tag{5.45}$$

$$x(2)=\frac{BG\hat{v}}{A+G\hat{v}+\overline{G}\overline{v}(1)}$$

$$=\frac{BG\hat{v}}{A+G\hat{v}+\overline{G}\left[(1-\overline{A})\delta_0+\overline{B}\left(\frac{n-1}{n}\right)\left[\frac{BG\hat{v}}{A+G\hat{v}+\overline{G}\delta_0}-\Theta\right]^+\right]} \tag{5.46}$$

$$\overline{v}(2)=(1-\overline{A})\overline{v}(1)+\overline{B}\left(\frac{n-1}{n}\right)[x(2)-\Theta]^+$$

$$=(1-\overline{A})\left[(1-\overline{A})\delta_0+\overline{B}\left(\frac{n-1}{n}\right)\left[\frac{BG\hat{v}}{A+G\hat{v}+\overline{G}\delta_0}-\Theta\right]^+\right]$$

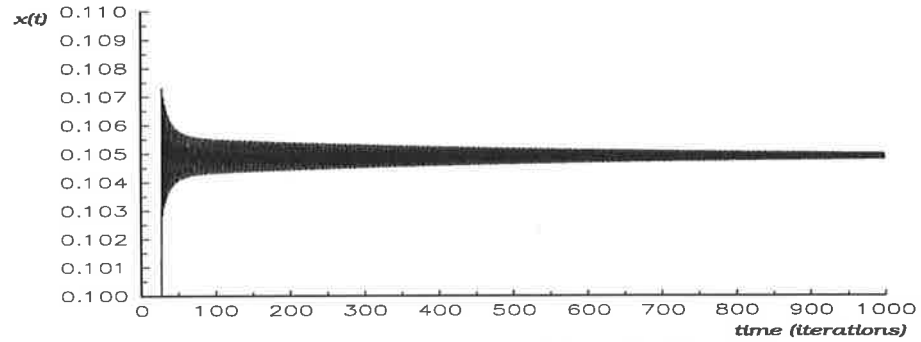
$$+\overline{B}\left(\frac{n-1}{n}\right)\left[\frac{BG\hat{v}}{A+G\hat{v}+\overline{G}\left[(1-\overline{A})\delta_0+\overline{B}\left(\frac{n-1}{n}\right)\left[\frac{BG\hat{v}}{A+G\hat{v}+\overline{G}\delta_0}-\Theta\right]^+\right]}-\Theta\right]^+ \tag{5.47}$$

Alternatively, the inhibitory potential at t = 2 must be below its initial value (i.e., the peak of the inhibitory potential decays). That is, if $\overline{v}(2)<\overline{v}(0)$ then the oscillations will decay. The reason for this is that the final steady state $\hat{v}$ is always less than $(\hat{x}-\Theta)$. The rate of the oscillation decay will depend on how close $\delta_c/2$ is to the lower critical value δ_c^L (the closer the two are, the slower the rate of decay).

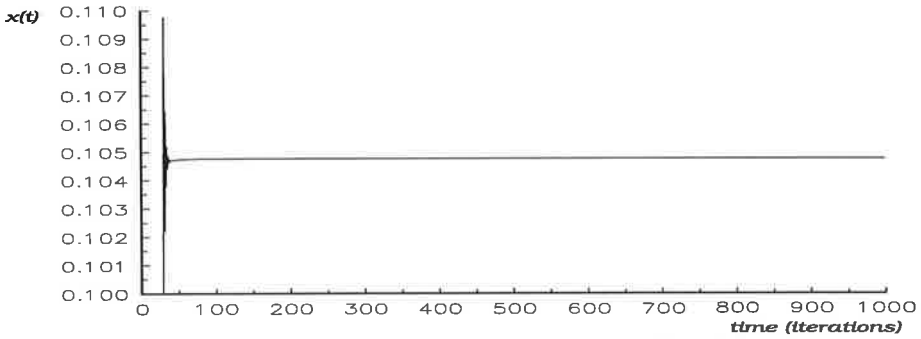
The critical δ_c^L can be found by solving $\overline{v}(2)-\delta_0=0$. Although we have not found a closed form analytic solution to δ_c^L (computer simulations show that for the chosen set of parameters δ_c^L~0.0049032, $\overline{G}^L$~25794.5) we have been able to show that if $\overline{B}/\overline{A}=1$ and $\delta t=1$, then $\overline{v}(2)-\delta_c/2<0$ (i.e., $\delta_c/2$ leads to convergence). Thus $\delta_0=\delta_c/2$ guarantees convergence. Since $\delta_c/2$ can be determined from (5.38) and $\overline{G}$ from (5.43), we can then determine the magnitude of the input that can drive the layer to stability. Since $\delta_c/2>\delta_c^L$, the predicted input magnitudes will be somewhat pessimistic (i.e., they will be higher than the minimum required to drive the layer to stability).

Thus for stability, we want $\hat{x}\geq\Theta+\delta_c/2$. Therefore

$$\frac{BG\hat{v}}{A+G\hat{v}+\overline{G}\overline{v}}\geq\Theta+\frac{\delta_c}{2} \tag{5.48}$$



(a) Open loop (no postsynaptic feedback)



(b) Closed loop (postsynaptic feedback)

FIGURE 5.4. Decay in oscillations at $\delta_0 = \delta_c/2$ in the open-loop and the closed-loop layer. Initial conditions for the simulation are: $x(0) = 0$; $\bar{v}(0) = 0$; $v(0) = 0$; $u(0) = 1$; $y(0) = 0$. Note that the postsynaptic feedback has increased the rate of oscillation decay.

Substituting for $\delta_c/2$ from (5.38) gives

$$\frac{BG\hat{v}}{A + G\hat{v} + \overline{G}\overline{v}} \geq \Theta + \frac{\Theta}{2}\left(\frac{BG\hat{v}_i}{G\hat{v}_i[B - \overline{A}(B - \Theta)] + \overline{A}A\Theta} - 1 \right) \tag{5.49}$$

Rearranging and substituting for $\bar{v}$ and $\overline{G}$ ($\bar{v} = (\overline{B}/\overline{A})[(n-1)/n]\delta_c/2$, $\overline{G}$ is given by (5.43)) gives

$$\hat{v} \geq \frac{-A_2 \pm \sqrt{(A_2^2 - 4A_1A_3)}}{2A_1} \tag{5.50}$$

where

$$A_1 = G^2([B - \overline{A}(B - \Theta)](B - \Theta) - \Theta B) \tag{5.51}$$

$$A_2 = 2BGS - 2\Theta AGR - 2G\Theta S - 2\Theta\overline{G}\overline{v}GR - A\Theta BG$$
$$- \Theta BG\overline{G}\overline{v} + \Theta GRA + \Theta GR\overline{G}\overline{v} + \Theta SG \tag{5.52}$$

$$A_3 = -2\Theta A S - 2\Theta \overline{G}\overline{v} S + \Theta S A + \Theta S \overline{G}\overline{v} \tag{5.53}$$

where

$$S = \overline{A}\Theta \tag{5.54}$$

$$R = [B - \overline{A}(B - \Theta)] \tag{5.55}$$

Now from equations (5.18) and (5.50) we can determine the minimum average intensity that an input spatial pattern needs to have to drive the layer to a stable state. Figure 5.5 shows the simulation result for increased levels of facilitation.

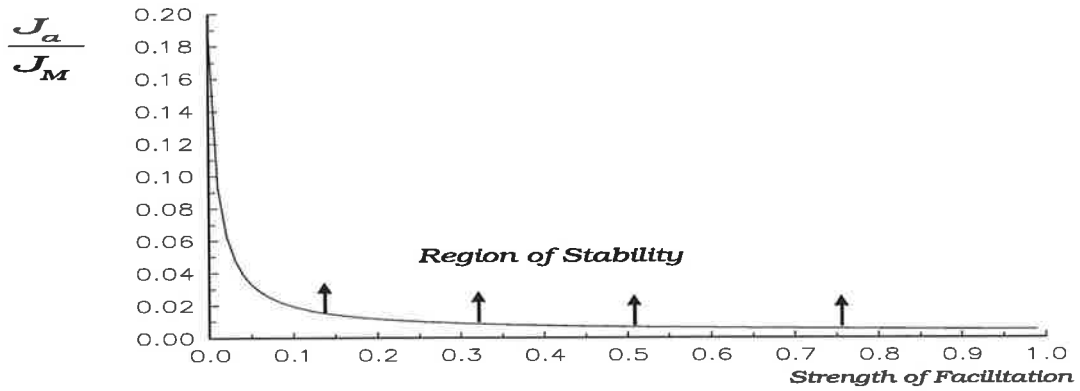


FIGURE 5.5. Stability of the facilitated shunting competitive neural layer as a function of the average input strength.

The above figure shows that an increased level of facilitation will enable the layer to stably process input patterns whose average strength is low. Note that the worst case scenario is a zero contrast input pattern since all the cells will be engaged in competition.

5.7 Simulation of Designed Layers

Below we provide simulations of the designed PFE-SCNL at the lower lateral gain $\overline{G} = 5000$ for three cases: (i) feedforward-only computations with zero postsynaptic feedback; (ii) feedforward-feedback case; and (iii) a preferred implementation where the fast cellular variable is iterated several times for each iteration of the slow synaptic variables.

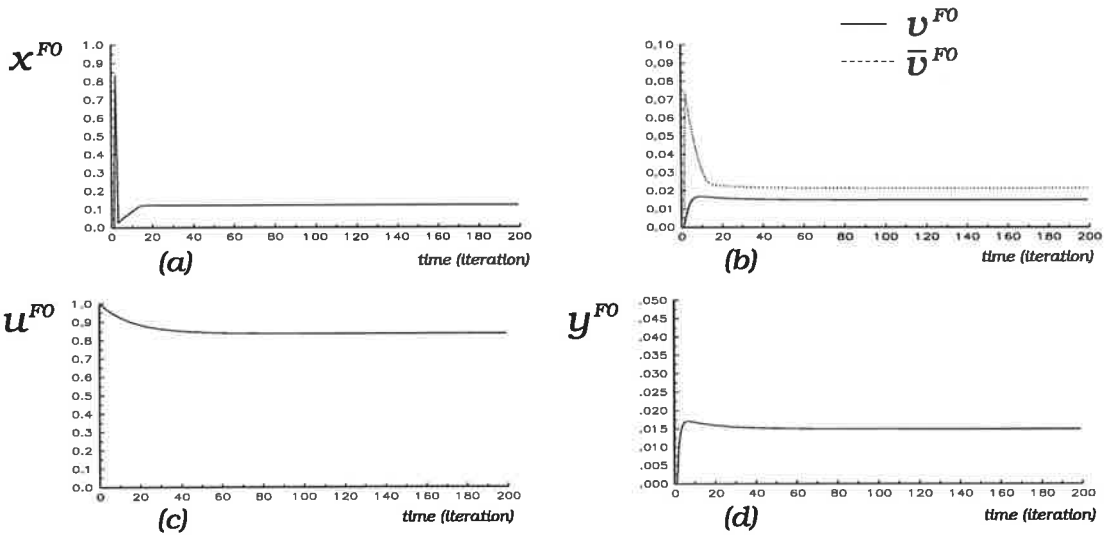


FIGURE 5.6. Simulation of the designed PFE-SCNL (zero postsynaptic feedback, $\overline{G} = 5000$): (a) Cellular dynamics (x_i); (b) Synaptic EPSP (v_i) and the lateral feedback inhibitory postsynaptic potential ($\overline{v}_i$); (c) Stored transmitter (u_i); and (d) Mobilized transmitter (y_i). The steady state values (after 200 iterations) are: $\hat{x}_i = 0.121420$; $\hat{v}_i = 0.014925$; $\overline{\hat{v}}_i = 0.02140$; $\hat{u}_i = 0.83582$; $\hat{y}_i = 0.014295$.

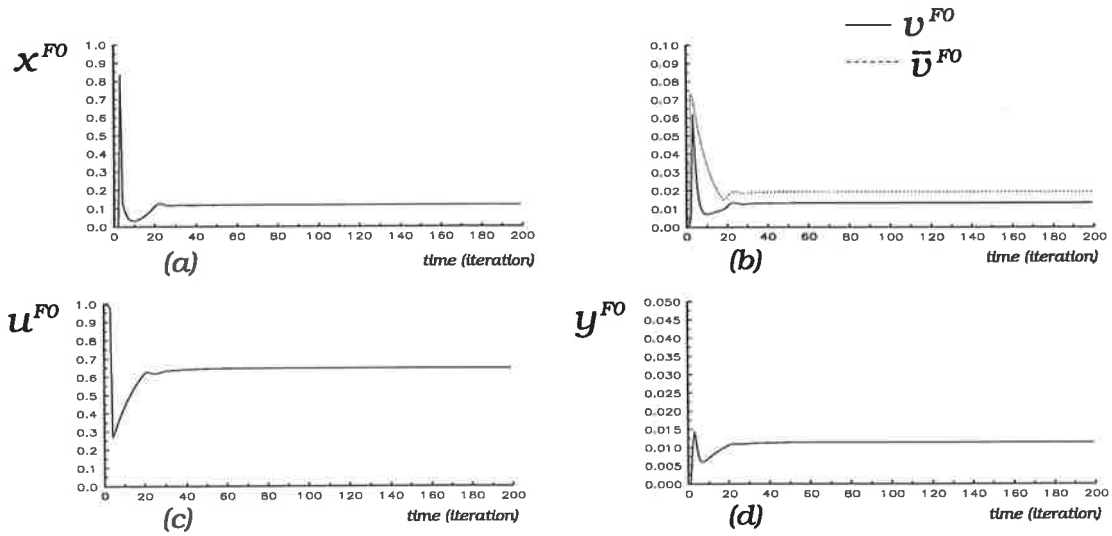


FIGURE 5.7. Simulation of the designed PFE-SCNL (with postsynaptic feedback, $\overline{G} = 5000$): (a) Cellular dynamics (x_i); (b) Synaptic EPSP (v_i) and the lateral feedback inhibitory postsynaptic potential ($\overline{v}_i$); (c) Stored transmitter (u_i); and (d) Mobilized transmitter (y_i). The steady state values (after 200 iterations) are: $\hat{x}_i = 0.119448$; $\hat{v}_i = 0.0133134$; $\overline{\hat{v}}_i = 0.019429$; $\hat{u}_i = 0.6350$; $\hat{y}_i = 0.011146$.

The simulation results in Figure 5.7 demonstrates that the steady state variables approach their calculated values. Note that the sharp spike shown in graph 5.7(a) is

due to an initial fast rise in x_i, versus a much slower rise in $\overline{v}_i$. This undesirable effect can be removed by iterating the variables x_i and $\overline{v}_i$ several times for each iteration of the other three variables (typically 5-10 times; see Figure 5.8).

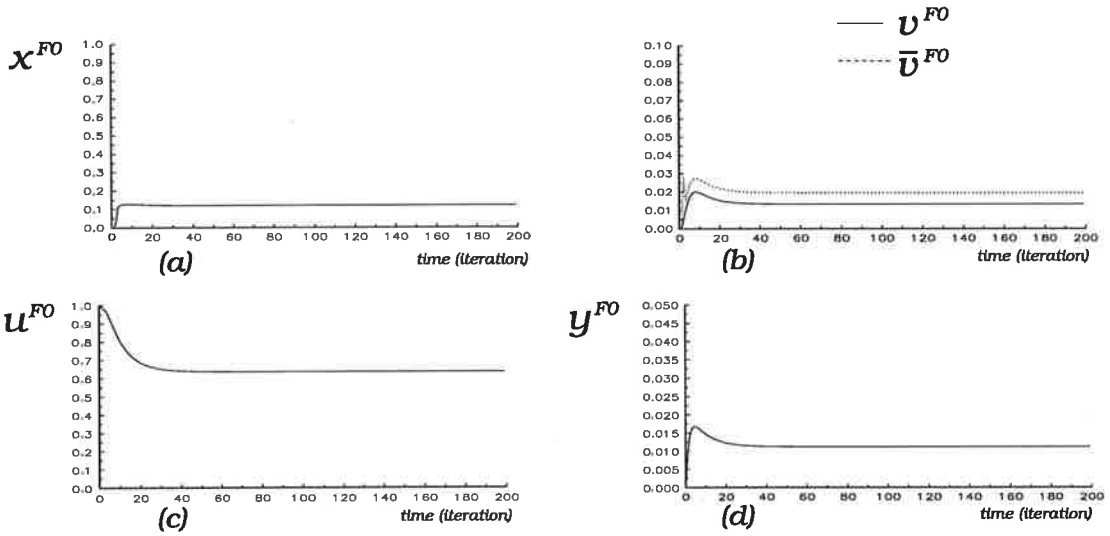


FIGURE 5.8. Simulation of the designed PFE-SCNL (with postsynaptic feedback and 10 iterations of the variables x_{ij} and $\overline{v}_{ij}$; $\overline{G} = 5000$): (a) Cellular dynamics (x_i); (b) Synaptic EPSP (v_i) and the lateral feedback inhibitory postsynaptic potential ($\overline{v}_i$); (c) Stored transmitter (u_i); and (d) Mobilized transmitter (y_i). The steady state values (after 200 iterations) are: $\hat{x}_i = 0.119448$; $\hat{v}_i = 0.0133134$; $\overline{v}_i = 0.019429$; $\hat{u}_i = 0.63500$; $\hat{y}_i = 0.011146$.

Note that the steady state values are equal to those of Figure 5.7, but the spike that appears in Figures 5.6 and 5.7 has been removed by iterating the variables x_i and $\overline{v}_i$ (equations (5.12) and (5.13) respectively) 10 times for each iteration of equations (5.17), (5.21) and (5.22).

5.7.1 Simulation of Selective Information Transfer

PM-SCNLs may be presynaptically facilitated or inhibited and this modulation may be pattern-specific or spatially specific. Each of these modulatory mechanisms affect the neural signal transmission in a different way and give rise to different end results. We will presently limit our discussion to presynaptic facilitation, i.e., we will consider the two facilitatory modes of regulating the signal flow into a PFE-SCNL.

Below we provide computer simulation results of the designed PFE-SCNL for each mode of synaptic facilitation (spatially-specific and pattern-specific).

Simulation of Spatially-specific Facilitation

Figure 5.9 shows the bottom-up 2-D spatial input and the resultant steady state of the layer for the case of zero facilitation. In Table 5.1 we show a 5x5 matrix, I(i,j), that depicts the amplitudes of the bottom-up inputs from a 5x5 region that is centred at the leftmost vertical edge of one of the ships (the ship shown in the middle of Fig 5.9(a)). In order to expose the synaptic and cellular dynamics, we will present simulation data for three input pathways whose bottom-up input amplitudes are shown in bold. The entry in position (4,0) corresponds to a high amplitude bottom-up input whose pathway is not on the ship's boundary. That is $J(4,1) = 1.0$. The entry in position (3,3) corresponds to a pathway that is on the leftmost edge of the ship's boundary and whose input magnitude is also high, $J(3,3) = 1.0$. The third bold entry, at position (4,4) corresponds to an input pathway that is on the ship's boundary (located on the horizontal edge at the far left corner) but whose input magnitude is relatively low, i.e., $J(4,3) = 0.1$.

In the simulations shown below, all parameters are as per design, with $\overline{G} = 5000$.

(a) J_{ij}

(b) Steady state cellular activation level ($\hat{x}_{ij}$) for the case of zero facilitation ($F_{ij} = 0 \quad \forall(i,j)$.)

FIGURE 5.9. Bottom-up input and the resultant steady state cellular activation level for the case of zero facilitation.

The simulation runs for 100 iterations of the layer (each differential equation is iterated using the Euler's first order approximation method with $\Delta t = 1$) and is split into two parts: iterations in the interval 0 - 49 are without the facilitatory gain control; iterations in the interval 50 - 100 include the facilitatory gain control.

TABLE 5.1

0.00	*0.00*	*1.00*	*0.00*	*0.00*
0.00	*0.00*	*1.00*	*0.00*	*0.00*
0.00	*0.00*	**1.00**	*0.00*	*0.00*
1.00	*0.00*	*0.00*	**0.10**	*0.40*
0.00	*0.00*	*0.00*	*0.00*	*0.00*

The table shows the amplitude of the bottom-up inputs, centred at the leftmost vertical edge of the ship shape (refer to the text for further explanation).

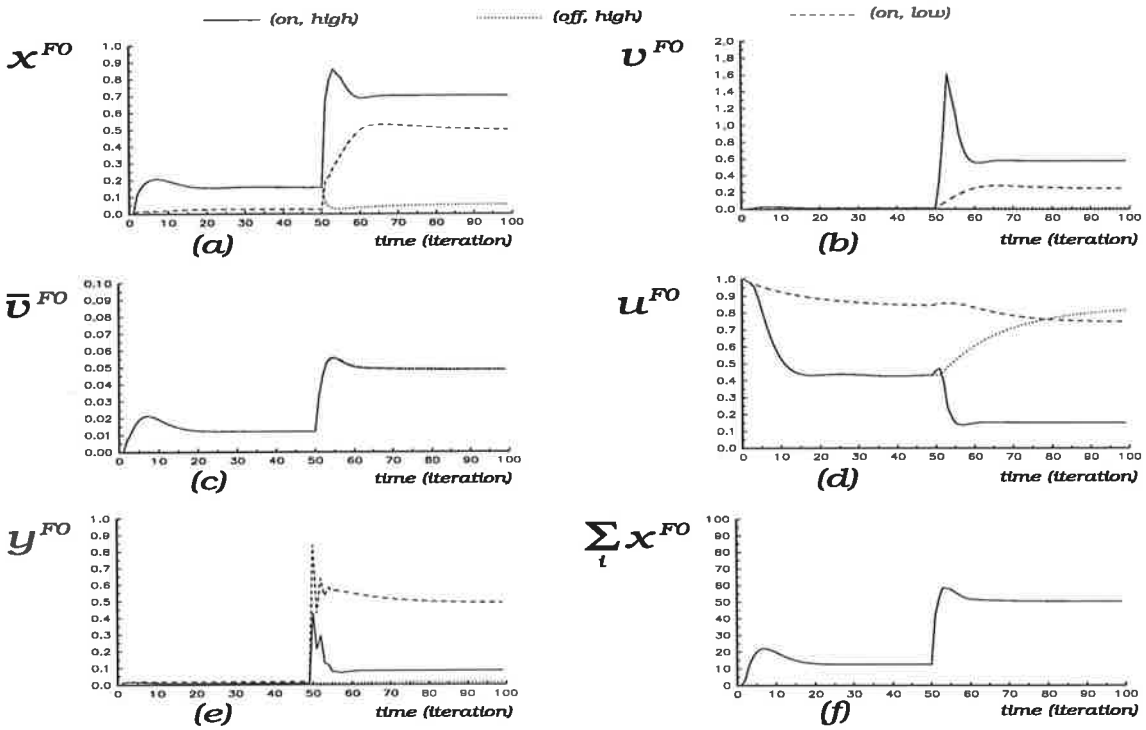


FIGURE 5.10. Synaptic and cellular dynamics for three input pathways, before, during and after the application of the spatially-specific facilitatory gain control.

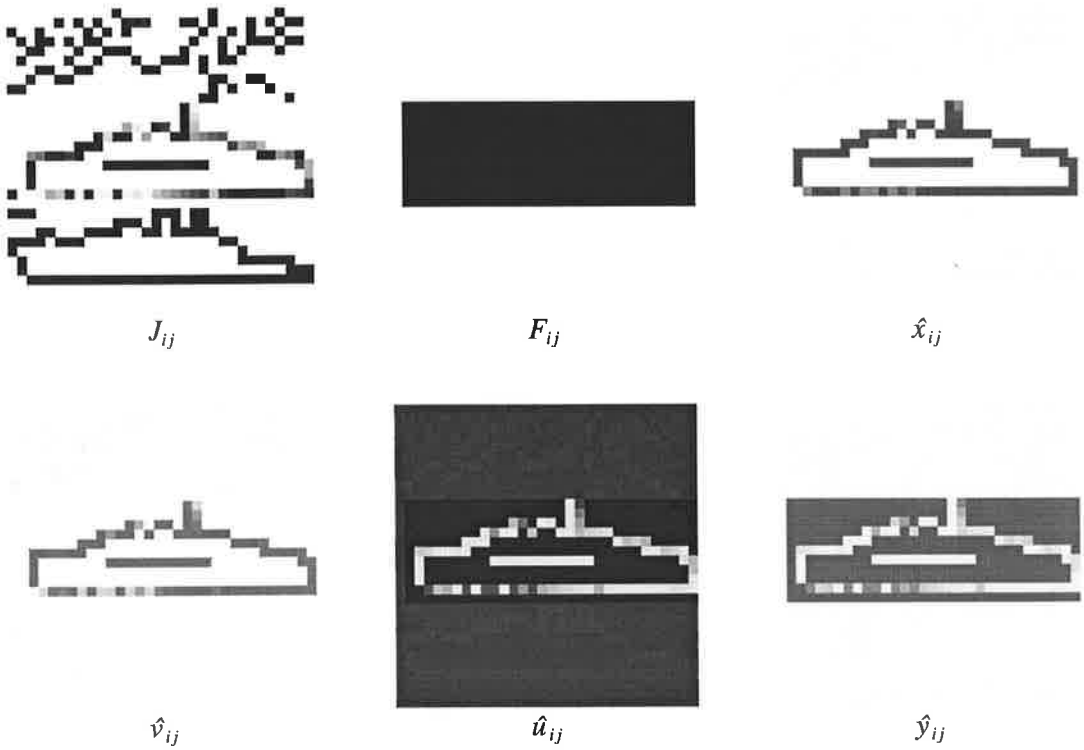


FIGURE 5.11. Steady state after the introduction of the spatially-specific facilitatory gain control signals.

Figure 5.10 shows the synaptic and cellular dynamics for the three input pathways: (i) on, high - for a cell whose location is on the reference shape and whose bottom-up amplitude is high; (ii) on, low - for a cell whose location is also on the reference shape but whose bottom-up amplitude is low; and (iii) off, high - for a cell whose location is not on the reference shape but whose bottom-up amplitude is high.

As can be seen in Figures 5.10 and 5.11, the bottom-up signals whose synapses are facilitated give rise to relatively high activation levels of their target cells, while the unfacilitated cells have their activity suppressed by lateral competition.

Simulation of Pattern-specific Facilitation

This simulation repeats the one above with the exception that we are now using a specific 2-D spatial pattern to provide the synaptic facilitation. Hence the final steady is expected to be different but only in those locations where the pattern-specific and the spatially specific facilitatory gain control signal differ. It is thus expected that the horizontal line that appears within the boundary of the input target ship will be eliminated from the layer once the shown pattern-specific facilitation is introduced into the layer.

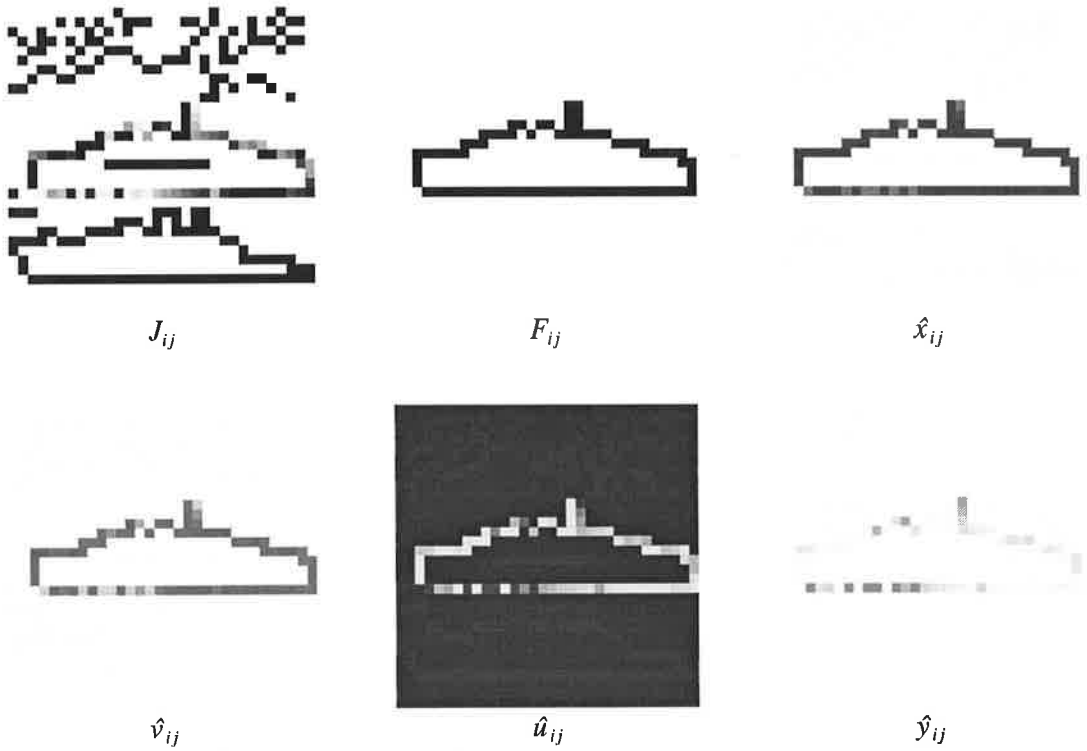
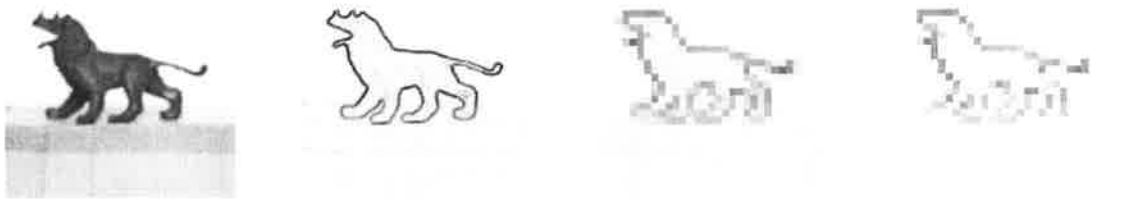


FIGURE 5.12. Steady state after the introduction of the pattern-specific facilitatory gain control signals.

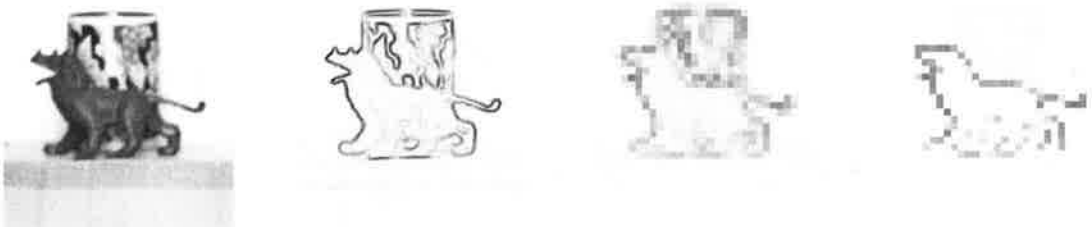
In the above simulations we have considered simple input spatial patterns that were not derived from real-world images. To demonstrate the robustness of pattern selective information transfer on real-world image data, below we show the simulation results of presynaptic facilitation on real-world visual image data.

Figure 5.13 shows an object (a toy dinosaur) whose edges are used as the input to layer. The grey level images (8 bits, 256x256 pixels, captured by a Vidicon camera and then digitized) were preprocessed by a 3x3 Sobel edge operator to obtain object edges. The resultant edge processed images were then scaled (by a simple averaging procedure) to the size of the layer (32x32 cells). This reduces the resolution of the target and the test images by a factor of 8. The edge processed image of the target object is initially fed to the layer. The resultant steady state (shown in Figure 5.13(d)) was then used as the facilitatory spatial pattern on cluttered images that contained the same object. Note that as a result of the competition in the layer, some of the weaker edges in the edge images of the target object did not survive and were absent in the memory, as shown in Figure 5.13(d).



(a) Object 1 (256x256) (b) Objects 2 edges (c) Input (32x32) (d) Steady state of the layer

FIGURE 5.13. Target object whose resultant edges are used as the 2-D facilitatory spatial pattern.



Test image 1 Image 1 edges Layer input Layer output

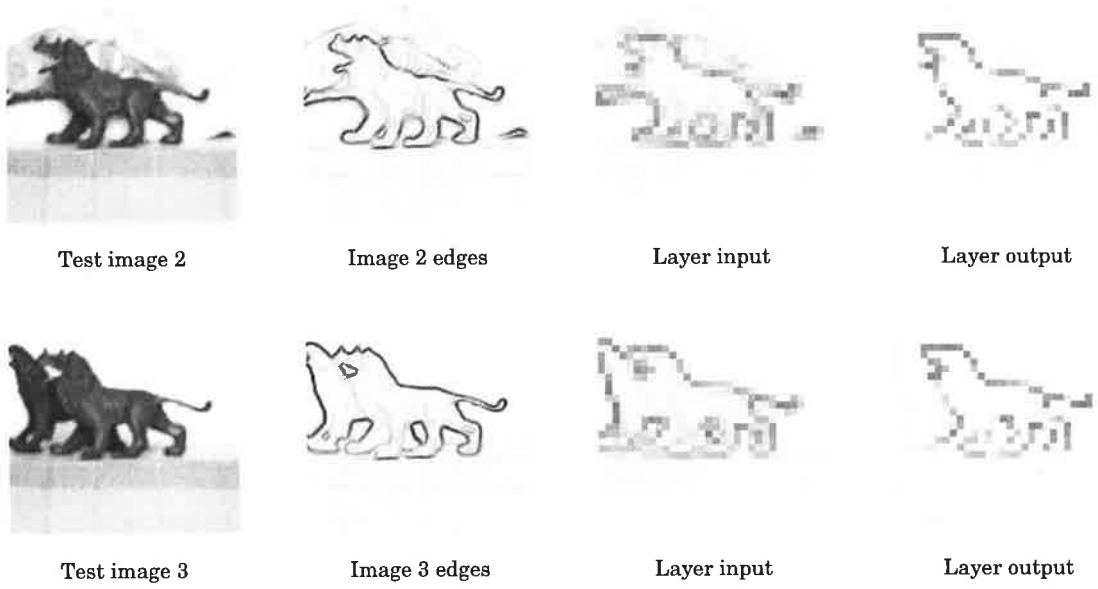


FIGURE 5.14. Example of selective information transfer on cluttered visual images.

In the next section we provide a computer simulation of the designed Feedforward Excitation-Feedback Presynaptic Facilitation neural circuit and expose its tuning properties on inputs whose intensity may differ along a given object's boundary.

5.7.2 Simulation of the FFE-FBPF Neural Circuit

To demonstrate the nature of tuning in the FFE-FBPF neural circuit of Chapter 4, we present simulations results on the two spatial patterns shown in Figures 5.17 and 5.18. The shape shown in Figure 5.17 is used a top-down reference that is fed into F1. This shape is also embedded in the cluttered bottom-up input of Figure 5.17, but some of its elements are very much weaker in amplitude than those of the non-relevant shape and clutter. The aim is to use the spatial pattern across F1 (the reference shape) to extract the same pattern from the input such that the match between F1 and F0 exceeds the required threshold of 0.98 (which is currently measured by the cosine of the angle between the two multidimensional vectors X_{ij}^{F0} and X_{ij}^{F1}). In the simulation, we first present the input of Figure 5.18 for the first 50 iterations of the circuit equations and then introduce the reference shape into F1. We will follow the dynamics of all variables in the circuit for three cellular positions in both PFE-SCNLs, i.e., Fields F0 and F1 (refer to TABLE 5.1). Both Fields have the same parameters with the exception that the bottom-up transmitter mobilization rate from Field F0 to Field F1 is five times that of the input synapses to F0. Synaptic pathways between the two Fields may also in general be modulated (as in the self-regulated attentional neural circuit of section 6.3.1 in Chapter 6).

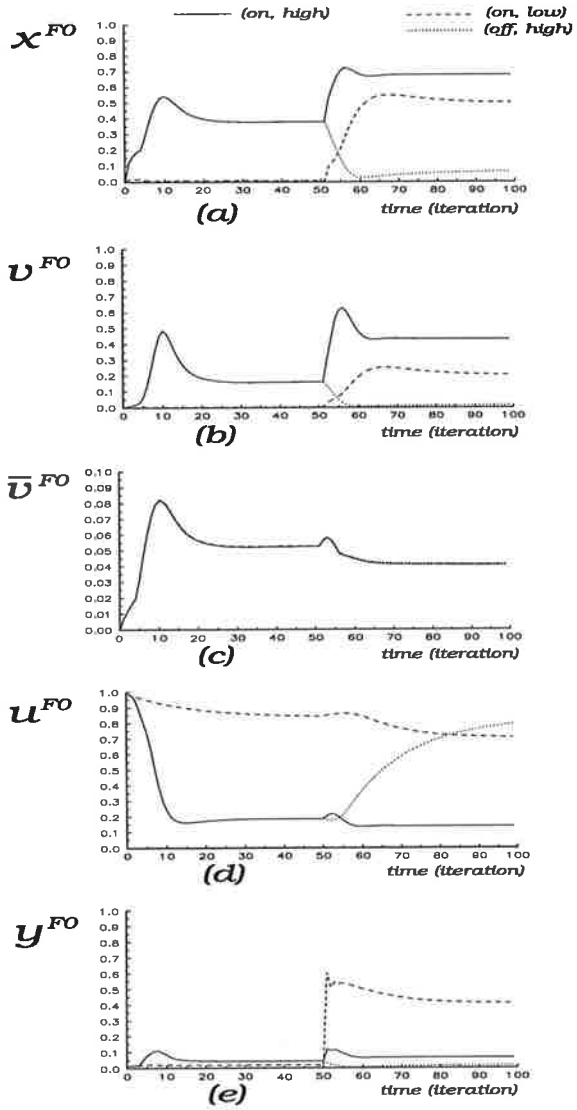


FIGURE 5.15. Dynamics of the five variables in Field F0, at three cellular positions (before, during and after the introduction of the top-down reference into F1).

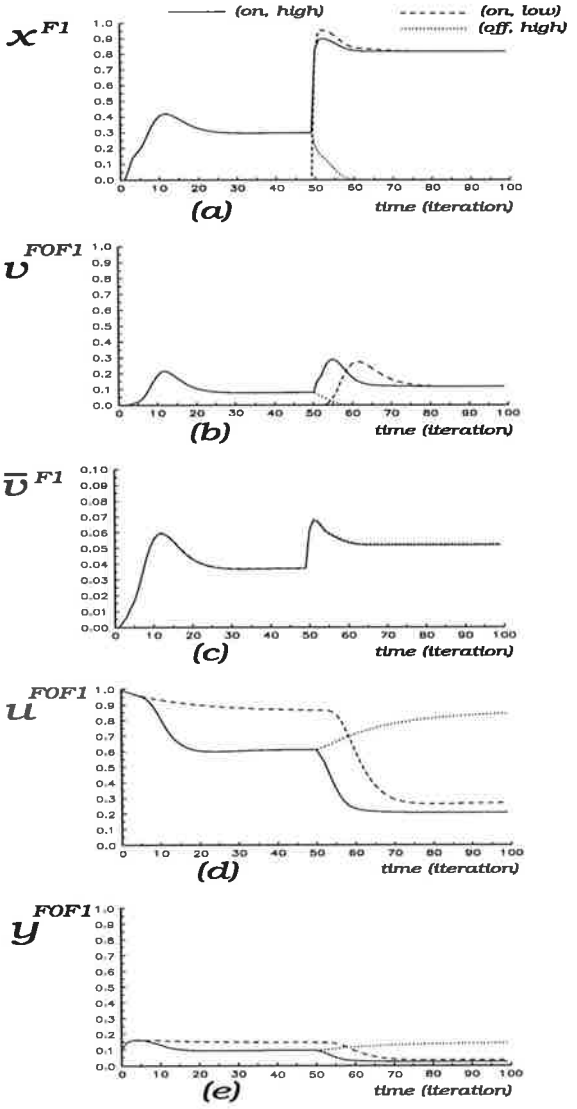
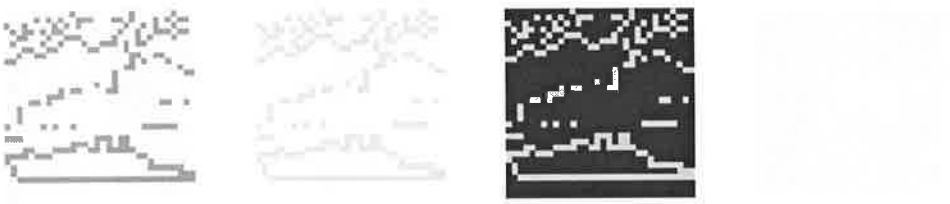


FIGURE 5.16. Dynamics of the five variables in Field F1 (at the same cellular locations as for F0).

FIGURE 5.17. 2-D top-down reference input into F1 used to regulated the circuit dynamics (shown in reverse contrast: black = 1; white = 0).

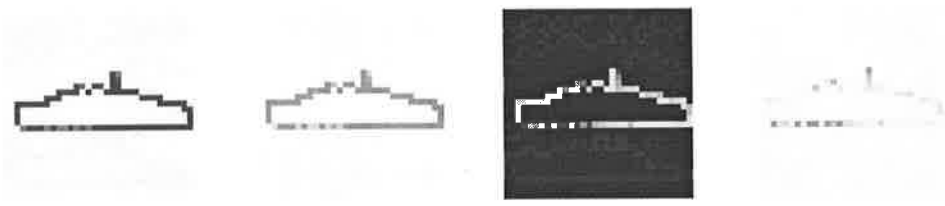
FIGURE 5.18. 2-D cluttered bottom-up input used to demonstrate the tuning characteristics of the FFE-FBPF neural circuit.

Figures 5.15 and 5.16 show the dynamics of various variables in the circuit, for each of the Fields, at three cellular positions: (i) on, high - for a cell whose location is on the reference shape and whose bottom-up amplitude is high; (ii) on, low - for a cell whose location is also on the reference shape but whose bottom-up amplitude is low; and (iii) off, high - for a cell whose location is not on the reference shape but whose bottom-up amplitude is high. The actual values (centred at the leftmost vertical edge of the ship shape) are shown bold in Table 5.1.



$\hat{x}_{ij}^{F0}$ $\hat{v}_{ij}^{F0}$ $\hat{u}_{ij}^{F0}$ $\hat{y}_{ij}^{F0}$

FIGURE 5.19. Steady state variables of Field F0 before the introduction of the top-down reference input into F1.



$\hat{x}_{ij}^{F0}$ $\hat{v}_{ij}^{F0}$ $\hat{u}_{ij}^{F0}$ $\hat{y}_{ij}^{F0}$

FIGURE 5.20. Steady state variables of Field F0 after the introduction of the top-down reference input into F1 (the final steady state match between F0 and F1 is 0.9950).

The results shown in Figure 5.20 show why the circuit has an interesting tuning property. If we refer to the two columns on the right, we can see that, in the steady state condition, the synaptic transmission gain (mobilized transmitter, variable $\hat{y}_{ij}^{F0}$) is higher in magnitude in those locations where the facilitated input signal is lower in magnitude, while the stored transmitter shows the opposite effect. The uniform top-down pattern specific facilitatory feedback signal and the pre-postsynaptic interactions have adjusted the two synaptic variables such that the synaptic gain is non-uniform across the input. This continual readjustment of the synaptic transmission gain (without long term memory modifications) provides the FFE-FBPF neural circuit (and the SAART neural network) the type of flexibility that does not appear to be available in the current artificial neural networks. This flexibility is needed to overcome the problem of having a fixed number of receptive fields that in general may not be able to handle the vast variety of cluttered backgrounds in which the same object can appear.

Figure 5.21 shows the global dynamics of the circuit during the simulation run. As can be seen in graph (a), the match between the spatial patterns across the two interacting neural Fields has fallen at the instant that the top-down input is introduced, but is soon elevated to a very high level (above 0.98, as measured by the cosine of the angle between the two multidimensional vectors that represent the two spatial patterns). Graph (b) shows that the circuit converges to a stable state (as measured by the time rate of change of the match).

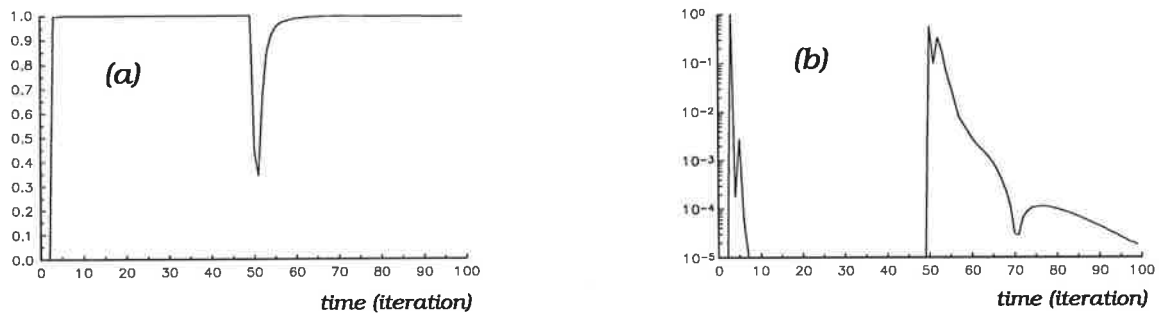


FIGURE 5.21. Dynamics of the FFE-FBPF neural circuit: (a) degree of match between F0 and F1 (threshold = 0.98); (b) rate of change of the match (steady state threshold = 0.0001).

5.8 Conclusions

We have presented a mathematical procedure for the design of modulated competitive neural layers which is based on finding the steady state solution to the synaptic variables when there is no postsynaptic feedback. The method has proven to

be quite effective in determining the critical values at which the layer will oscillate. Computer simulations of the designed layer have verified that the layer is stable under various input conditions.

Chapter 6

Self-regulation, Spatial Attention and Memory Guided Search

"Segmentation is a difficult task, especially when there are many objects in a scene partially occluding one another. The problem is essentially that global information is needed to make decisions at the local level concerning what goes with what. At lower levels of processing such as V1, however, the receptive fields are relatively small and it is not possible locally to decide which pieces of the image belong together. If lower levels can use information that is available at higher levels, such as representation of whole objects, then feedback connections could be used to help tune lower levels of processing".

Churchland, P.S., Ramachandran, V.S., and Sejnowski, T.J. (1994)

6.1 Introduction and Overview

In this chapter we extend the neuro-engineering design principles of Chapter 4 and propose a design logic for two dimensional neural circuits that are based on several interacting 2-D Presynaptically Modulated Shunting Competitive Neural Layers. Hence the mathematics that was introduced in Chapters 4 and 5 is also applicable here. However, in this Chapter we will propose several new ways of using PM-SCNLs in more complex neural circuits. As a guide to the desired neurodynamics and circuit properties, we will partially address the various aspects of cognitive visual functions (whose cognitive and electrophysiological data was discussed in Chapter 2) with a particular emphasis on the dynamics of short term memory, attentional modulation and attentional selection. The generated neural circuits, while increasing both in complexity and capability, are based on very simple psychophysically and neurophysiologically supported concepts of biological visual information processing.

The chapter addresses the following issues in mechanisms and circuits of primate vision and visual memory:

(i) neural circuit for the detection of resonant (matching) states;

(ii) reverberatory memory and self-regulated attentional modulation during learning;

(iii) feedforward (bottom-up) neural network architecture for pre-attentive attentional selection, shifts of visual spatial attention, translation invariant 2-D pattern representation, learning and recognition;

(iv) neural circuit for top-down (memory) guided attentional selection and visual search in a cluttered bottom-up input.

At the end of the chapter we will provide a minimum 2-D neural circuit that can simulate and explain a recent neurobiological experiment on the neural basis of memory guided visual search in a monkey (Chelazzi *et al.* 1993).

6.2 Detection of Match/Mismatch and Resonant States

In Chapters 4 we have demonstrated that the facilitatory presynaptic feedback from one neural Field of competitive neurons to another can cause selective resonance (or a match in spatial activity) between the two if both have a common spatial pattern. In all the neural circuits discussed and simulated so far, whenever it was required to measure the degree of match, we have used a simple algorithmic procedure (i.e., by calculating the cosine of the angle between two multidimensional vectors). This computation, although performed by a slightly different algorithmic method in ART based neural networks, forms a crucial step in ART's orienting subsystem as it compares the bottom-up inputs with the top-down memory. The result of this comparison is then used to control whether the network is to learn the input or whether it is to be reset (in which case a memory search gets engaged). While being computational simple and easy to implement, these algorithmic procedures for evaluating the degree of match is not suitable for more general neural network designs where it may be required to use the degree of match/mismatch to regulate system dynamics. It leaves us now to propose a neural mechanism for the detection of resonant or matching states.

Figure 6.1 shows how we can detect whether spatial patterns across two interacting neural Fields F0 and F1 are matched or mismatched. As shown, cells from each neural Field project both excitatory and modulatory signals to two competing neurons, M (the match neuron) and N (the mismatch neuron). For simplicity we have only shown one set of synaptic pathway into the match/mismatch cells, although there are as many such pathways as the number of neurons in the two interacting Fields.

The activity of cell M indicates the degree of match while the activity of cell N indicates the degree of mismatch between the neural patterns of activity across the two Fields. Equations (6.1) and (6.2) describe the activities of the two competing neurons, M and N.

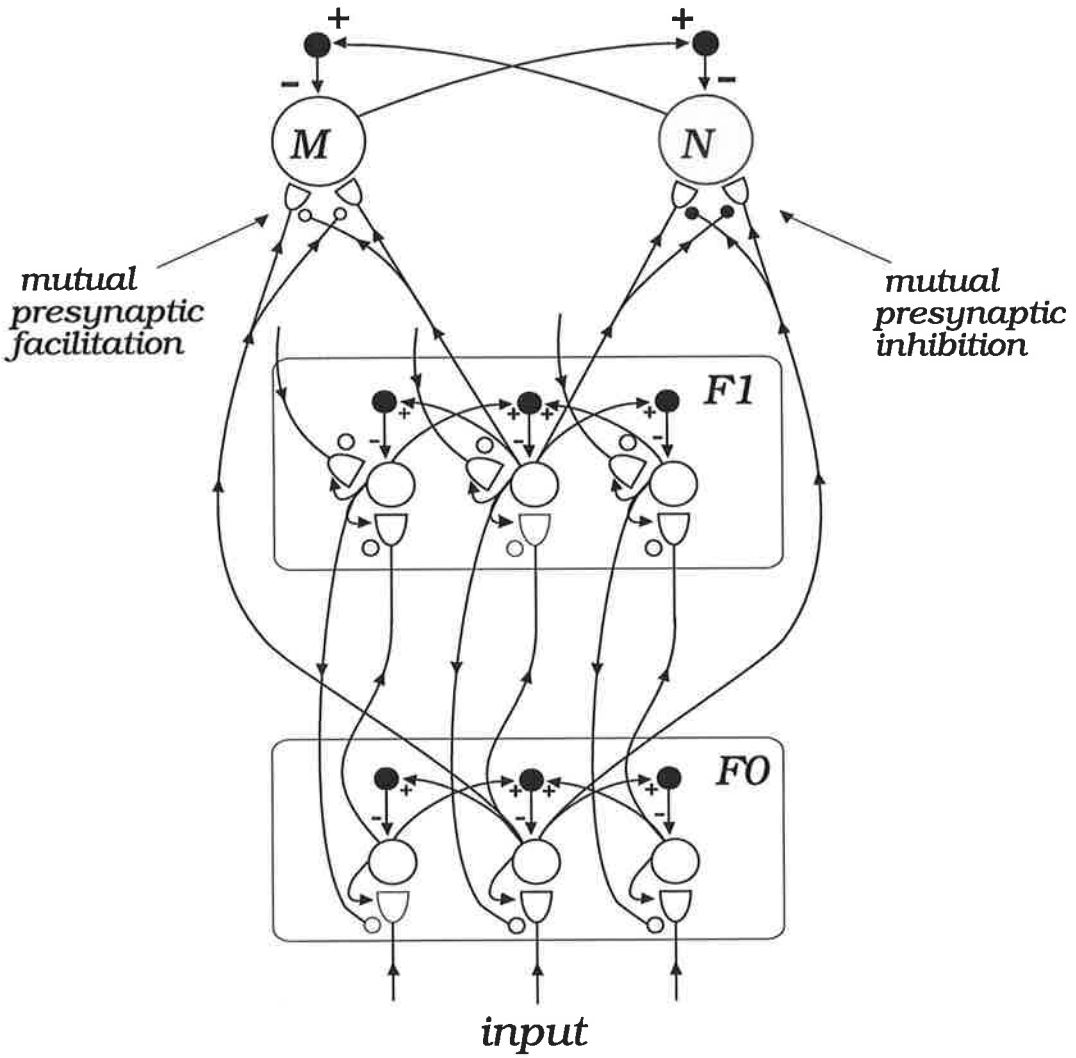


FIGURE 6.1. Neural mechanism for the detection of match/mismatch between the spatial patterns across two interacting neural Fields F0 and F1.

$$\frac{dM}{dt} = -A_M M + (B_M - M)G_M V_M - M\overline{G}_M \overline{v}_M \tag{6.1}$$

$$\frac{dN}{dt} = -A_N N + (B_N - N)G_N V_N - N\overline{G}_N \overline{v}_N \tag{6.2}$$

where A_M and A_N are the passive decay rates of the match and the mismatch neuron respectively; B_M and B_N are their respective upper saturation levels; G_M and G_N are their respective excitatory postsynaptic gains; $\overline{G}_M$ and $\overline{G}_N$ are their respective lateral inhibition gains; V_M and V_N are their respective total excitatory postsynaptic potentials; $\overline{v}_M$ and $\overline{v}_N$ are the potentials on their respective slow inhibitory interneurons that mediate the competition between the two cells.

These equations may be solved at equilibrium to give

$$M = \frac{B_M G_M V_M}{A_M + G_M V_M + \overline{G}_M \overline{v}_M} \tag{6.3}$$

$$N = \frac{B_N G_N V_N}{A_N + G_N V_N + \overline{G}_N \overline{v}_N} \tag{6.4}$$

The competitive interactions between the match/mismatch neurons is mediated by slow inhibitory neurons whose activation levels, $\overline{v}_M$ and $\overline{v}_N$, are given by

$$\frac{d\overline{v}_M}{dt} = -\overline{A}_M \overline{v}_M + \overline{B}_M f(N) \tag{6.5}$$

$$\frac{d\overline{v}_N}{dt} = -\overline{A}_N \overline{v}_N + \overline{B}_N f(M) \tag{6.6}$$

where f(.) is a thresholding function (linear above a threshold). Below we describe the synaptic dynamics into the match/mismatch neurons.

Synaptic dynamics of the match cell M

The total excitatory postsynaptic potential V_M acting on neuron M is given by

$$V_M = \sum_i \sum_j v_{ij}^{F_0 M} + \sum_i \sum_j v_{ij}^{F_1 M} \tag{6.7}$$

where $v_{ij}^{F_0 M}$ is the $(i,j)^{th}$ synaptic input from Field F0 to cell M and is given by (6.8), while $v_{ij}^{F_1 M}$ is the $(i,j)^{th}$ synaptic input from Field F1 to cell M and is given by a similar equation.

$$\frac{dv_{ij}^{F_0 M}}{dt} = -D v_{ij}^{F_0 M} + f\left(x_{ij}^{F_0}\right)\left[\rho_v^{F_0 M} + K_v^{F_0 M} f(M)\right]\left[y_{ij}^{F_0 M} - Y^{F_0 M}\right]^+ \tag{6.8}$$

where D is the passive decay rate of the EPSP. The second term in the above equation represents two contributions to the EPSP: (i) an input driven contribution via the term $f\left(x_{ij}^{F_0}\right)\rho_v^{F_0M}\left[y_{ij}^{F_0M} - Y^{F_0M}\right]^+$ that is independent of the postsynaptic activity; and (ii) a contribution that is due to the correlated pre-postsynaptic activity (given by the term $f\left(x_{ij}^{F_0}\right)K_v^{F_0M} f(M)\left[y_{ij}^{F_0M} - Y^{F_0M}\right]^+$). As in Chapter 4, the latter is assumed to exert a greater effect (i.e., by choosing $K_v^{F_0} > \rho_v^{F_0}$).

Stored transmitter

$$\frac{du_{ij}^{F_0M}}{dt} = \alpha_u^{F_0M}\left(1 - u_{ij}^{F_0M}\right) - \left[\beta_u^{F_0M} + \rho_u^{F_0M} f\left(x_{ij}^{F_0}\right) + K_u^{F_0M} f(M) f\left(x_{ij}^{F_0}\right)\right]\left(u_{ij}^{F_0M} - y_{ij}^{F_0M}\right) \tag{6.9}$$

The term $-K_u^{F_0M} f(M) f\left(x_{ij}^{F_0}\right)\left(u_{ij}^{F_0M} - y_{ij}^{F_0M}\right)$ in the above equation says that the stored transmitter is depleted by the correlated firing of the postsynaptic neuron M and the synaptic input signal $f\left(x_{ij}^{F_0}\right)$. Note that all synapses have a fixed transmitter production rate ($z_{ij}^{F_0M} = 1, \forall(i,j)$).

Mobilized transmitter

$$\frac{dy_{ij}^{F_0M}}{dt} = -\gamma_y^{F_0M} y_{ij}^{F_0M} - f\left(x_{ij}^{F_0}\right)\left[\rho_y^{F_0M} + K_y^{F_0M} f(M)\right]\left[y_{ij}^{F_0M} - Y^{F_0M}\right]^+ + \left[\beta_y^{F_0M} + H_y^{F_0M} f\left(x_{ij}^{F_1}\right)\right]\left(u_{ij}^{F_0M} - y_{ij}^{F_0M}\right) \tag{6.10}$$

where $\beta_y^{F_0M}$ is the tonic level of transmitter mobilization in the excitatory synaptic pathways abutting cell M. The term $H_y^{F_0M} f\left(x_{ij}^{F_1}\right)\left(u_{ij}^{F_0M} - y_{ij}^{F_0M}\right)$ in equation (6.10) says that the $(i,j)^{th}$ synapse from Field F0 to cell M is facilitated by the $(i,j)^{th}$ cellular activity of Field F1, where $H_y^{F_0M}$ is the facilitatory gain. Similar set of equations apply to the $F1 \rightarrow M$ synapses.

Synaptic dynamics of the mismatch cell N

The excitatory set of synaptic pathways into the mismatch cell N also have the same dynamics but, instead of mutual presynaptic facilitation, these pathways are engaged in mutual presynaptic inhibition. That is, the corresponding cells in Fields F0 and F1 presynaptically inhibit the transmitter mobilization process in the synapse of the other. The following equation represents the dynamics of the transmitter mobilization process in a synapse from Field F0 to cell N.

$$\frac{dy_{ij}^{F_0N}}{dt} = -\gamma_y^{F_0N} y_{ij}^{F_0N} - f\left(x_{ij}^{F_0}\right)\left[\rho_y^{F_0N} + K_y^{F_0N} f(N)\right]\left[y_{ij}^{F_0N} - Y^{F_0N}\right]^+$$

$$+ \frac{\beta_y^{F_0N}}{\left[1 + \overline{H}_y^{F_0N} f\left(x_{ij}^{F_1}\right)\right]}\left(u_{ij}^{F_0N} - y_{ij}^{F_0N}\right) \tag{6.11}$$

The term $\frac{1}{\left[1+\overline{H}_y^{F_0N} f\left(x_{ij}^{F_1}\right)\right]}$ in the above equation is the inhibitory term that reduces the transmitter mobilization rate of the excitatory $(i,j)^{th}$ synapse from Field F0 to cell N when a neuron in the corresponding location of Field F1 is activated above its threshold. When the mobilized transmitter in all the active excitatory synaptic pathways to cell N is reduced, the total EPSP on cell N is also reduced, while the EPSP on cell M is increased, thus enabling cell M to become active. However, this can only occur if the spatial patterns of activity across Fields F0 and F1 are well matched. The degree of match that needs to be met before cell M becomes active can be controlled by the tonic level of transmitter mobilization in the synapses. A higher level of tonic transmitter mobilization in the excitatory synaptic pathways to cell N means that if cell M is to be activated, then the degree of match between the spatial patterns across Fields F0 and F1 needs to be very high.

6.2.1 Simulation of the Match/Mismatch Neural Circuit

Below we provide two computer simulations that demonstrate the match/mismatch neural circuit on oscillatory neural layers. However, rather than simulating the circuit shown in Figure 6.1, we provide computer simulations of the circuit shown in Figure 6.2 whose neurons are modelled by oscillatory dynamics (refer to Chapter 4, section 4.5.2).

Note that the circuit shown below differs from that shown in Figure 6.1. In the circuit below, both neural Fields receive the same bottom-up input (whereas in the circuit of Fig. 6.1, only Field F0 receives the bottom-up input, while Field F1 receives its excitatory input from F0). The purpose of simulating this circuit is to demonstrate that: (i) the facilitatory presynaptic feedback can synchronize the two oscillatory neural Fields; and (ii) the match neuron will become active only when the two are synchronized. As in section 4.5.2 (Chapter 4), each oscillatory neural Field consists of 20 excitatory neurons and 20 slow inhibitory interneurons. Each excitatory neuron in the two Fields is given random initial conditions. Cells in each Field are also subjected to the same fixed bottom-up input that is constant in amplitude and time at each spatial location. Neurons within each oscillatory neural layer are coupled via *lateral presynaptic exci-*

tation (see Chapter 4). This ensures cellular synchronization across each layer. Facilitatory presynaptic feedback from one Field to the other ensures synchronization of the two. These neural Fields are thus similar to those described in section 4.5.2 of Chapter 4. Note that the circuit is equally applicable to the case where pulsating Field F0 provides bottom-up excitatory pulses to F1, provided that there is no significant time delay introduced between the two. The non-oscillatory version of the circuit shown in Figure 6.1 will be used in section 6.3.

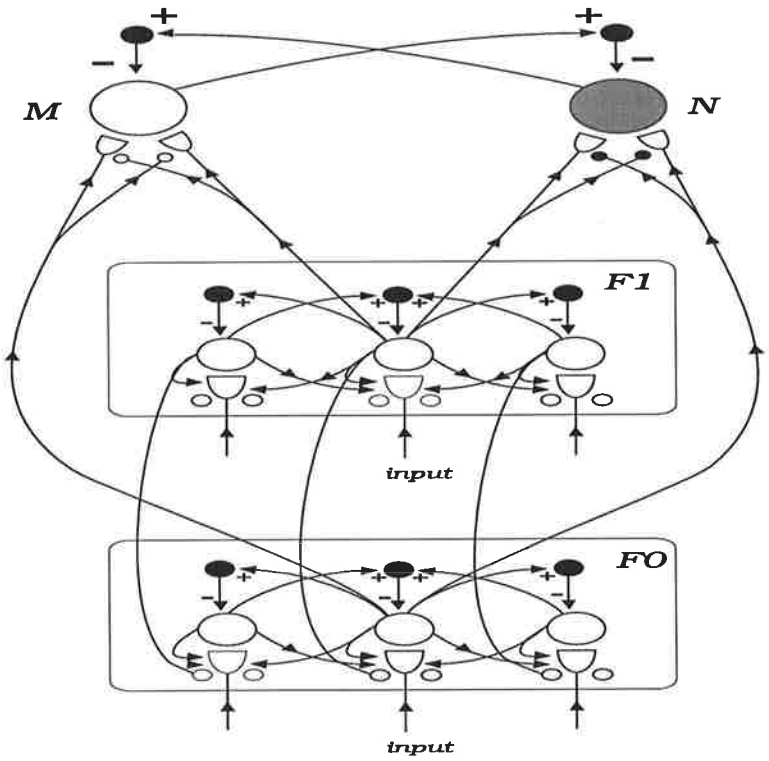


FIGURE 6.2. Simulated match/mismatch neural circuit on oscillatory neural layers.

All differential equations in the circuit (including equations (6.1) and (6.2)) were iterated using Euler's first order approximation method with time increment $\Delta t = 1$. For this time step, the circuit parameters are as follows: $A_M = A_N = 0.01$; $B_M = B_N = 1$; $G_M = G_N = 0.0005$; $\overline{G}_M = \overline{G}_N = 5$; $\overline{A}_M = \overline{A}_N = 0.001$; $\overline{B}_M = \overline{B}_N = 0.001$; threshold for postsynaptic activation $\Theta = 0.3$, $f(x) = \max(x - \Theta, 0)$; $D = 0.5$; $\rho_v^{F_0M} = \rho_v^{F_1M} = 10$; $K_v^{F_0M} = K_v^{F_0M} = 80$; $Y^{F_0M} = Y^{F_1M} = 0.2$; $\alpha_u^{F_0M} = \alpha_u^{F_1M} = 0.05$; $\beta_u^{F_0M} = \beta_u^{F_1M} = 0$; $\rho_u^{F_0M} = \rho_u^{F_1M} = 0.012$; $K_u^{F_0M} = K_u^{F_1M} = 1.2$; $\gamma_y^{F_0M} = \gamma_y^{F_1M} = 0.01$; $\rho_y^{F_0M} = \rho_y^{F_1M} = 0.001$; $K_y^{F_0M} = K_y^{F_1M} = 1$; $\beta_y^{F_0M_1} = \beta_y^{F_1M_1} = 0.004$; $H_y^{F_0M} = H_y^{F_1M} = 0.2$. All the parameters in the synaptic pathways to the mismatch cell are as above, with the exception of the following: $\rho_v^{F_0N} = \rho_v^{F_1N} = 100$; $H_y^{F_0N} = H^{F_1N} = 2$; $\overline{H}_y^{F_0N} = \overline{H}^{F_1N} = 2$. The parameters for the two neural Fields (F0 and F1) are the same as for those in section 4.5.2 of Chapter 4, and are given in Appendix B.1.

For the above model to be applicable to oscillatory neural layers, the threshold in the inhibitory modulating presynaptic pathways must be lower than on the facilitatory pathways. The reason for this is that if the excitatory and the presynaptic inhibitory signals from Fields F0 and F1 arrive simultaneously, presynaptic inhibition may not have sufficient time to prevent the mismatch cell N from firing when the spatial patterns across Fields F0 and F1 are well matched. Alternatively, one can use slower intermediate variables (as in section 4.2.3) to depress the relevant synapses. In the simulation results shown below, the presynaptic inhibitory term in equation (6.11) is

$$\frac{1}{\left[1+\overline{H}_y^{F_0N} f\left(x_{ij}^{F_1}\right)\right]} = \frac{1}{\left[1+\overline{H}_y^{F_0N} x_{ij}^{F_1}\right]} \tag{6.12}$$

That is, there is no threshold on the output signals driving presynaptic inhibition. Figures 6.3 and 6.4 show the simulation results that were obtained with the oscillatory version of the circuit for two cases: (i) interacting neural Fields, where one Field (Field F1) presynaptically facilitates the other (Field F0); and (ii) independently oscillating neural Fields (i.e., there is no interaction between F0 and F1).

Results in Figure 6.3 show that the two neural layers eventually become synchronized. This synchronization takes some time, as sensed by neuron M, which begins to fire at t = 4500 iterations. Since the competition between the two sensing neurons is mediated by inhibitory neurons whose decay rate is slow, it take some time for the excitatory inputs on neuron M to build to a level sufficiently high to overcome the inhibitory input from neuron N. However, when this happens, the flip between M and N is very fast.

Figure 6.4 shows that if the two neural layers do not interact, then they remain out of phase with one another. These simulations thus demonstrate the powerful effect that the facilitatory presynaptic feedback from F1 has on the dynamics of F0. Simulation results in Figure 6.4 (open loop system) show that although cells within each neural layer synchronize, the two layers remain asynchronous (as sensed by neuron N). Although it is not the intention of this thesis to use oscillatory neural layers in circuit designs, above simulations show that, at least for simple circuits, it is feasible to use our proposed neuro-engineering design principles to model oscillatory neural layers. However, since there is no established parameter design procedure nor suitable mathematical methods for stability analysis of oscillatory neural layers, their application is currently limited.

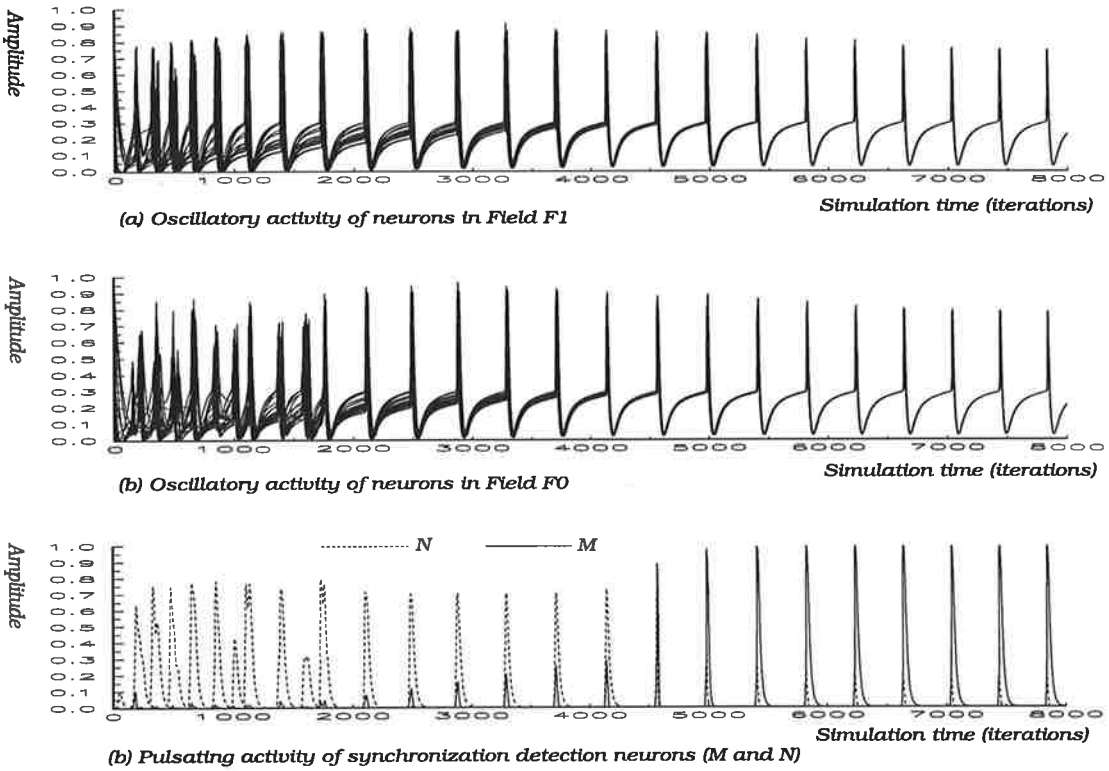


FIGURE 6.3. Detection of synchronized oscillations.

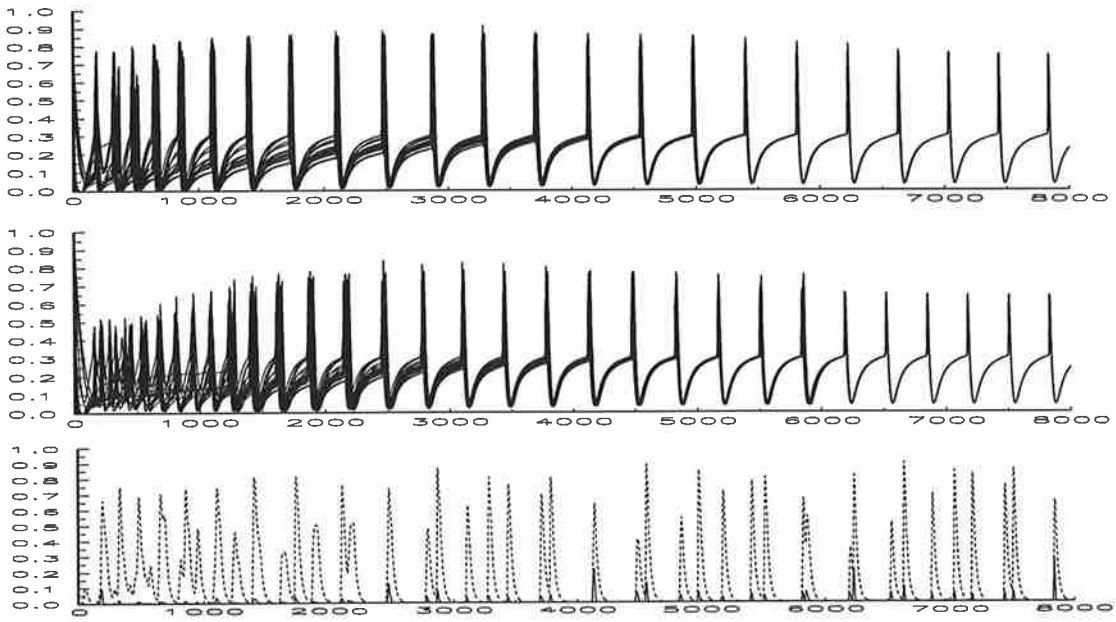


FIGURE 6.4. Detection of asynchronous oscillations.

6.3 Storage and Self-regulation of Memory

In this section we propose how presynaptic modulation may be used in the attentional regulation and the storage of memory, particularly the storage of the Immediate Short Term Memory (ISTM). We define Immediate Short Term Memory as the memory of the most recent spatial pattern (or neural activity) that is retained after the input offset. One of the ways that the ISTM may be retained after the input offset is by an excitatory reverberatory loop between two neural Fields. In Chapter 3 we have seen how the reverberatory loop is implemented in the ART-3 neural network. Current neural network models and theories (including ART) do not provide a plausible solution to how such reverberations may be modulated by attentional factors. Grossberg (1976) has proposed that one source of non-specific arousal modulating signals may be due to a mismatch between a bottom-up input and a top-down expectancy. However, due to the lack of plausible design principles by which such modulatory effects may be embedded in a real-time neural network, none of the ART based neural networks developed to date model such mechanisms. In this section we propose neuro-engineering design principles for the design of self-regulated attentional neural circuits whose inputs and reverberations may be presynaptically modulated by the facilitatory/inhibitory synaptic gain control signals.

To begin with, we first consider the schematic of a reverberatory neural circuit shown in Figure 6.5(a). The circuit consists of three presynaptically modulated shunting competitive neural layers (or Fields), denoted by $F0, F1$ and $F2$. Each of the shown Fields is assumed to be modelled by equilibrium steady state activity (i.e., non-pulsating dynamics). The bottom-up 2-D spatial pattern enters into the circuit via Field F0, which provides excitatory synaptic inputs to $F1$. Field F1 interacts with Field $F2$ through a mutual excitatory reverberatory loop. Field $F1$ also provides top-down presynaptic facilitation of Field $F0$. The excitatory reverberatory loop between $F1$ and $F2$ will maintain the ISTM of the bottom-up input once the input is removed. It is our aim to investigate how the reverberation between Fields F1 and F2 may be modulated and how the strength of this modulation affects the integrity of the stored ISTM when the input is either offset or is replaced by another 2-D spatial pattern that may or may not embed the original pattern in a cluttered background.

Let us now suppose that the strength of the reverberation between $F1$ and $F2$ is presynaptically modulated by facilitatory and inhibitory synaptic gain control signals, as shown in Figure 6.5(b). If F1 and F2 are coupled by a strong facilitatory gain control signal, then the ISTM activity will not only be higher in magnitude but will also persist for a longer period of time when the input is removed. Since in general there may be a need to modulate the strength of these reverberations, thus modulating the duration

of ISTM, we need two modulatory signals (F - facilitatory and I - inhibitory) to presynaptically modulate transmitter mobilization in the respective pathways. These modulatory signals may be generated elsewhere in the system by a pair of competing neurons or circuits. For example, signals F and I may be generated by at least two types of circuits: (a) by a neural circuit that detects novelty/familiarity of the input stimulus; or (b) by a neural circuit that detects whether the spatial pattern across F0 and F1 are matched or mismatched.

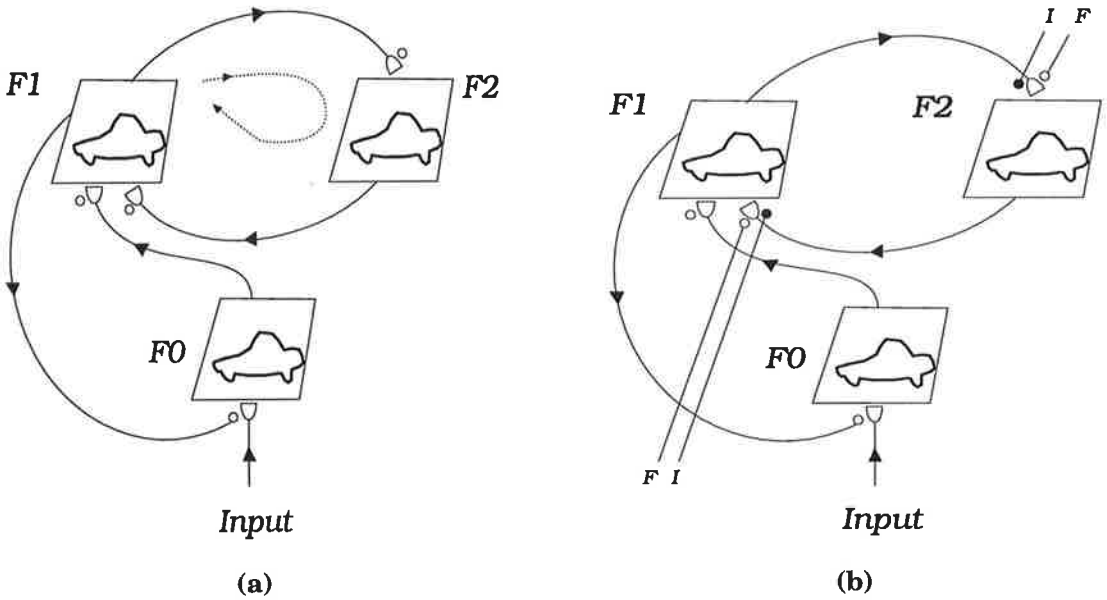
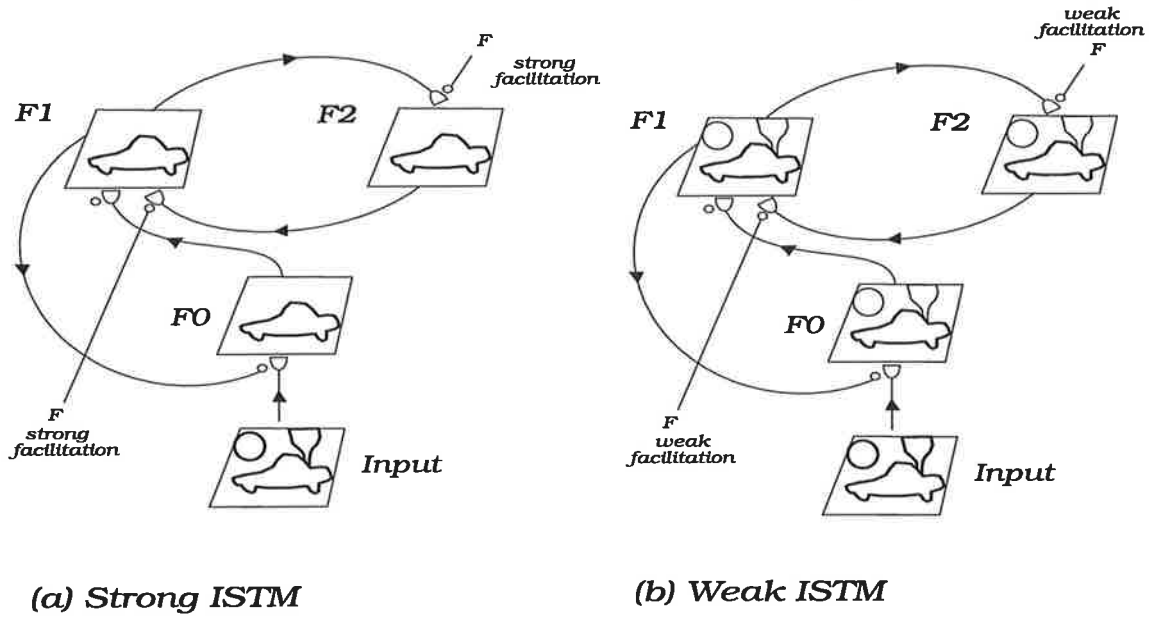


FIGURE 6.5. Schematic of a reverberatory attentional neural circuit for the maintenance of the Immediate Short Term Memory (ISTM). Reverberatory activity between F1 and F2 maintains ISTM, the strength of which may be presynaptically modulated, as shown in (b).

As an example, let us assume that the presynaptic signals are driven by the novelty/familiarity part of the system (as yet unspecified). Then, if the bottom-up input is familiar the strength of the reverberation should be decreased (i.e., pay less attention to familiar stimuli - a well established result from psychology). Otherwise we would want the opposite, i.e., an increase in the strength of the reverberation. However, this should occur only after the familiar stimulus is recognized (otherwise one cannot conclude whether the stimulus is familiar or novel) and when it is determined that the familiar stimulus is of no immediate interest. Thus, when the other parts of the system reach a consensus that the stimulus is novel, we should expect that strong facilitatory signals are issued to increase the strength of ISTM reverberation. Otherwise, it is assumed that a weak facilitatory and a strong presynaptic inhibitory signal is issued.



(a) Strong ISTM *(b) Weak ISTM*

FIGURE 6.6. Strong versus weak ISTM in a reverberatory attentional neural circuit.

If a bottom-up input stimulus is retained in the ISTM by a strong reverberatory loop between F1 and F2 and if the sensitivity of F1 to the inputs from F2 is much larger than it is to the bottom-up input from F0, then as schematized in Figure 6.6(a), the neural pattern of activity across F1 (the car boundary) should be able facilitate the same object boundary when it is embedded in a cluttered bottom-up input. If on the other hand the retained ISTM is not very strong, more of the non-matching bottom-up input will enter into the circuit, as schematized in Figure 6.6(b), thus corrupting the original ISTM. As indicated, the final spatial pattern that ends reverberating in the circuit depends on the relative strengths of the original ISTM activity and the bottom-up input. Because the strength of the facilitatory top-down presynaptic feedback from Field F1 to Field F0 is indirectly affected by these modulatory mechanisms, it may be concluded that the circuit's level of attention is regulated by the combination of the F-I signals.

Figure 6.7 shows a three layered neural circuit that maintains the ISTM in the reverberatory loop between Field F1 and F2. For simplicity we have shown only one set of modulatory pathways that abut the bottom-up inputs into Field F2.

Figures 6.8 and 6.9 show the simulation results that were obtained with the above neural circuit for two examples. In each of the simulations we first store a reference 2-D shape (shape 1) in the reverberatory neural activity by presenting it to the circuit for ten iterations of the circuit equations. This is then followed by a new input for another ten iterations. In the first example shown in Figure 6.8, the second

input embeds shape 1 in a cluttered background, whereas in the second example shown in Figure 6.9, the second input contains a new shape. We should thus expect the final steady states for the two examples to differ.

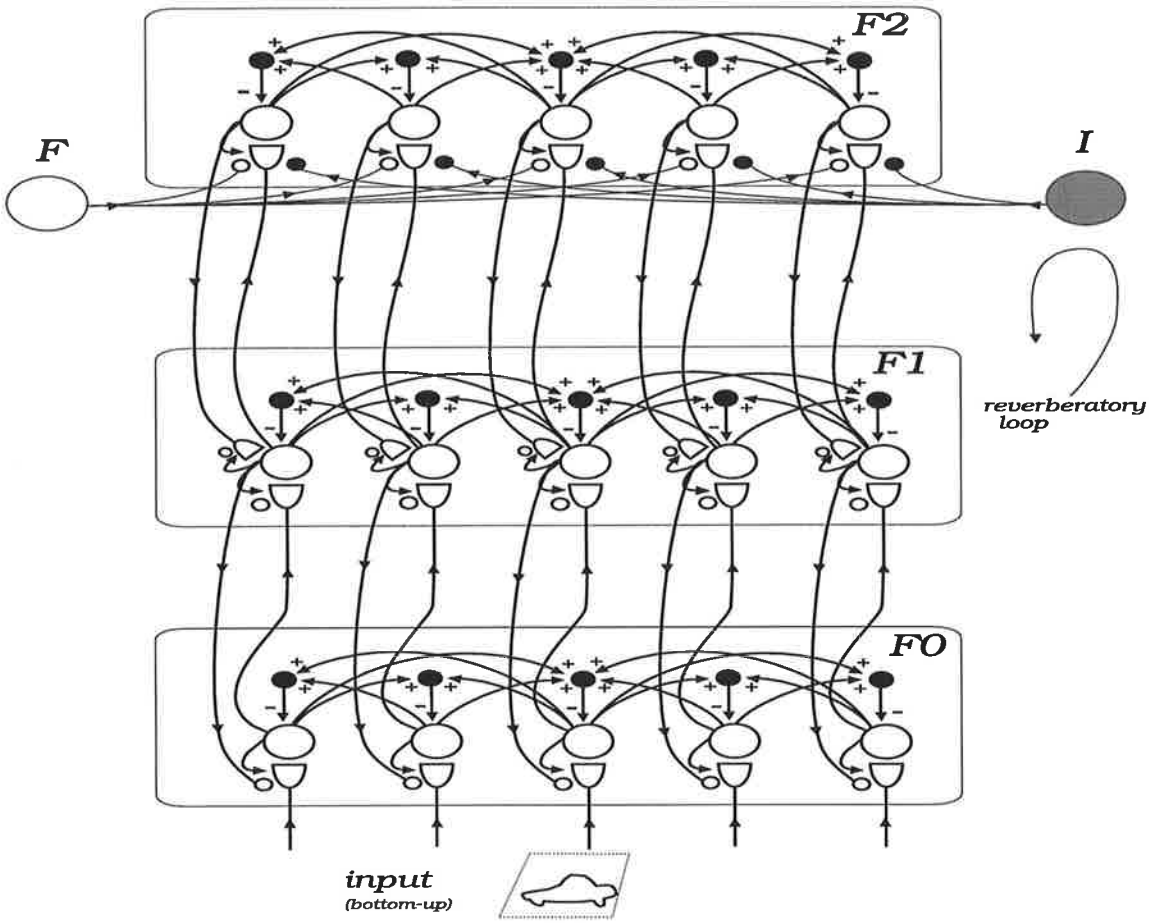


FIGURE 6.7. Attentional neural circuit for ISTM storage in reverberatory neural Fields.

As can be seen in Figure 6.8 at time t = 20 iterations, the steady state response of the reverberatory neural circuit shows that the stored ISTM has not been corrupted by the second input. The reverberating pattern has also facilitated the transmission of the matching portions of the bottom-up input, thus causing a steady state match between the ISTM across F1 and the attended portion of the input pattern at F0. The strong circuit reverberation and the competitive effects in each Field have thus prevented the degradation of the ISTM. If the strength of the reverberation was much weaker, then the activity across the three Fields (F0, F1 and F2) would be a combination of the ISTM and the cluttered background.

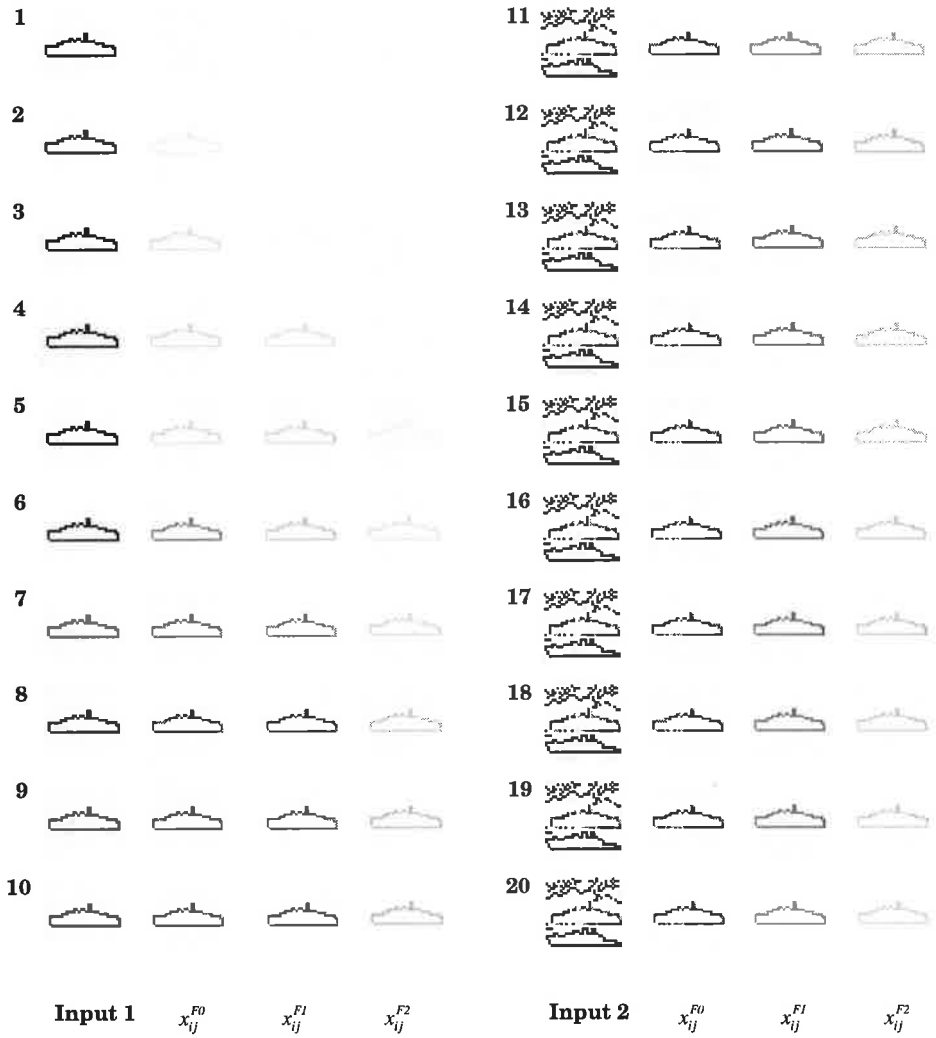


FIGURE 6.8. Steady-state simulation results on the effect of strong ISTM on the bottom-up neural signal transmission in a cluttered input.

The relative intensity of each portion would thus, in the steady state, depend on the strength of the initial ISTM reverberation (as was indicated previously in Fig 6.6(b)). However, because the resultant 2-D spatial pattern across F1 is always engaged in the facilitation of the corresponding bottom-up pathways into F0, the steady state response will always end up in a match between F0 and F1 (provided that every subsequent input contains the original pattern). This reverberatory neural circuit is thus more flexible and capable than the reverberatory loops as currently implemented in ART based neural networks.

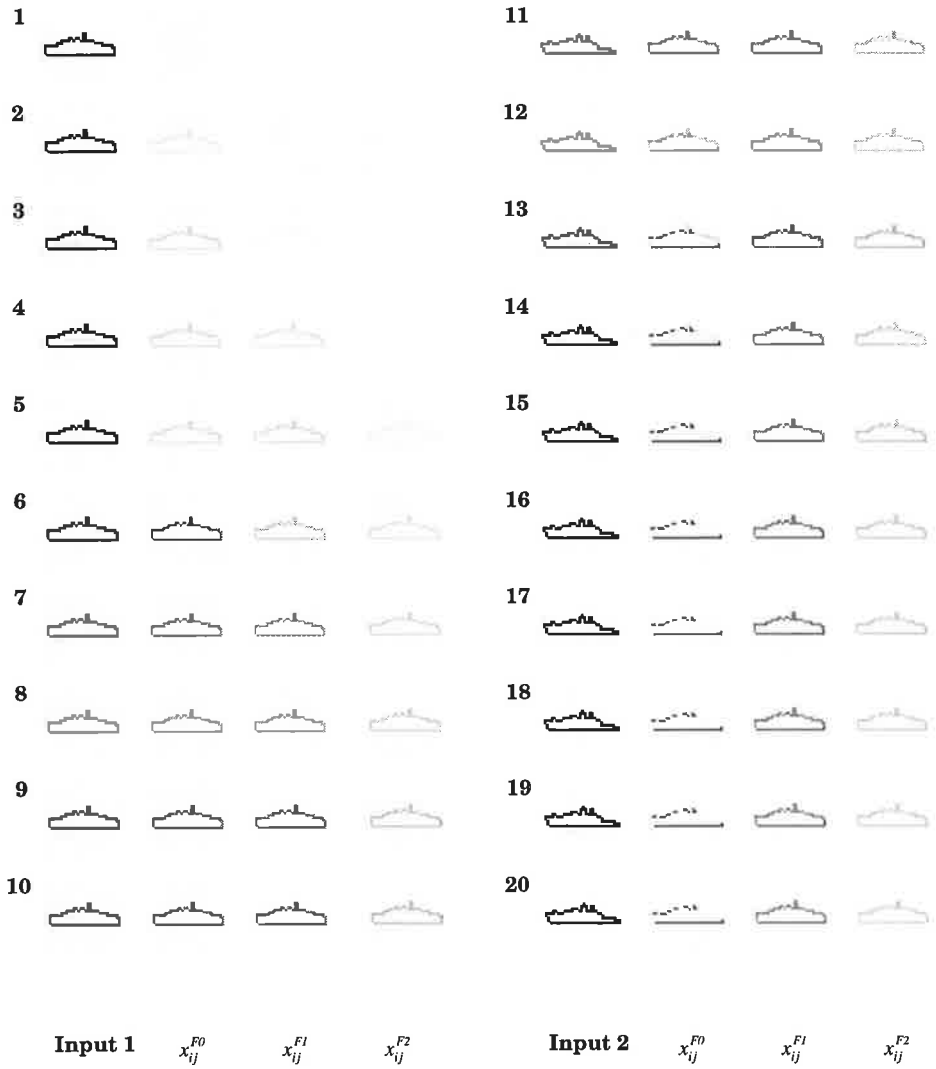


FIGURE 6.9. Steady-state simulation results on the effect of strong ISTM in the presence of a new input stimulus.

The simulation data in Figure 6.9 shows that because the strong reverberating ISTM (due to shape 1) cannot match with the portion of the bottom-up input that it can facilitate, a steady state is reached where the degree of match is below the required tolerance level (which was chosen to be 0.98). Figure 6.10 shows the degree of match between Fields F0 and F1 for the two examples considered above. Note that at the instant the second input is presented to the circuit, the degree of match in the first example (graph (a)) has fallen below the required match threshold but was soon elevated and remains above the threshold, whereas in the second example it has fallen and remains below the match threshold.

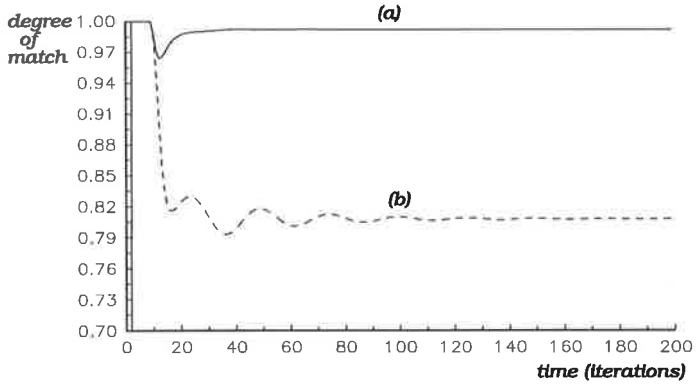


FIGURE 6.10. Degree of match between the reverberating ISTM and the selected portion of the bottom-up input. (a) circuit dynamics during a matched phase; (b) circuit dynamics during a mis-matched phase.

In the next section we propose how the modulatory synaptic signals may be generated and used in a simple attentional neural circuit.

6.3.1 Self-regulated Attentional Neural Circuit

In the previous section we have discussed the simple case of memory storage as a spatial pattern of a reverberatory activity between two competitive neural Fields and have not provided the mechanism by which the circuit can autonomously self-regulate its degree of attention. We now extend the above concepts by first proposing how one can design a new circuit that can detect whether its input is familiar or novel. We then propose a simple example to indicate how the resultant signals may be used to provide self-regulated attentional modulation during learning. It is well known that novel stimuli attract a higher degree of attention and that a memory of an external sensory event is highly dependant on the degree of attention that the event has received at the time of its occurrence, (Rock and Gutman, 1981). A high degree of attention presumably ensures stronger memory encoding, and vice versa.

The schematic shown in Figure 6.11 indicates that in a neural circuit that consists of three competitive Fields, two pairs of match/mismatch neurons are needed to regulate various synaptic pathways. One pair regulates the bottom-up synapses in the pathway $F0 \rightarrow F1$, while the other pair regulates the top-down synapses in the pathway $F2 \rightarrow F1$. The memory in the circuit is held in the top-down synapse of Field F2 and is learned. As shown, the signals from Field F1 to the top-down synapses of Field F2 provide the spatial pattern that is to be stored in the top-down memory. Figure 6.12 shows the equivalent neural circuit, whose learning mechanism in the top-down memory pathways is schematized in Figure 6.13.

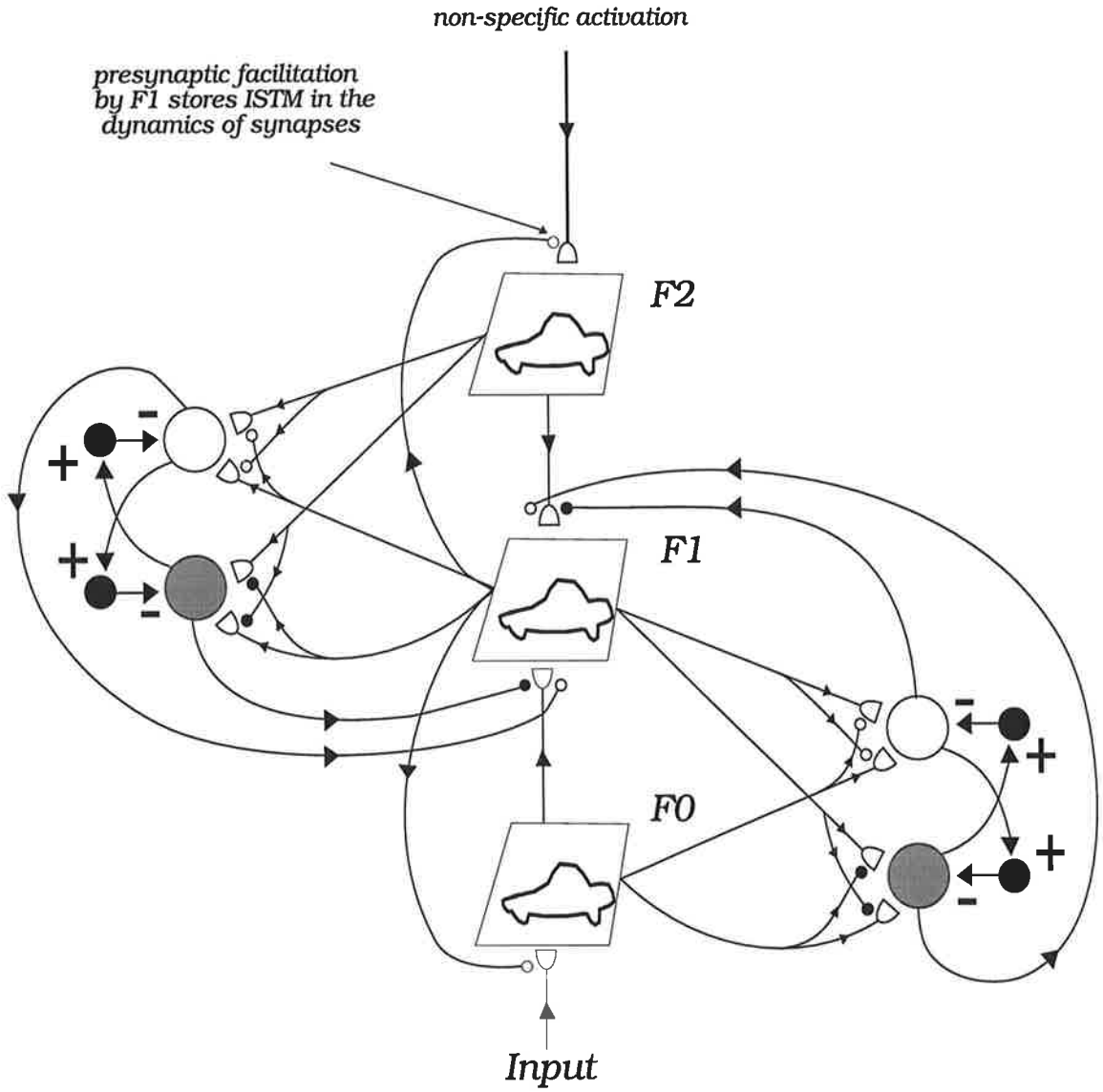
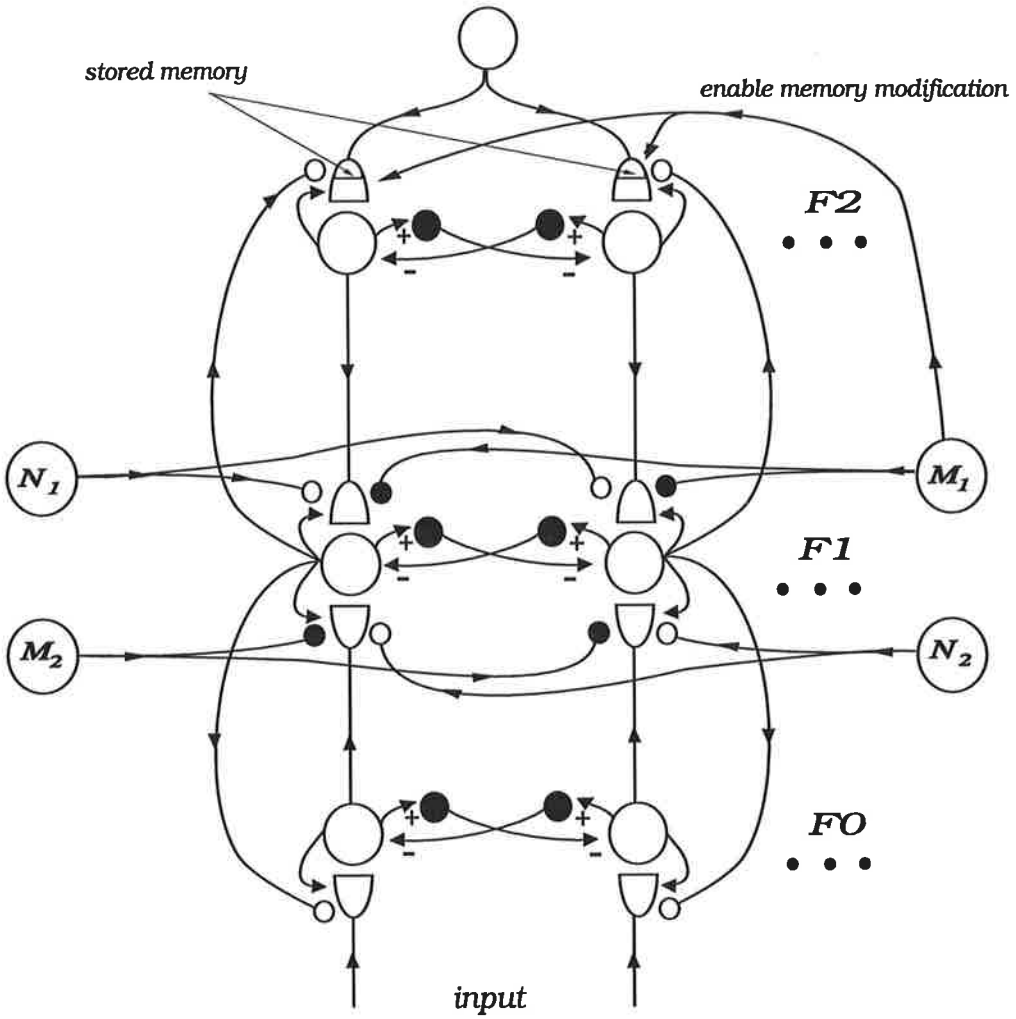


FIGURE 6.11. Neural scheme for familiarity/novelty detection and self-regulated attentional modulation. The memory within the circuit is learned and is stored within the top-down synapse abutting Field F2. Two pairs of match/mismatch neurons regulate the synaptic transmission gains in the bottom-up and the top-down synapses to Field F1. For a novel stimulus, the cellular activity within the circuit will be high since the mismatch neurons will facilitate the shown synapses, thus increasing the strength of signals into the neurons of Field F1. Since there is no previously established top-down memory, Field F2 will initially be inactive. As the memory strength increases it begins to activate Field F2. The match neurons then inhibit the transmitter mobilization of the top-down synapses into Field F1, thus reducing the cellular activity and hence the incremental rise in the memory strength.



M_1 - *Measure of match between F0 and F1*

N_1 - *Measure of mismatch between F0 and F1*

M_2 - *Measure of match between F1 and F2*

N_2 - *Measure of mismatch between F1 and F2*

FIGURE 6.12. Simple neural circuit capable of single memory storage and attentional regulation by familiarity/novelty of the bottom-up input.

The following three equations describe the facilitatory/inhibitory modulation of the synapses in the relevant pathways. All other equations remain same as per presynaptically modulated shunting competitive neural layer. Equations for the two sets of match/mismatch neurons are as given in section 6.2.

Top-down transmitter mobilization into Field F2

$$\frac{dy_{ij}^{F_2}}{dt} = -\gamma_y^{F_2} y_{ij}^{F_2} - T_{ij}^{F_2}\left[\rho_y^{F_2} + K_y^{F_2} f\left(x_{ij}^{F_2}\right)\right]\left[y_{ij}^{F_2} - Y^{F_2}\right]^+ + \beta_y^{F_2}\left(u_{ij}^{F_2} - y_{ij}^{F_2}\right) \tag{6.13}$$

where $T_{ij}(= 1, \forall(i,j))$, is the top-down input into Field F2, which initially does not excite Field F2 because the top-down synaptic memory variables (z_{ij}^{F2}, u_{ij}^{F2} and y_{ij}^{F2}) into the Field are initialized to below a threshold for transmitter release (these are learned, as shown in Figure 6.13).

Bottom-up transmitter mobilization into Field F1 (pathway $F0 \rightarrow F1$)

$$\frac{dy_{ij}^{F_0F_1}}{dt} = -\gamma_y^{F_0F_1} y_{ij}^{F_0F_1} - f\left(x_{ij}^{F_0}\right)\left[\rho_y^{F_0F_1} + K_y^{F_0F_1} f\left(x_{ij}^{F_1}\right)\right]\left[y_{ij}^{F_0F_1} - Y^{F_0F_1}\right]^+ + \frac{\left[\beta_y^{F_0F_1} + H_y^{F_0F_1} f(N_2)\right]}{\left[1 + \overline{H}_y^{F_0F_1} f(M_2)\right]}\left(u_{ij}^{F_0F_1} - y_{ij}^{F_0F_1}\right) \tag{6.14}$$

This equation says that the transmitter mobilization rate in the synaptic pathways from Field F0 to Field F1 is regulated by the activity of match/mismatch neuron pair M_2 and N_2. Neuron M_2 inhibits the transmitter mobilization rate, while neuron N_2 facilitates the synapse. Since the two competing neurons measure the match/mismatch between the spatial patterns across Field F1 and F2, it may be said that the degree of this match will determine the amplitude of the bottom-up inputs into Field F1.

Top-down transmitter mobilization into Field F1 (pathway $F2 \rightarrow F1$)

$$\frac{dy_{ij}^{F_2F_1}}{dt} = -\gamma_y^{F_2F_1} y_{ij}^{F_2F_1} - f\left(x_{ij}^{F_2}\right)\left[\rho_y^{F_2F_1} + K_y^{F_2F_1} f(x_{ij}^{F1})\right]\left[y_{ij}^{F_2F_1} - Y^{F_2F_1}\right]^+ + \frac{\left[\beta_y^{F_2F_1} + H_y^{F_2F_1} f(N_1)\right]}{\left[1 + \overline{H}_y^{F_2F_1} f(M_1)\right]}\left(u_{ij}^{F_2F_1} - y_{ij}^{F_2F_1}\right) \tag{6.15}$$

This equation says that the transmitter mobilization rate in the synaptic pathways from Field F2 to Field F1 is regulated by the activity of match/mismatch neuron pair M_1 and N_1. Hence the degree of match between spatial patterns across Fields F0 and F1 will determine the strength of the top-down signals from Field F2 to Field F1.

The memory trace of the input stimulus is stored in the transmitter dynamics of the top-down synapses that abut Field F2. Specifically, the transmitter production level, z_{ij}^{F2}, in the $(i,j)^{th}$ top-down synapse adapts to the sum of signals from Fields F1 and F2. This adaptation is specified by the following gated long term memory equation.

$$\frac{dz_{ij}^{F2}}{dt} = -\delta(z_{ij}^{F2} - u_{ij}^{F2}) + rT_{ij}\varepsilon[0.5f(x_{ij}^{F1}) + 0.5f(x_{ij}^{F2}) - z_{ij}^{F2}] \tag{6.16}$$

where r is the learning rate ($r \leq 1$), T_{ij} (= 1, $\forall(i,j)$) is a top-down driving signal (i.e., all top-down synapses into Field F2 receive the same bias) and ε is a gating signal that enables memory modification. Because the top-down synapses are initialized to below the threshold for transmitter release, the top-down input into Field F2 does not excite its neurons until the synaptic memory strength has increased (i.e., Field F2 remains inactive until its top-down synapses have learned a spatial input pattern). The term $-\delta(z_{ij}^{F2} - u_{ij}^{F2})$ in the above equation says that the transmitter production rate decays via the storage transmitter (the importance of this term will be discussed in the next chapter). The factor of 0.5 in the second term ensures that the top-down transmitter production level (LTM) is bounded.

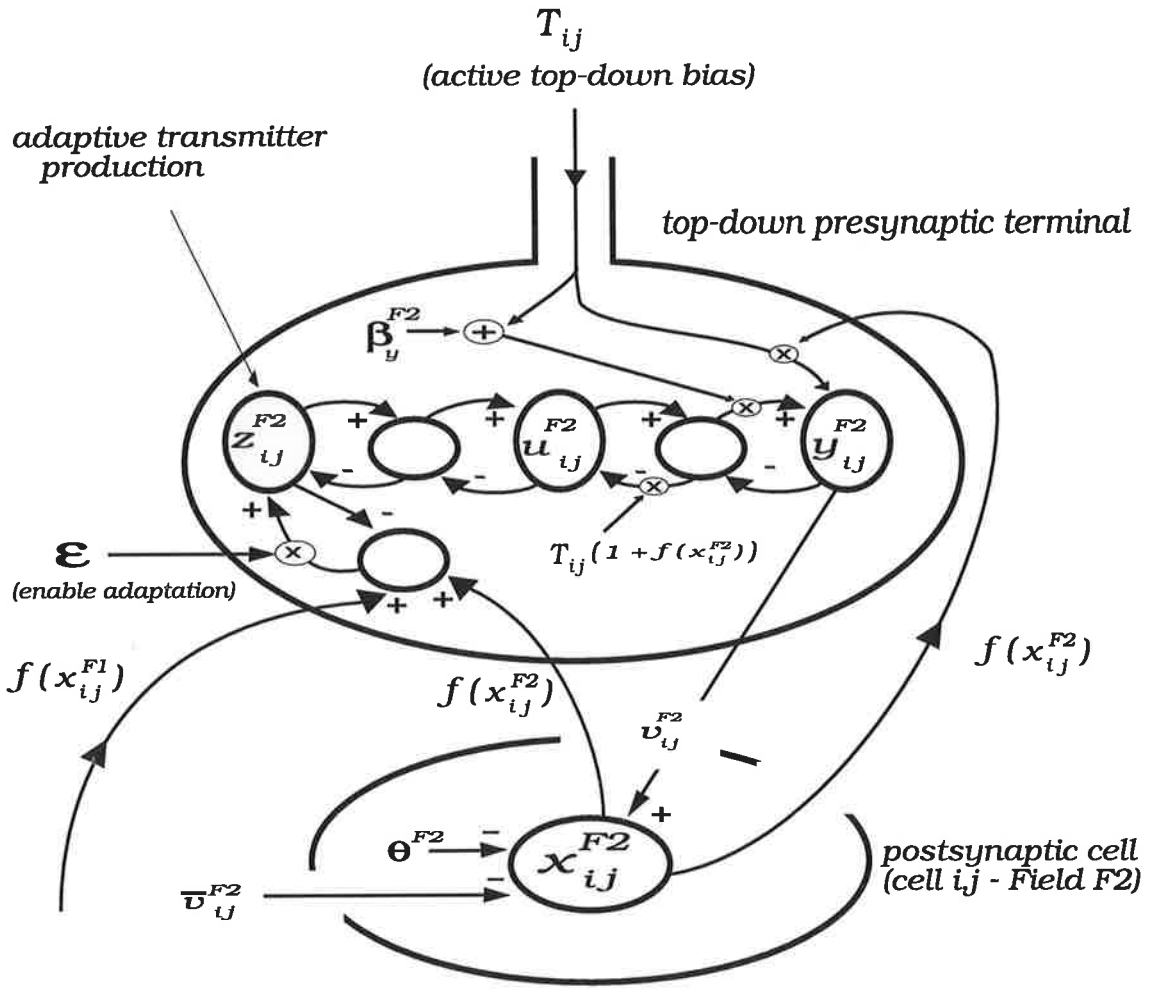


FIGURE 6.13. Synaptic mechanism for regulated storage of top-down memory.

Memory modification is allowed only when $\varepsilon = 1$. The learn enable signal (ε) is specified as

$$\varepsilon = \left\{ \begin{array}{cc} 1 \quad if \quad M1 \geq \Theta^M \quad \text{and} \quad \Delta(M1 + N1) < \delta_1 \\ 0 \quad otherwise \end{array} \right\} \tag{6.17}$$

where Θ^M is the threshold that must be exceeded by the match neuron $M1$, $\Delta(M1 + N1)$ is the time rate change of the combined activity of $M1$ and $N1$ (which measures the stability of the match); δ_1 is the threshold below which the network is assumed to be in a steady state. This equation says that the combined rate of change of activity must be below a threshold for the memory modification to be enabled. Note that the learn enable signal is obtained from the activity of the lower two neural Fields (F0 and F1).

Detection of Novelty/Familiarity

Since Field F1 also receives excitatory top-down signals from Field F2, the net spatial pattern of neural activity across Field F1 is therefore a combination of the top-down memory and the bottom-up input, Figure 6.14.

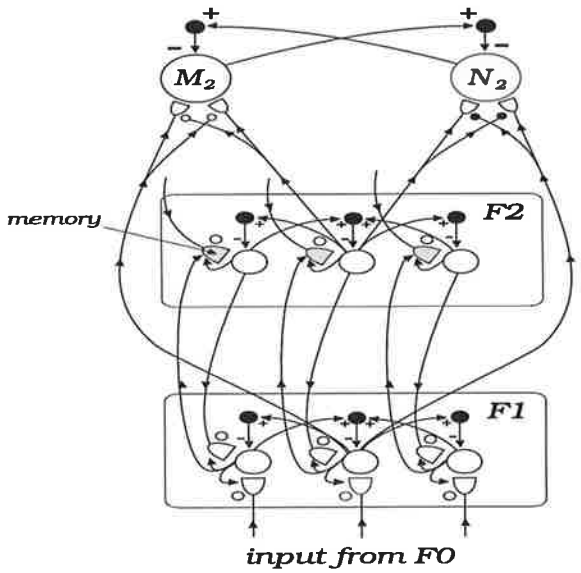


FIGURE 6.14. Detection of novelty/familiarity in a self-regulated competitive artificial neural circuit.

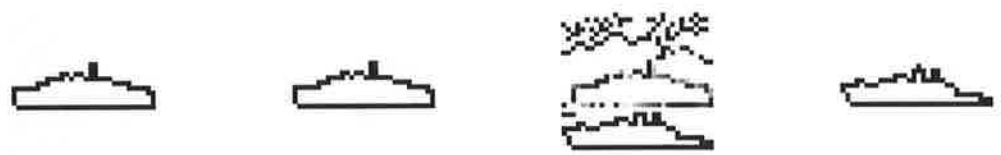
For a novel input, Field F2 will therefore be inactive for a period of time. The activity of Field F1 will therefore match that of F0 and hence lead to memory modification in the top-down synapses of Field F2. Neuron $N2$ (which measures the mismatch between the spatial pattern across Fields F1 and F2) will therefore be initially highly active, thus suppressing the activity of neuron $M2$. High activity of neuron $N2$ (or low activity of neuron $M2$) indicates that whatever is across Field F1 (i.e., the input pattern)

is not stored in memory. That is, the input stimulus if found to be novel. The net effect of this is that the transmitter mobilization rate in the bottom-up synapses from Field F0 to Field F1 is increased, thus increasing the overall cellular activity of F1.

6.3.1.1 Simulation of a Self-regulated Neural Circuit

Below we provide six computer simulations of the proposed self-regulated attentional neural circuit. The primary purpose of these simulations is to demonstrate the dynamics of self-regulation and its effect on memory and shape recognition in both a simple and a cluttered bottom-up input.

In the first two simulations we compare the cellular activities and the synaptic memory for two cases: (i) without regulation; and (ii) with self-regulation. These two simulations run for 100 iterations of the circuit equations. In the next two simulations, which run for 300 iterations, we repeat the first two simulations but this time we remove the input after the first fifty iterations and follow the circuit and memory dynamics. In the fifth simulation, which runs for 300 iterations, we first present a 2-D spatial input to the circuit and allow this to be stored in the top-down memory for the first 50 iterations. We then remove the input for 50 iterations and then represent it, but in a cluttered background for another 150 iterations. This is then followed by the input offset for another 50 iterations. In the sixth and the final simulation, we repeat the fifth simulation, but on the second presentation of the input we introduce a new 2-D shape. The results for these six simulations are graphed in Figures 6.16-6.21. These graphs are obtained from one cellular position in each of the neural Fields and at the location where the first 2-D shape has an active input. The six simulations use the following 2-D inputs:



2-D input for Simulations 1-4. | First input used in Simulation 5 and 6. | Second input used in Simulation 5. | Second input used in Simulation 6.

FIGURE 6.15. The two dimensional inputs used in the simulation of a self-regulated attentional neural circuit.

The simulation data for each of the simulations is presented graphically for one cellular pathway. The following is a list of the variables that are plotted in each of the graphs: (a) cellular activity of one cell in each of the neural Fields; (b) the dynamics of the top-down synaptic memory variables into Field F2; (c) the synaptic dynamics of the bottom-up input synapses into Field F0; (d) the activity of the match/mismatch

neuron pair (M1,N1); and (e) the activity of the match/mismatch neuron pair (M2,N2). Since the operation of the circuit has been described above, below we will only briefly compare the results for each pair of simulations.

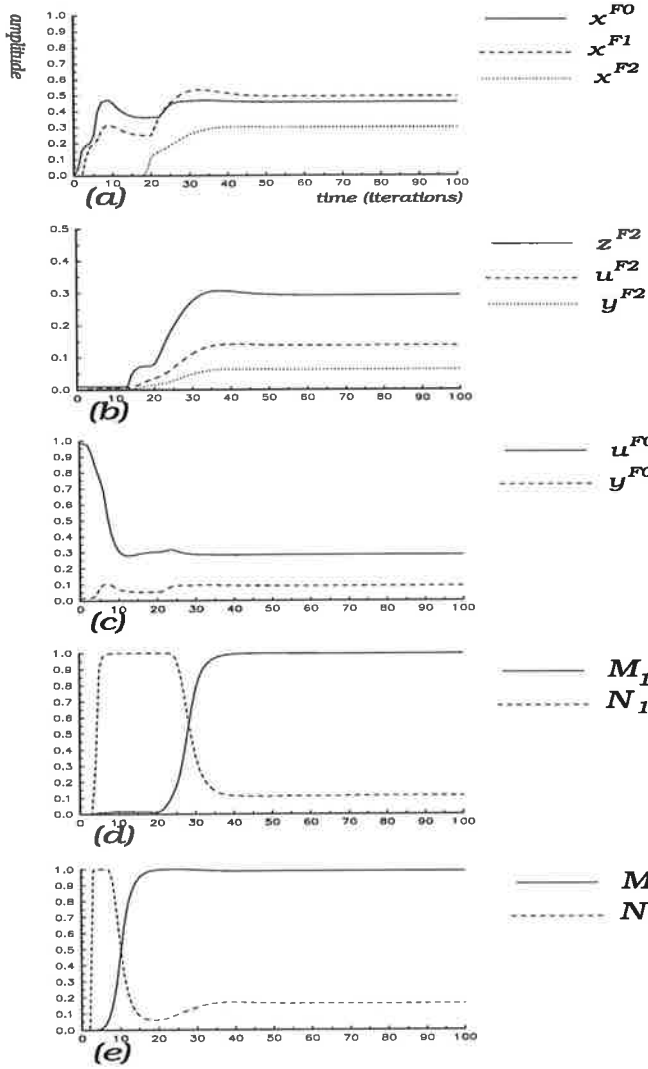


FIGURE 6.16. Cellular and synaptic memory dynamics during learning in an unregulated attentional neural circuit.

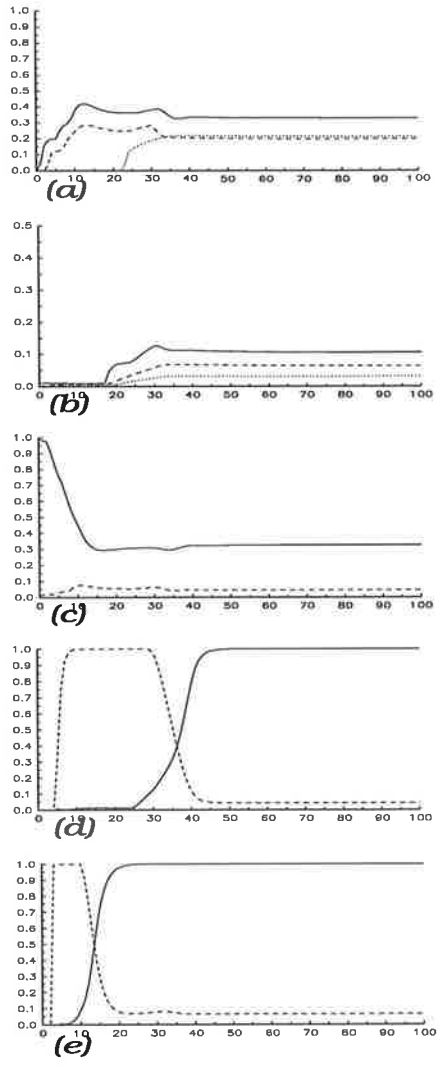


FIGURE 6.17. Cellular and synaptic memory dynamics during learning in a self-regulated attentional neural circuit.

Comparison of the above two simulation results shows that the cellular activity of the circuit and the top-down memory strength for the self-regulated circuit is lower. The data shows that as soon as the top-down memory is strong enough and begins to activate Field F2, it causes a rise in the activity of the match neuron M1 (which becomes fully activated at t = 40 iterations). Hence the rate of learning is also reduced (unlike the case for the unregulated circuit, where the memory converges to a much larger magnitude).

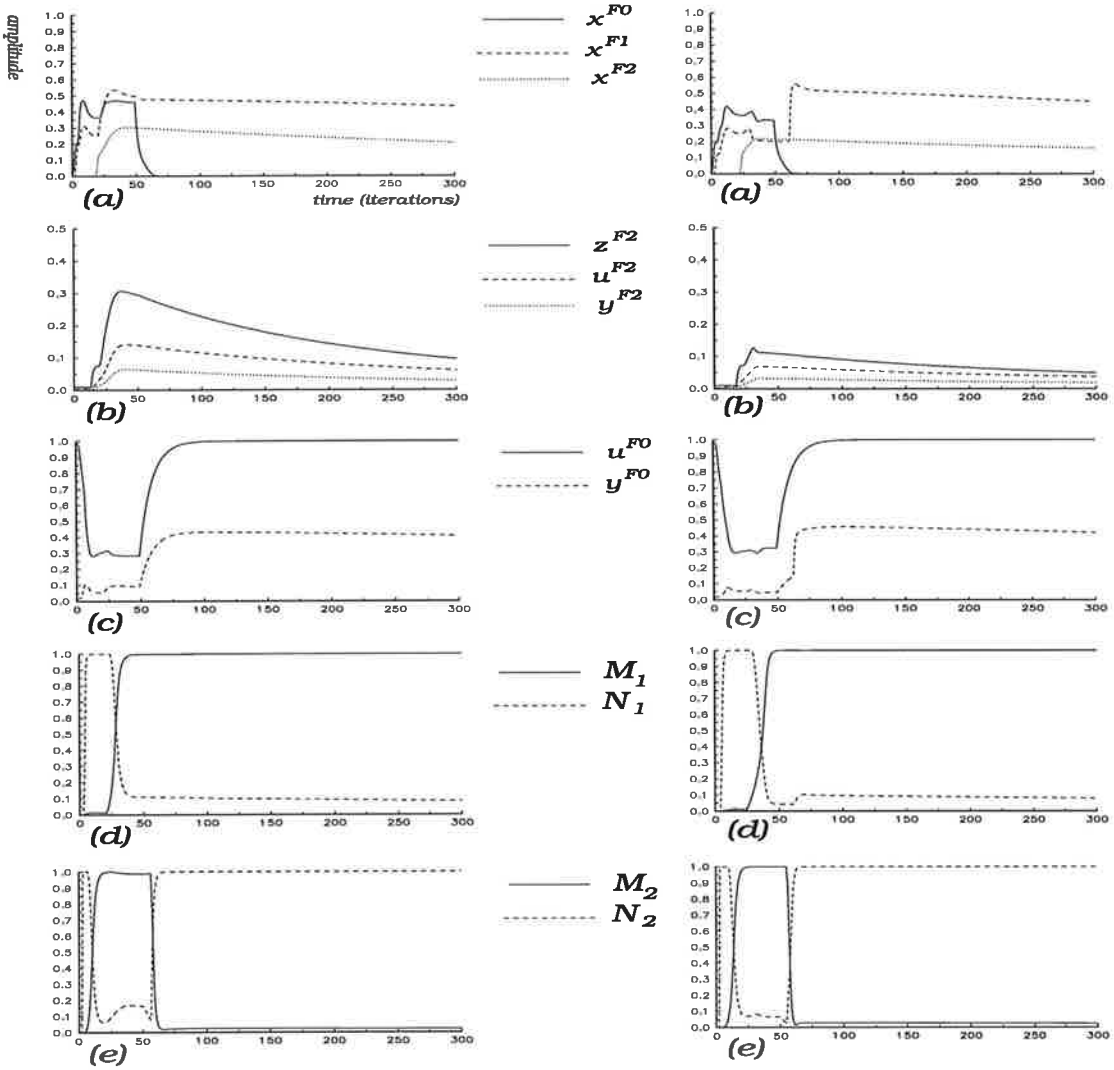


FIGURE 6.18. Cellular and synaptic memory dynamics during learning and memory retention in an unregulated attentional neural circuit.

FIGURE 6.19. Cellular and synaptic memory dynamics during learning and memory retention in a self-regulated neural circuit.

The above simulation data shows that when the circuit is self-regulated, the activity of Field F1 rises as soon as the input stimulus is removed (i.e., the circuit is receiving a strong top-down memory of the learned pattern from Field F2). Since Field F1 is very active and because the input is offset, the transmitter mobilization at the input to Field F0 rises, thus priming Field F0.

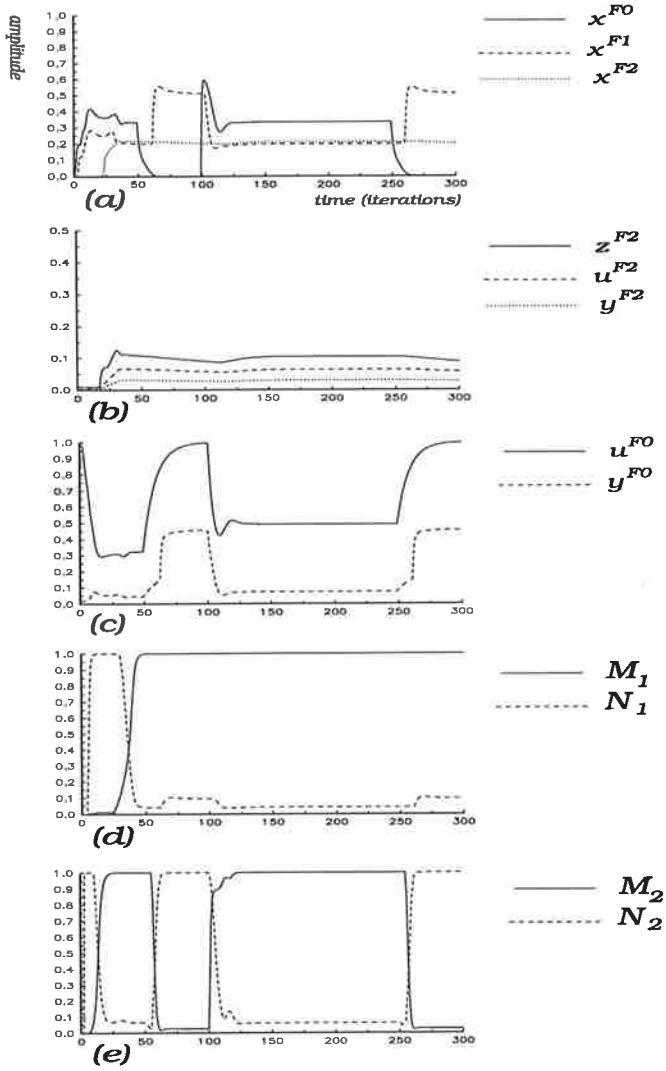


FIGURE 6.20. Cellular and memory dynamics in a self-regulated attentional neural circuit during learning, memory retention and 2-D shape recognition in cluttered background.

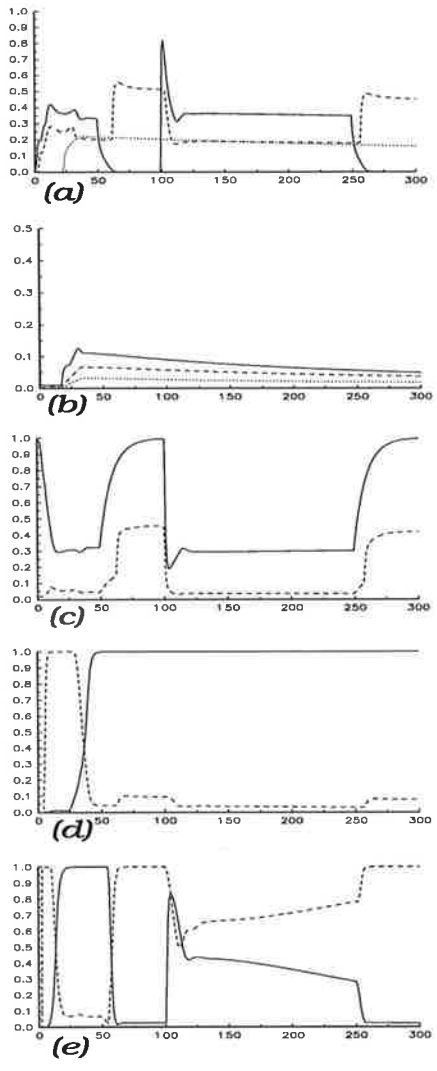


FIGURE 6.21. Cellular and memory dynamics in a self-regulated attentional neural circuit during learning, memory retention and mismatch with a new input shape.

Comparison of the above two simulation results shows that in the first case, Figure 6.20, the reverberatory memory has enabled the facilitation and hence the recognition of the shape when it is embedded in the cluttered background (as indicated by the high activity of neuron M2 at t = 100 iterations). In the second example (Figure 6.21), the activity of the same neuron drops as soon as the new shape is presented to the circuit (which implies that the new input does not match the reverberatory memory). The two dimensional representations of the above two simulations are shown in Figures 6.22 and 6.23.

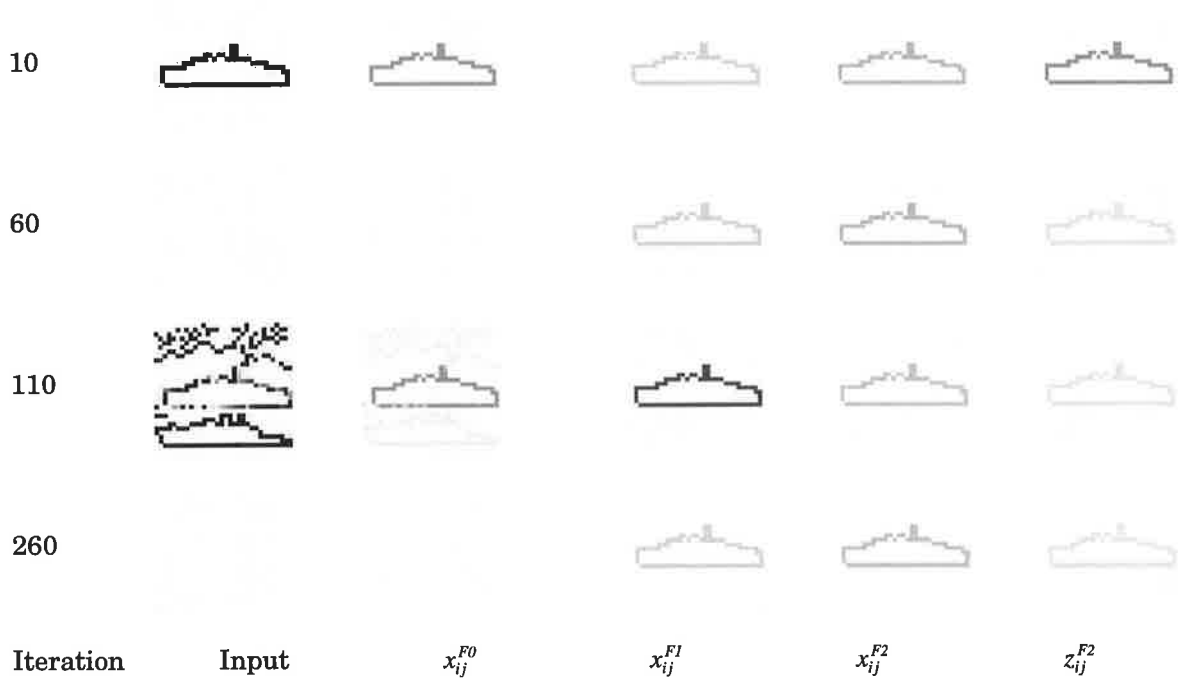


FIGURE 6.22. Two dimensional simulation results of a self-regulated attentional neural circuit during learning and pattern recognition in a cluttered input.

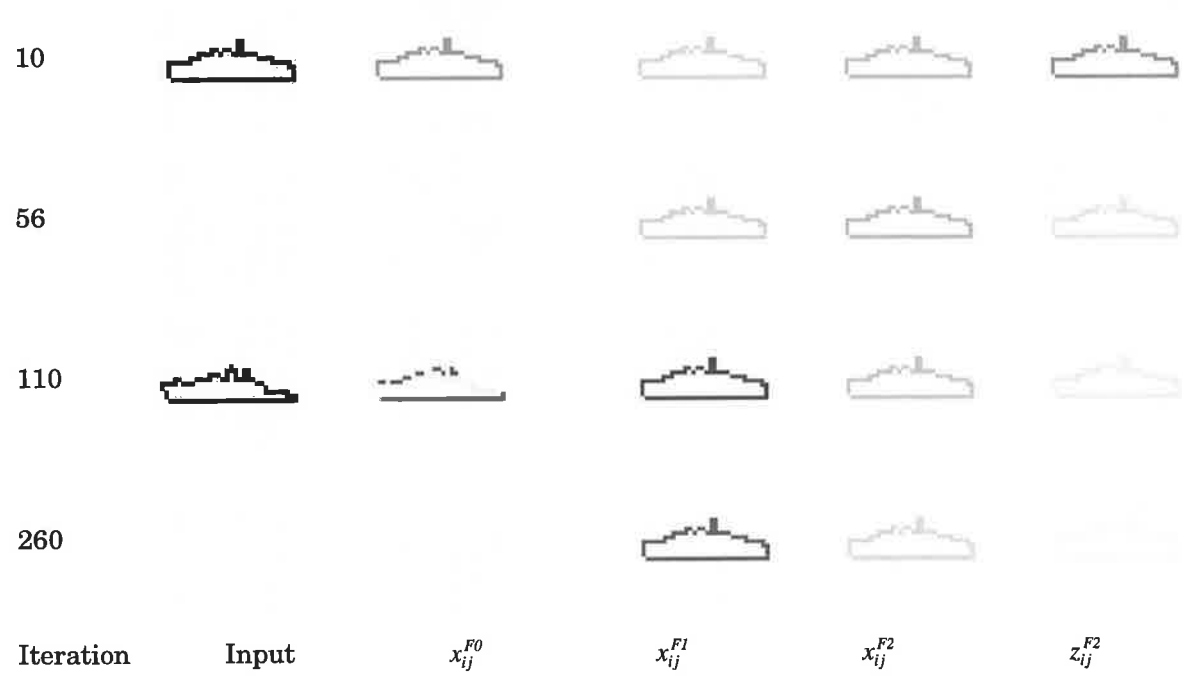


FIGURE 6.23. Two dimensional simulation results of a self-regulated attentional neural circuit during learning and mismatch with a new input.

6.4 Spatial Attention, Translational Invariance and Memory Guided Search

During the early phases of our research on biological vision and selective attention, we have developed a real-time competitive feedforward neural network model that is capable of translation invariant 2-D pattern recognition in simple input environments (Lozo *et al.*, 1994, 1995). The model, shown in Figure 6.24, uses ART-3 neural network of Carpenter and Grossberg (1990) for pattern recognition and an attention shifting module (Lozo *et al.*, 1993a, 1993b) whose window of attention shifts across the input space and selectively transfers the attended patterns to ART-3.

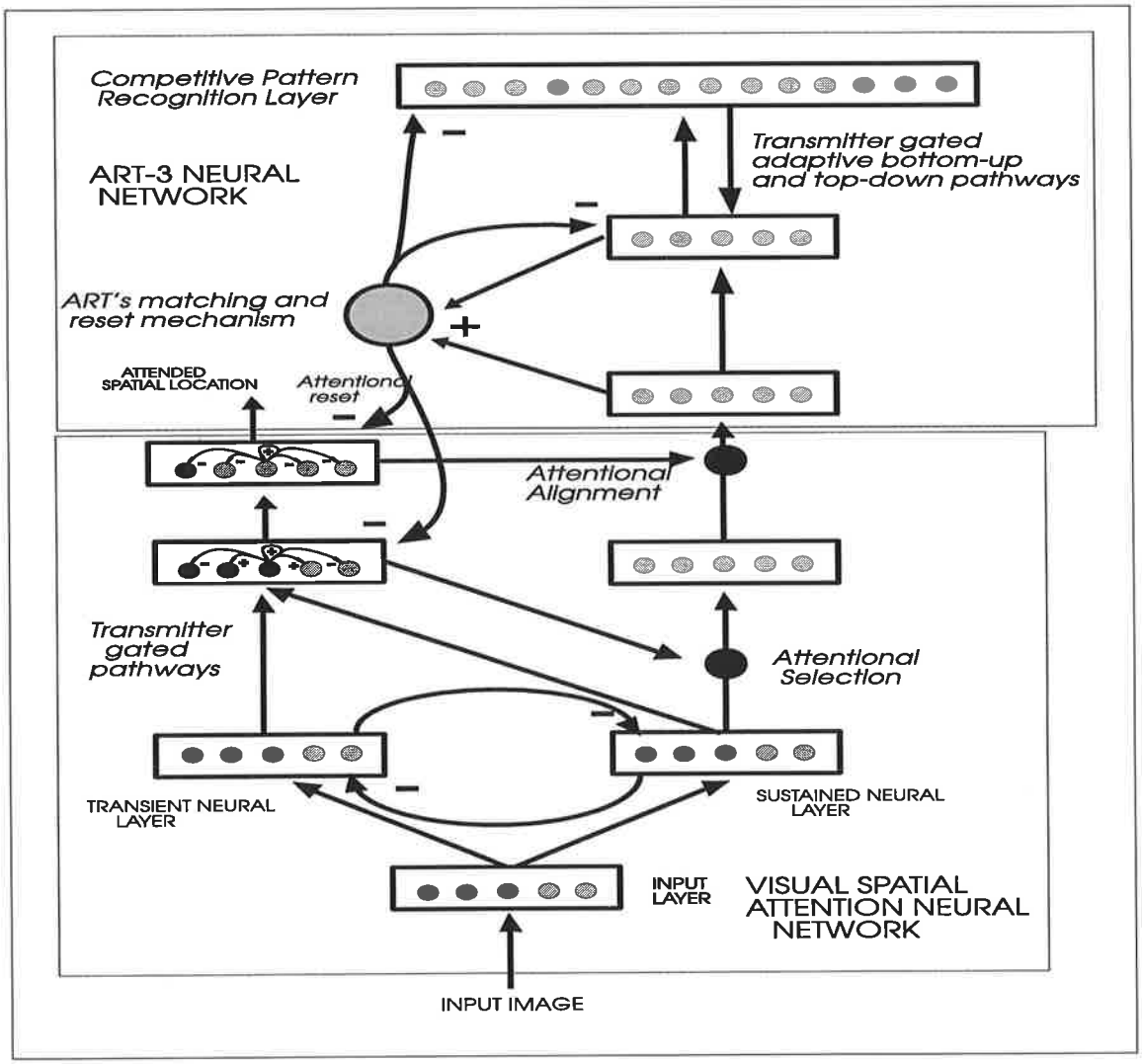


FIGURE 6.24. Simple feedforward real-time neural network model of bottom-up attentional selection and translation invariant 2-D pattern recognition (Lozo *et al.*, 1995).

The attention shifting module is based on the psychophysical data of the target cuing experiments (as reviewed in Chapter 2) and assumes that there is a neural structure in biological visual systems whose bottom-up and the top-down memory pathways are independent of external spatial locations but are referenced to an internal frame of reference. Our attention shifting module and the mechanism for achieving translation invariant representation is thus similar in principle to the dynamic routing circuit that was previously proposed by Anderson and Van Essen (1987) and subsequently implemented in a pattern recognition network by Olshausen *et al.* (1992, 1993), Van Essen *et al.* (1994). The main difference between the model by the above mentioned authors and that described here is that we have used a simple model of chemical synapses (described in Chapter 3) as a neural mechanism for biasing previously attended inputs against continually winning the competition for attention. In addition we have modelled the input transients. Apart from these small variations and the fact that we have used ART-3 neural network (rather than the Hopfield network), the method of achieving the translation invariant representation of the selected bottom-up input is identical. That is, before an object can be recognized, its centroid needs to be spatially aligned with the central axis of the pattern recognition system. The mechanism that achieves this spatial alignment is called 'visual spatial attention'.

The visual spatial attention neural network is driven by transient and sustained neural layers, the output of which feeds into two competitive layers that align the centroid of the attended bottom-up 2-D shape with the centre of ART-3's input layer. The first of the two attentional layers was modelled by Grossberg's cooperative-competitive feedback equation, whose lateral cooperation was selected to be the size of the attentional window (7x7 cells). The second attentional layer was a winner-take-all layer that gated the signals in the attentional window so as to form translational invariant input to ART-3's input layer. The central group of attentional neurons have an internal bias (i.e., "the fixation centre") and the lateral competitive interactions are distance dependent. Since the above model will be generalised in the next section (and in Chapter 8), below we briefly provide some simulation results that were obtained with the model.

Figure 6.25 shows the simulation results for two images that contain the same shapes but at different locations. The network is first exposed to the image shown in the left column. The activity of the attentional layer is shown on the left of each pair of columns while the output of the translation invariant layer is shown on the right of each pair of columns. These data are taken at different times (time increasing down the column). The data in the left pair of columns shows that the network was first attracted to character M (because it contains more information than character L) whose translation invariant representation is learned when the system reached a resonant steady after 29 iterations of the network. The attentional layer is then reset and attention shifts to character L which is learned at iteration 56.

The data in the right column was obtained when the second image was presented to the network. As before, the network was first attracted to the character M, but took only 21 iterations to reach a resonant state (because the shapes have been previously learned during the exposure to the first image and ART-3's recognition nodes become activated sooner). Character M was recognised after 21 iterations and character L after 39 iterations.

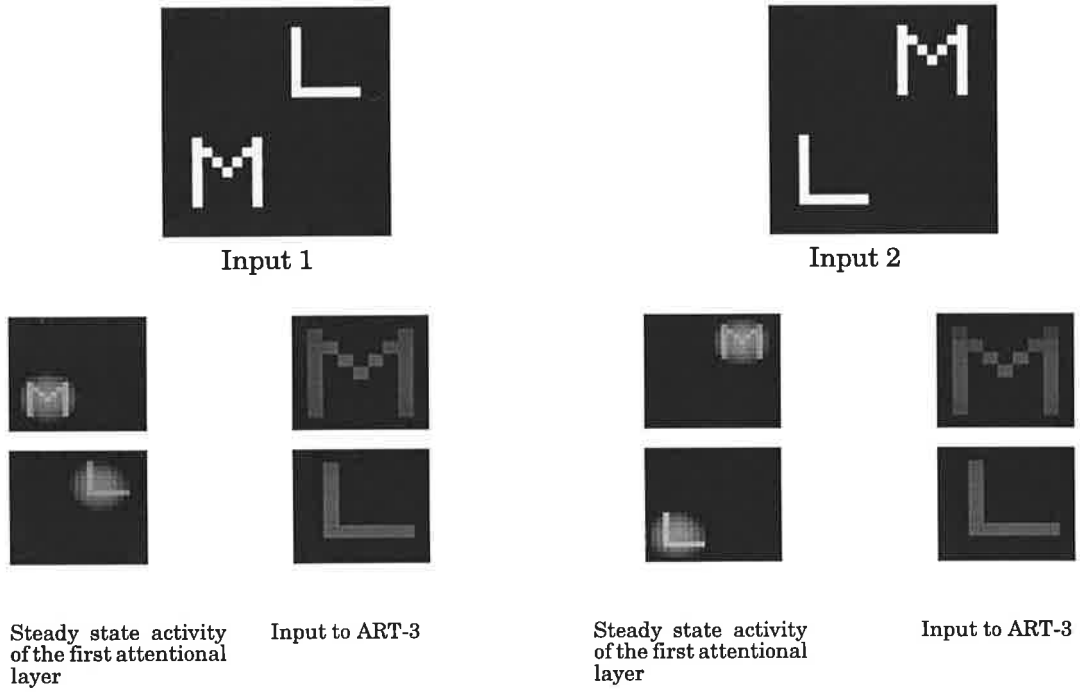


FIGURE 6.25. Translation invariant recognition of 2-D patterns in a simple visual input.

Pre-cueing Target Locations

To fully appreciate how the above network behaves when pre-cued to a valid (or non-valid) visual location of an impeding target stimulus, it was required to run the simulations over a large number of different input conditions. For example: precueing target locations at different distances from the fixation point, varying the validity of the cue, and repeating this for different cue-to-target inter-stimulus interval (stimulus onset asynchrony or SOA), and doing so at a number of contrast levels of both the cue and the target stimulus. In order to reduce this to a computationally manageable set of simulations that still reveal the important characteristics of the network, we have restricted the simulations to a single contrast level of the target stimulus (and the cue) which was presented at varying distances from the fixation centre and have repeated the simulation at each location a number of times and each time incrementing the cue-to-target time interval.

Note that the cued position was valid for each simulation (i.e., the target stimulus was always presented at the cued location). Since even this restricted set of simulations is still formidable when the 2-D version of the network is simulated on a 486 PC, the results shown below were obtained with the 1-D version of the network. The target location was cued by presenting the network with a constant amplitude signal whose spatial extend equalled in size to the impeding 1-D target stimulus and then removing the cue at the onset of the target stimulus.

Figure 6.26 shows the results of the cueing experiment. These simulation results resemble the psychophysical data from target pre-cueing experiments. Firstly, the results support the psychophysical data that shows that the benefit of target pre-cueing is a function of distance of the cued target from the fixation point. Secondly, our results also agree with the observations that the benefit of target pre-cueing on the recognition time increases with longer SOA's and that the greatest increase in the benefit is at short SOA's. The simulation data also shows that at long SOA's the benefit to reaction times begins to decrease (i.e., the recognition time begins to increase). This is attributed to the fact that at long SOA's the cue itself has had long time to sufficiently deplete the transmitter levels of the cued pathways so that by the time the target stimulus is presented, its transmitter gated signal is much weaker and takes longer time to surpass the level of inhibition from the fixational neurons. The increased reaction time at longer SOA's is actually observed with human subjects (Shulman *et al.*, 1979).

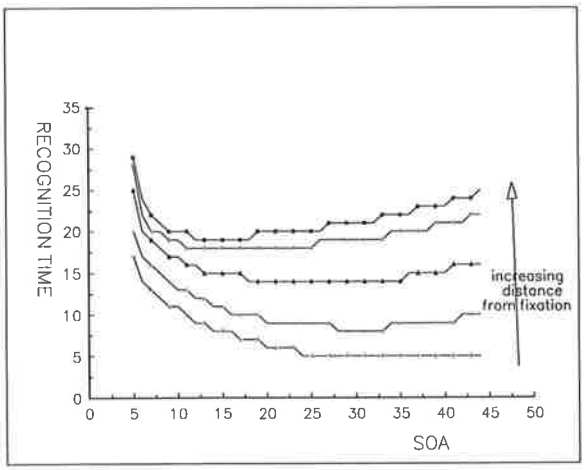


FIGURE 6.26. Simulation results of target pre-cueing at 5 different locations.

Note that the recognition time is represented by the number of iterations and is measured from the onset of the target-stimulus. Each successive SOA is incremented by presenting the target stimulus at successively longer delays from the onset of the cue (data shown is for 40 different SOA's).

The other significant result that can be noted from the above graphs is that the regression in the reaction time is faster and occurs sooner for far locations, that is, further the cued location is from the fixation centre sooner will it begin to feel the rise in reaction time. This is attributed to the fact that the level of fixational inhibition that needs to be surpassed by the transmitter gated post-synaptic potential increases as a function of distance from the fixation centre.

While being able to learn and/or recognize the bottom-up 2-D neural pattern of activity regardless of the pattern's spatial location in the input array and reproduce and explain the psychophysical results on target cueing experiments (Lozo *et al.*, 1995), the above model does not succeed when the desired 2-D pattern is surrounded by non-relevant neural activity in close proximity or when there is a spatial overlap between the input patterns. Thus even when a familiar pattern is at the input but is translationally displaced as well as being embedded in a complex background activity, the recalled memory in the above network cannot resonate with the input. In the next section we will consider how the mechanisms of presynaptic facilitation and top-down memory guided selective attention may be used to model an advanced neural model of attention selection where top-down attentional processes, mediated by the neural mechanisms of top-down presynaptic facilitation, may interact with and even override the bottom-up processes.

6.4.1 Advanced Neural Model of Visual Spatial Attention

In the primate visual system the selection of a desired input stimulus and the transfer of its neural representation to higher brain structures can be influenced by a number of factors. Some of these factors may depend on the characteristics of the stimulus (such as its colour or contrast; its spatial location, size or orientation; relative motion; continuity of its boundary or symmetry, etc). Although attentional selection can often proceed in a feedforward manner, in general it is a dynamic process that couples the bottom-up (pre-attentive) processes with the top-down attentive processes that may be under the influence of memory. For example, the memory of the colour of a person's clothing may often enables us to quickly spot that person in a crowd of people.

In Chapter 4 we have exposed only one top-down attentional mechanism (top-down presynaptic facilitation) that allows top-down memory to cause resonance in cases where the familiar input, although embedded in a cluttered bottom-up neural activity, was aligned with its top-down reference (top-down memory) pathways. However, the problem arises when there is a spatial misalignment between the memory pathways and the input pattern. On one hand, the misalignment may be total, such that the memory of the pattern cannot be activated, while on the other hand, a partial misalignment will lead to recognition failure.

As mentioned in the previous section, the visual spatial attention module of Figure 6.24 was developed during the early phases of this research (Lozo *et al.*, 1993a, b) and, although implemented with competitive neural layers, was not based on the presynaptically modulated shunting competitive neural layers. Because the model does not use top-down feedback it fails in cluttered inputs, particularly when inputs are closely spaced. We now propose design principles for a more advanced neural model of attentional selection and translation invariant 2-D pattern recognition. The proposed model couples the bottom-up attentional neural mechanisms with the top-down facilitatory presynaptic feedback pathways to enable the overall system to succeed in cases where the model shown in Figure 6.24 fails.

The basic concept of feedforward visual spatial attention and translation invariant representation of 2-D patterns can now be summarised in the context of presynaptically modulated shunting competitive neural layers, as shown in Figure 6.27.

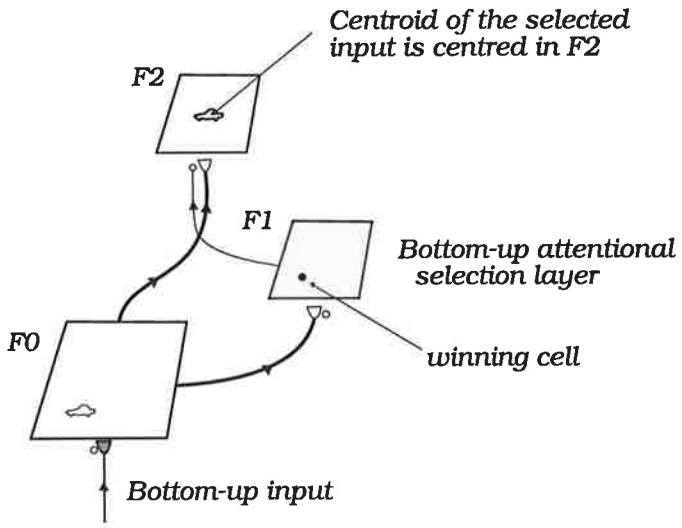


FIGURE 6.27. Basic neural mechanism of bottom-up visual spatial attention and translation invariant representation in the formalism of presynaptically modulated shunting competitive neural layers.

Field F0 receives its inputs from layers whose activity represents scenic edges and object boundaries. Field F1 (which may be modelled by a winner-take-all PFE-SCNL) is a 2-D bottom-up attentional selection layer that presynaptically facilitates signal transmission from Field F0 to Field F2. Each cell in Field F1 at location (m,n) samples its bottom-up inputs from a local cluster of MxM cells in F0 whose centroid is coincident with (m,n). The same cluster of cells also project their signals to Field F2, such that their centroid is coincident with the central axis of F2. That is, a contiguous 2-D region of MxM cells in F0 is mapped onto a central group of MxM cells in Field F2.

The winner-take-all competition between the cells in F1 ensures that, in a steady state, only one cluster of MxM cells in F0 transfers its signals to F2. Figure 6.28 shows the receptive fields of cells in Field F1.

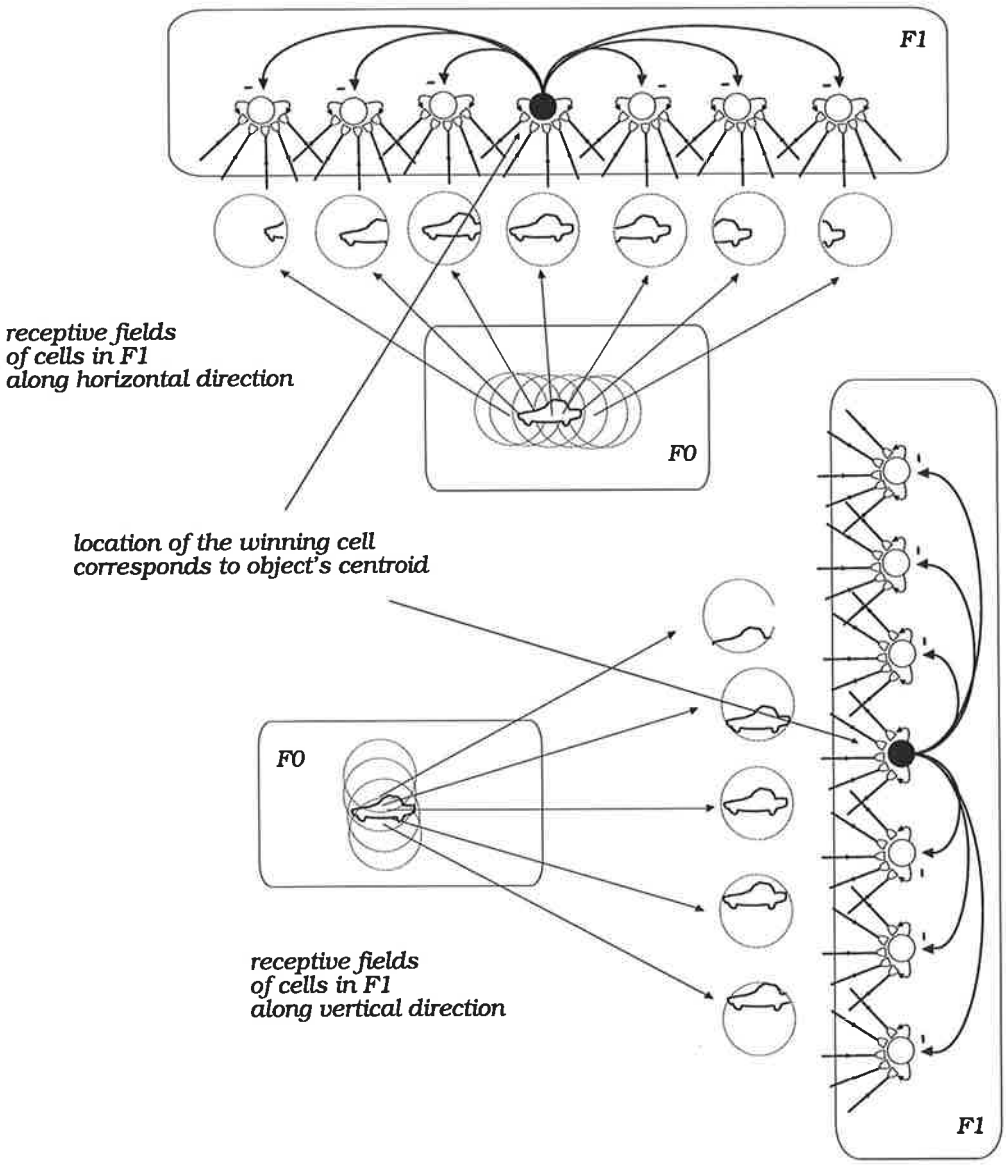


FIGURE 6.28. Input receptive fields of the bottom-up visual spatial attentional layer. For simplicity we have not shown the inhibitory interneurons which mediate the lateral competition.

Before proceeding on, let us pose for a moment and attempt to understand what is it that the bottom-up visual spatial attention layer is computing. Consider a single object being represented in the input array. If the boundary of this object, being represented by a spatial pattern of neural activity, appears alone across Field F0 such that all active cells that correspond to the object's boundary are equally active and, if its complete boundary fits within the receptive field of neurons in Field F1, then the

location of the winning cell in F1 corresponds to the centroid of the object. Hence, the location of the winning neuron in F1 can be used to align the centroid of the selected object with the centre of Field F2. This mechanism will thus ensure that the spatial pattern that is dynamically routed from Field F0 to Field F2 is independent of the spatial location of the input. Furthermore, since the receptive fields of cells in F1 are derived from a local sub-population of cells in F0, it can be said that local clusters of cells in F0 compete in F1. Alternatively, the bottom-up inputs compete for attentional selection via Field F1. It can now be easily envisaged how nearby neural activity (or highly varying strength along the objects boundary) in Field F0 can cause misalignments.

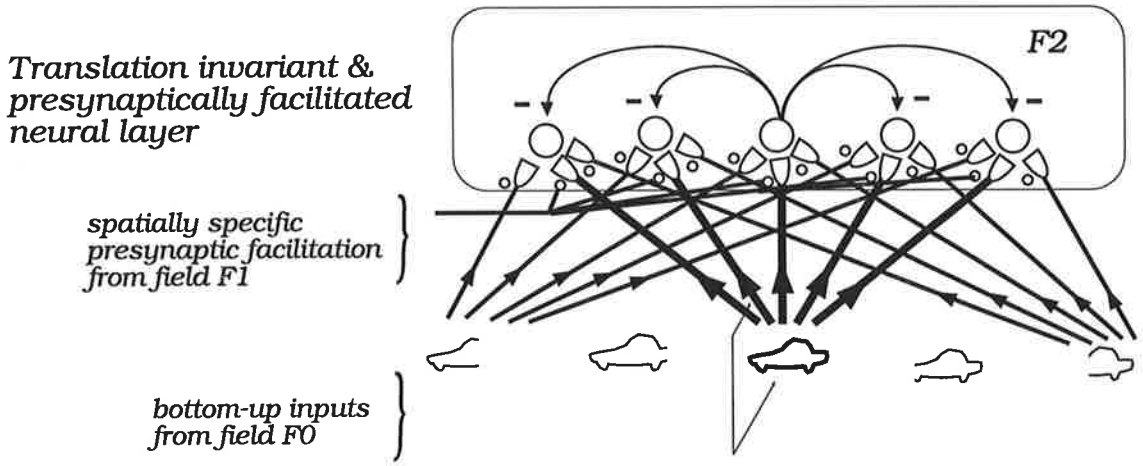


FIGURE 6.29. Bottom-up visual spatial attention layer presynaptically facilitates the transfer of neural signals from the input layer. Each cell of Field *F2* receives *MxM* input synapses from Field F0.

Since Field F1 facilitates the bottom-up synapses into Field F2, there is no reason to suppose why F2 should not send top-down facilitatory presynaptic feedback to facilitate the bottom-up synapses into Field F1. The feedforward-feedback interactions between the attentional selection layer and the translation invariant layer, shown in Figure 6.30, is of mutual presynaptic facilitation. That is, each neural layer or Field presynaptically facilitates the other. The attentional neural layer (Field F1) provides bottom-up spatially specific presynaptic facilitation of the translation invariant layer, while the latter provides pattern-specific top-down presynaptic facilitation of the former.

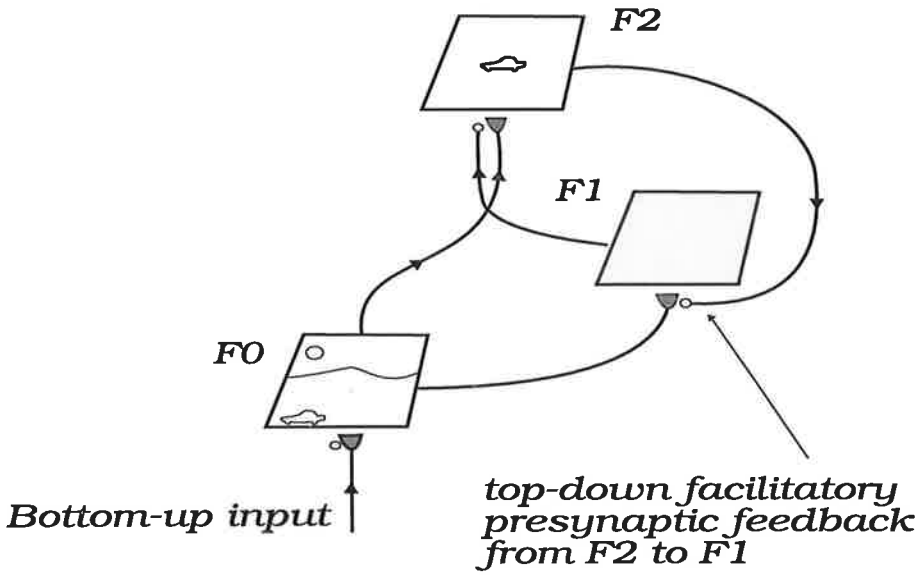


FIGURE 6.30. Schematic illustrating the bidirectional facilitatory interactions between the bottom-up attentional selection layer and the translation invariant layer.

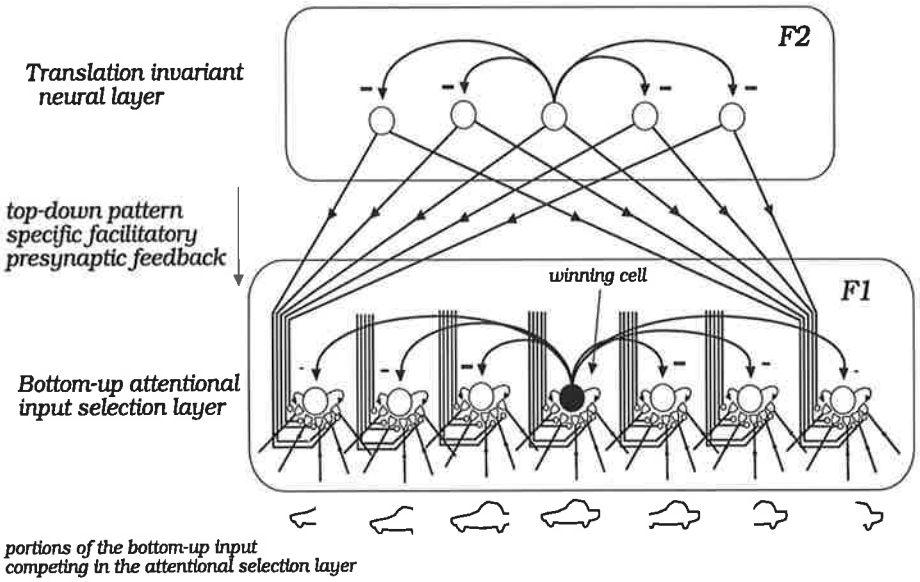


FIGURE 6.31. Pattern specific top-down presynaptic facilitation of the bottom-up attentional selection layer.

In the next section we propose how the top-down memory within the system may be used to influence the bottom-up attentional selection and thus effect the memory guided search process in cluttered bottom-up inputs.

6.4.2 Memory Guided Search in Cluttered Environments

Let us now suppose that Field *F2* projects excitatory inputs into another presynaptically modulated shunting competitive neural layer, Field F3, that also receives pattern-specific top-down memory, as shown in Figure 6.32.

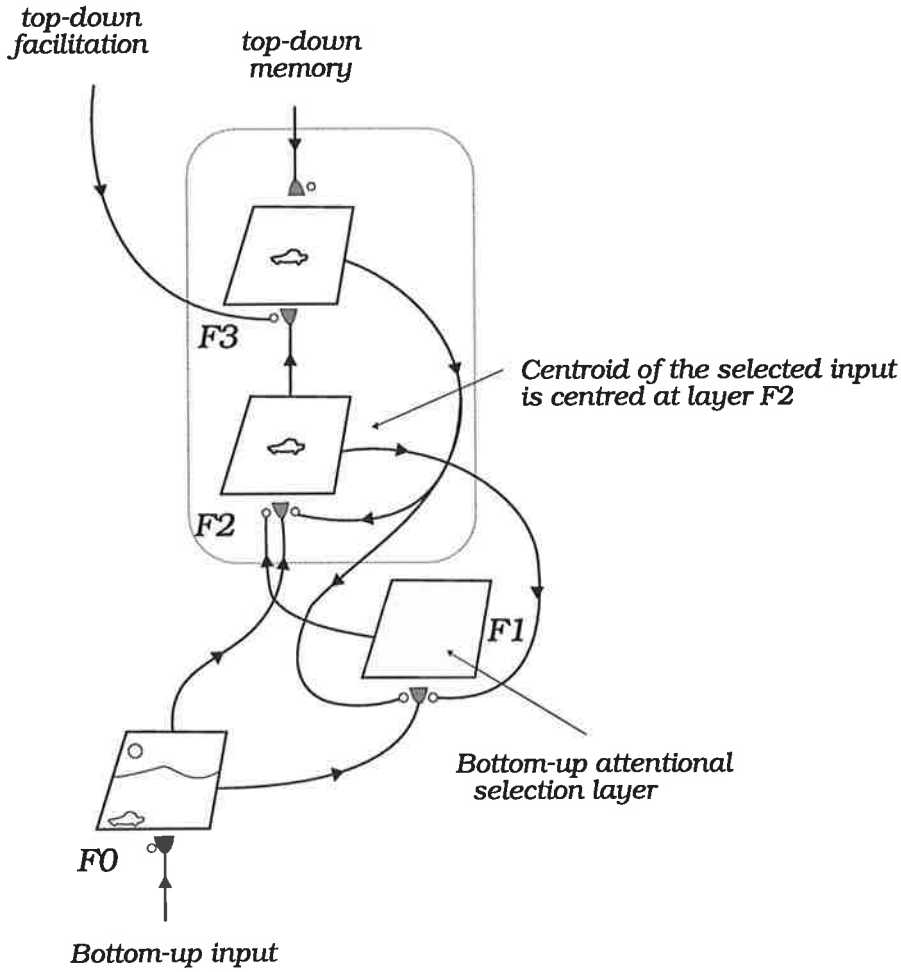


FIGURE 6.32. Influence of the pattern-specific top-down memory on the bottom-up attentional selection circuit.

If Field F3 is more sensitive to the top-down memory than it is to its bottom-up input, then the active memory across F3 can be used to influence the bottom-up attentional selection process by presynaptically facilitating the bottom-up attentional selection layer. This memory may thus drive the bottom-up attentional layer to search for a matching pattern in the input array and may also correct any misalignments that are initially due to the competing patterns across the bottom-up input Field *F0*. Note that the excitatory synaptic pathways $F2 \rightarrow F3$ may also be modulated, either by non-specific arousal mechanisms or by pattern (or spatially) specific facilitatory (inhibitory) presynaptic gain control signals.

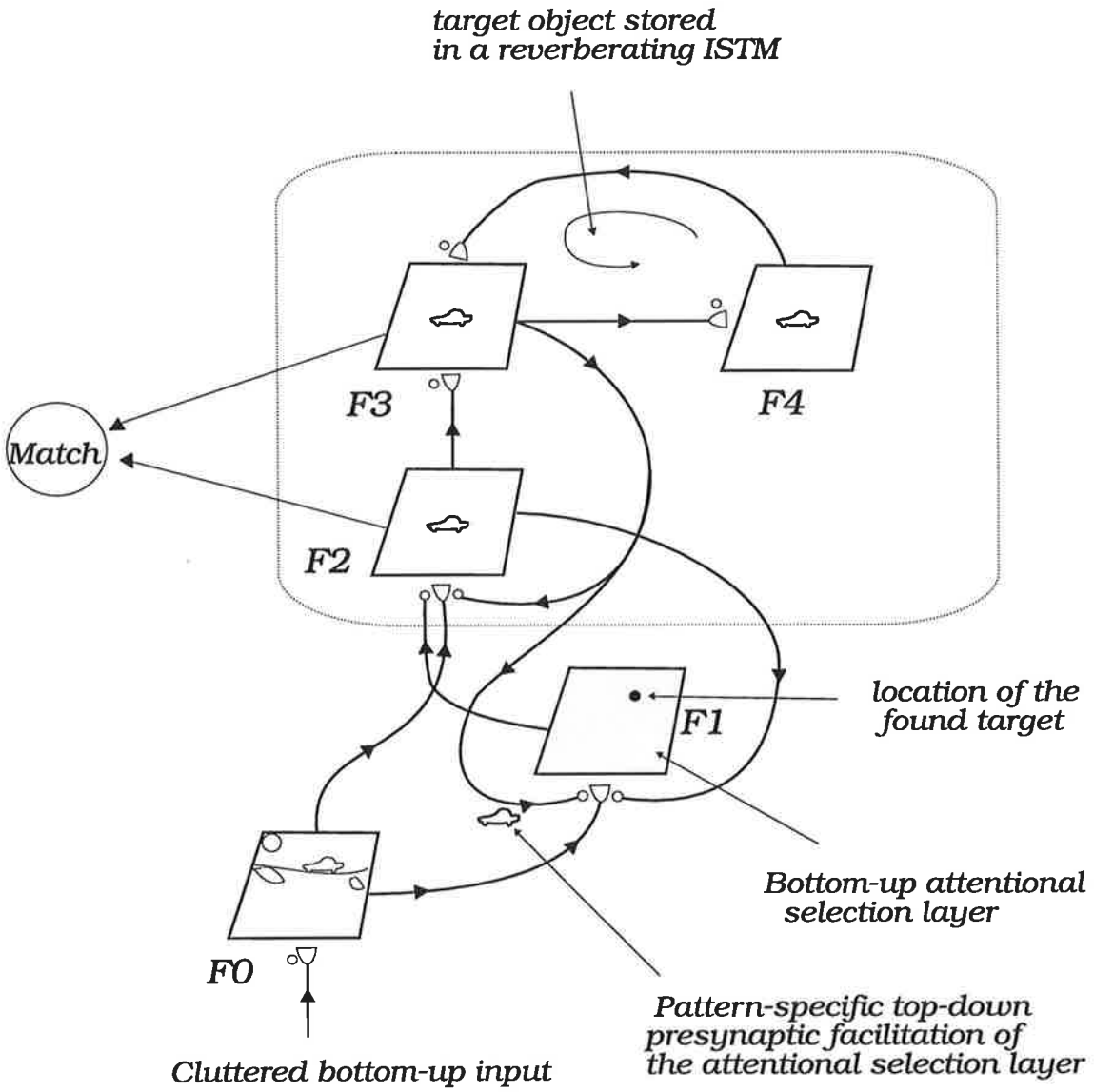


FIGURE 6.33. Schematic of a competitive neural circuit for memory guided search in a large and cluttered bottom-up input field. Note that the input Field F0 may in general have a much larger number of neurons than the translation invariant Field, F2. It is presently assumed that the receptive fields of the attentional selection neurons in Field F1 are all equal in spatial extent and that the target shape fits within these receptive fields.

To see how the immediate and pattern-specific short memory may participate in the top-down attentional selection and memory guided search in cluttered bottom-up inputs, we now use the reverberatory neural circuit of section 6.3 for immediate short term storage. With reference to Figure 6.33, let as assume that a certain 2-D shape is stored in the reverberatory neural circuit between Fields F3 and F4. Let us also assume that this reverberatory spatial pattern is centred (as shown) and that it facilitates each receptive field of the bottom-up attentional selection layer (Field F1).

If we now embed the original input shape (the car shape) in a cluttered background of other shapes, then only those cells in Field F1 whose input receptive fields are driven by the car shape will be activated. Hence, regardless of where the target shape is in the input array, the attentional selection layer, being guided by the top-down pattern-specific memory, will find it. When found, the target features are facilitated into Field F2. When the spatial patterns across Fields F2 and F3 are matched above the required threshold level, the circuit has found and recognized the target shape. The simple neural circuit described above is the minimum required to explain and simulate the recent neurobiological experiment on memory guided search in a monkey (Chelazzi *et al.,* 1993). However, we will provide the simulation of that experiment with a more advanced neural network that not only uses the above circuit but is also capable of self-organised real-time learning.

6.5 Conclusions

In this chapter we have provided several novel neural design principles and mechanisms for the design of self-regulated two dimensional attentional neural circuits that are capable of memory guided search. We have shown how the stored memories within a neural circuit may guide the bottom-up attentional processes to affect the bottom-up visual spatial attention and search for a desired 2-D pattern (or shape) in a much larger cluttered bottom-up input field. We have also provided a simple solution to how the activity of a neurobiological system may be decremented upon the presentation of familiar stimulus and how the neural activity is increased in the presence of novel stimuli. It is concluded that the top-down activated memory within a neurobiological system has a greater role than is currently appreciated. In the next chapter we propose a new self-organising neural network that can store a large number of memories and use them to facilitate the bottom-up signal transmission.

Chapter 7

SAART Neural Network

"How, then, can we decide when a network model is wrong? Such a task is deceptively simple. If the model's levels are incorrectly chosen, then it is wrong. If its interactions are incorrectly chosen, then it is wrong. And so on. The only escape from such a critique would be to demonstrate that a different set of levels and interactions can be correctly chosen, and shares similar functional properties with the original model. The new choice of levels and interactions would, however, constitute a new model. The old model would still be wrong."

S. Grossberg (1987)

7.1 Introduction and Overview

This chapter describes a novel real-time and self-organising neural network called Selective Attention Adaptive Resonance Theory (SAART). SAART extends the capability of ART neural networks of Carpenter and Grossberg (1987a, 1987b, 1990) to complex and cluttered visual sensory environments. The most significant property of the SAART network that is not shared by ART based models, such as the ART-3 neural network (Carpenter and Grossberg, 1990), is the capability of selective attention to familiar inputs. That is, the SAART neural network can recognize previously learned 2-D patterns and 2D shapes of 3D objects when they subsequently appear complete (and spatially aligned with the stored memory) in a cluttered bottom-up input. One of the new interactions in the network, *top-down presynaptic facilitation*, enables the SAART neural network to use its established memory to selectively filter the desired object shape from a cluttered input. This modulatory top-down selective gain control feedback, when combined with pre-postsynaptic interactions, also provides the network with invariance to spatial illumination effects that severely affects most other artificial neural networks.

7.2 SAART Neural Network Architecture

Figure 7.1 shows the architecture of the SAART neural network that is capable of real-time stable learning (and recognition of familiar neural events) in a complex and noisy bottom-up neural activity.

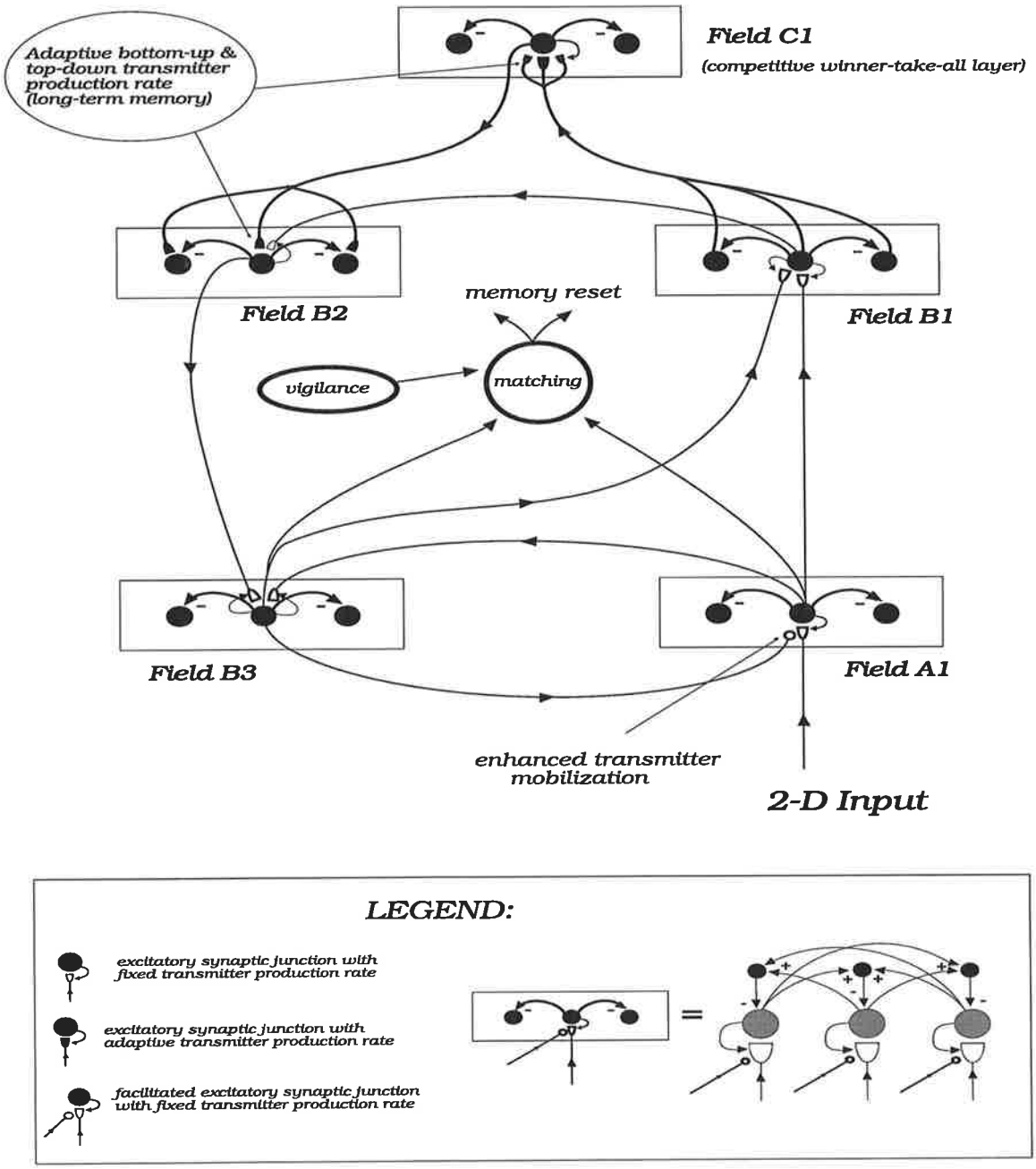


FIGURE 7.1. SAART neural network architecture. All neural Fields, with the exception of Field C1, are two dimensional. For simplicity we have shown only one set of synaptic pathways to and from memory cells in Field C1.

The network consists of five Presynaptically Facilitated Excitatory Shunting Competitive Neural Layers (PFE-SCNLs), or neural Fields ($A1, B1, B2, B3$, and $C1$).

SAART's processing Fields are interconnected via dynamic synaptic pathways whose internal dynamics represents the decaying memory of the most recent neural activity that entered the network via Field A1. Memories of past and regularly experienced neural events are evolved (learned) and retained in the adaptive bottom-up long-term memory (BU-LTM) pathways ($B1 \rightarrow C1$) and the adaptive top-down long-term memory (TD-LTM) pathways ($C1 \rightarrow B2$).

The top down pathways originating from cells in Field B3 enhance transmitter mobilization at input synapses of the corresponding cells in Field A1 (i.e., the interactions between these two neural Fields represent the FFE-FBPF neural circuit of Chapter 4). The bottom-up input is thus amplified by the recalled memory that appears across Field B3 (as well as Fields B1 and B2) when neurons in Field C1 have exceeded a "physiologically significant firing frequency" (at least one fifth of their maximum activation level). Field B2 is designed so that its cells are more sensitive to the top-down signals from Field C1 than to signals from Field B1 and also have a greater saturation level (by at least a factor of 10). This prevents the corruption of the long term memory (residing in the pathways $B1 \rightarrow C1$ and $C1 \rightarrow B2$) from intrusion by noise, clutter and non-matching inputs. Long term memory is updated whenever the 2-D spatial patterns of cellular activity across Fields A1 and B3 are matched to within a pre-set tolerance level (or vigilance) and when the network reaches a stable state (i.e., when $dR/dt \approx 0$ where R is the degree of match). Desensitization of postsynaptic receptors on cells of Field C1 provides a temporary bias against the activation of the same memory cell after the network is reset.

Since the bottom-up neural activity at the input to Field A1 is gated by the mobilized transmitter in the presynaptic terminal, only those cells whose synaptic gain (amount of neurotransmitter that is in the mobilized state) is elevated will be highly active. All other cells in Field A1, although they may receive strong bottom-up input, will have their activity suppressed by the competitive interactions that occur between the cells across the whole Field A1. Since Field B3 contains a sum of two spatial patterns or neural activities (the bottom-up activity from Field A1 and the top-down activity from Field B2) the resultant neural activity that ends resonating in the loop $B3 \rightarrow B1 \rightarrow B2 \rightarrow B3$ will depend not only on what has entered this loop from Field A1 but also on the signals entering this loop from Field C1 via Field B2. Very weak inputs to Field B2 from Field C1 (i.e., weak top-down memory) will not have any effect on the system, allowing the bottom-up input activity to enter and resonate in the loop, which then leads to learning of the new input stimulus. However, strong inputs from Field C1 will totally overwrite the contents of Field B2 (and hence Fields B1 and B3) that

was initially due to the input signals that are passed through Fields A1 and B1. This strong top-down activity will circulate (reverberate) in the $B3 \rightarrow B1 \rightarrow B2 \rightarrow B3$ loop while at the same time it will amplify the corresponding bottom-up neural signals that may be available at the input to Field A1. Whether this reverberation leads to stable resonance (equilibrium) and learning depends on whether the reverberating neural activity across Field B3 can select a portion of the bottom-up activity across Field A1 with which it can resonate. Comparison of the activities across Fields B3 and A1 indicates whether the two are in resonance or not. The long term memory, which resides in the synapses of adaptive pathways $B1 \rightarrow C1$ and $C1 \rightarrow B2$ is updated whenever the cellular activities (spatial patterns) across Fields B3 and A1 are matched to within a desired vigilance criterion and when the network is in a stable resonating state. If the reverberating activity cannot be matched to the selected portion of the input, the network will reset and then attempt to access another previously learned memory or will learn the total input activity across Field A1. The neural activity that may be reverberating in the $B3 \rightarrow B1 \rightarrow B2 \rightarrow B3$ loop decays very rapidly after the input offset. This fast decay, due to the low level of tonic transmitter mobilization in the synaptic pathway $B3 \rightarrow B1$, is in contrast to a more durable memory that is usually associated with short term memory.

Inspection of Figure 7.1 shows that ART models are embedded in the SAART model. Although the implementation may look different, in fact the network reverts to the ART model when several parts are removed (layer B3 and the top-down feedback from Field B3 to Field A1) and then integrating B1 and B2 into one layer and comparing its activity with that of Field A1. By further elimination of the top-down memory pathways and the matching (orienting or the vigilance) subsystem, we end up with a basic Instar model (Grossberg, 1972). In the process of eliminating the mentioned parts of the system, we also reduce the generality of the model. Although the SAART model is more general than ART, it is a long way from being a general neural model of cognitive vision. Its major deficiency is an inability to deal with the various spatial transformations of the sensory input. Thus, ART is a special case of a more general SAART.

7.2.1 Implementation of SAART's Processing Fields

All of SAART's processing Fields are implemented with the same basic set of five equations that we have proposed in Chapters 4 and 5 as a general model of presynaptically modulated shunting competitive neural layers. Thus almost all pre-post-synaptic interactions that were discussed in the previous chapters, as well as the parameter design procedure of Chapter 5 is valid for most of SAART's processing Fields. However, since Field C1 contains winner-take-all memory cells whose input and output pathways are modified during learning, the implementation of Field C1 requires extra equations. Although most of the network's processing Fields obey similar dynamics,

they vary in the way they receive and interact with their inputs. Hence we will describe the dynamics of pre-postsynaptic interactions for each layer in the network. For completeness, we provide all equations of the network.

7.2.1.1 SAART's Field A1

The following five equations define the dynamics of pre-postsynaptic interactions at each cellular position of Field A1, as schematized in Figure 7.2.

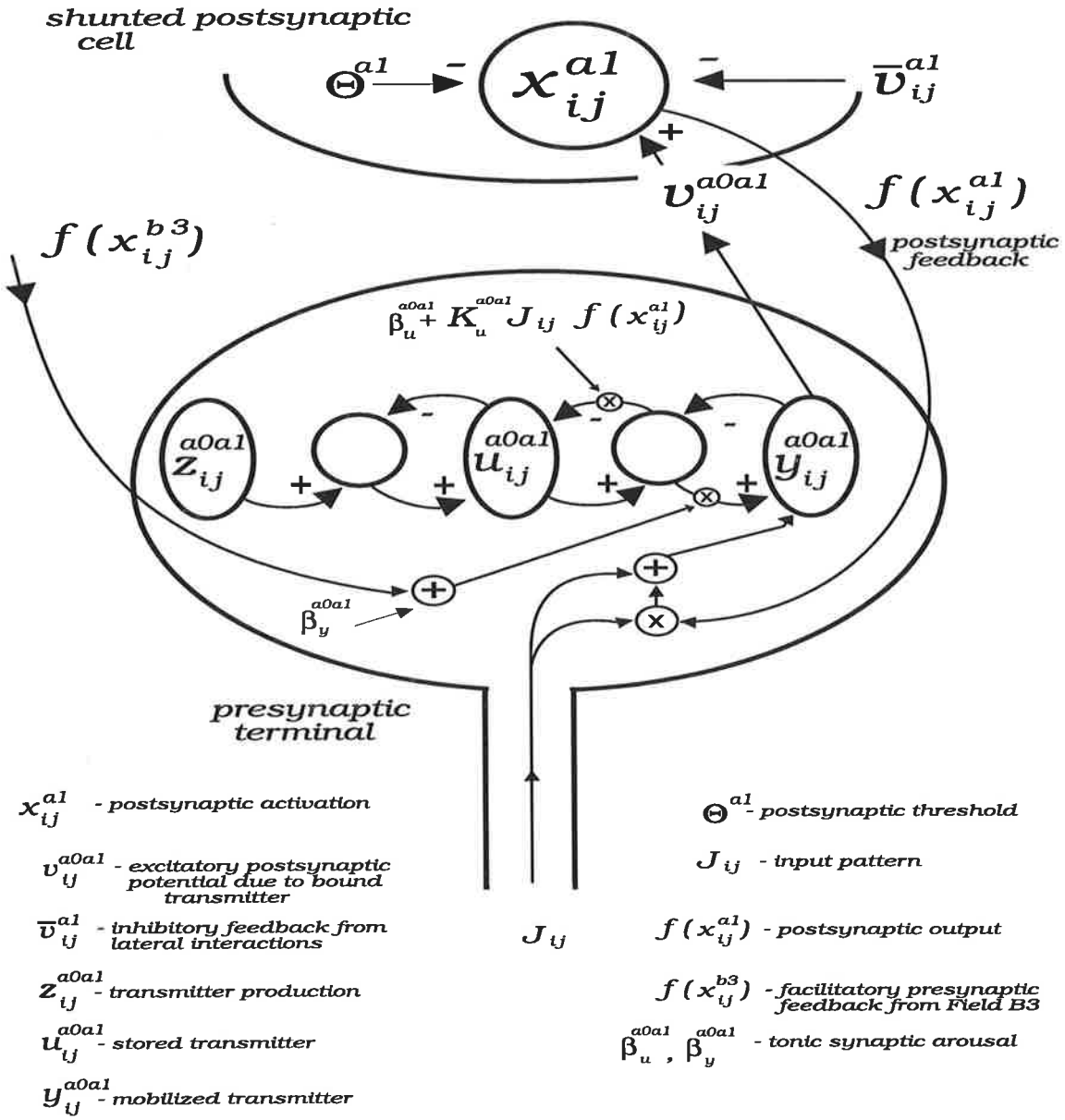


FIGURE 7.2. Pre-postsynaptic interactions in SAART's Field A1.

Postsynaptic activity at equilibrium (Field A1)

The equilibrium activity of each cell in Field A1 is given by

$$x_{ij}^{a1} = \frac{B^{a1} G^{a0a1} v_{ij}^{a0a1}}{A^{a1} + G^{a0a1} v_{ij}^{a0a1} + \overline{G}^{a1} \overline{v}_{ij}^{a1}} \tag{7.1}$$

where v_{ij}^{a0a1} is the excitatory postsynaptic potential; $\overline{v}_{ij}^{a1}$ is the lateral inhibitory postsynaptic potential; G^{a0a1} is the gain in the excitatory synaptic pathway, while $\overline{G}^{a1}$ is the gain of lateral inhibition.

Inhibitory postsynaptic potential (IPSP) due to lateral competition

The activity of the inhibitory layer is described by the following equation:

$$\frac{d\overline{v}_{ij}^{a1}}{dt} = -\overline{A}^{a1} \overline{v}_{ij}^{a1} + \overline{B}^{a1} \sum_{(k,m) \neq (i,j)} f(x_{km}^{a1}) \tag{7.2}$$

Excitatory postsynaptic potential due to the bound transmitter

The excitatory input is gated by the mobilized transmitter in the synaptic terminal to produce an EPSP that is given by (7.3).

$$\frac{dv_{ij}^{a0a1}}{dt} = -D^{a0a1} v_{ij}^{a0a1} + J_{ij}[y_{ij}^{a0a1} - Y^{a0a1}]^{+}[\rho_v^{a0a1} + K_v^{a0a1} f(x_{ij}^{a1})] \tag{7.3}$$

where ρ_v^{a0a1} and K_v^{a0a1} are constants which determine the relative contributions to the EPSP from the input signal alone and from the correlated pre-postsynaptic activation respectively. The first term in the above equation represent the passive decay of EPSP. The second term says that the total excitatory input is gated by the mobilized transmitter y_{ij}^{a0a1}. The transmitter level must be above a threshold Y^{a0a1} before it can be released to contribute to the EPSP. The bottom-up input into the synapse (which may be from a previous neural layer) is denoted by J_{ij}. The postsynaptic feedback ($f(x_{ij}^{a1})$) interacts with the bottom-up input to release the transmitter under the condition of correlated pre-postsynaptic firing (i.e., multiplicative pre-postsynaptic interaction). The released transmitter binds to the postsynaptic cell and is converted into an excitatory postsynaptic potential (EPSP), represented by v_{ij}^{a0a1}. Top-down feedback signal from Field B3 ($f(x_{ij}^{b3})$) primes the presynaptic terminal by enhancing (facilitating) transmitter mobilization. In the absence of bottom-up inputs, the priming signal can enhance transmitter mobilization but does not cause any release.

Stored transmitter

$$\frac{du_{ij}^{a0a1}}{dt} = \alpha_u^{a0a1}(z_{ij}^{a0a1} - u_{ij}^{a0a1}) - [\beta_u^{a0a1} + K_u^{a0a1} J_{ij} f(x_{ij}^{a1})](u_{ij}^{a0a1} - y_{ij}^{a0a1}) \tag{7.4}$$

The second term in equation (7.4) says that the stored transmitter is depleted by the correlated pre-postsynaptic activity.

Mobilized transmitter

$$\frac{dy_{ij}^{a0a1}}{dt} = -\gamma_y^{a0a1} y_{ij}^{a0a1} - J_{ij}[y_{ij}^{a0a1} - Y_{ij}^{a0a1}]^+ [\rho_y^{a0a1} + K_y^{a0a1} f(x_{ij}^{a1})]$$
$$+ [\beta_y^{a0a1} + H_y^{a0a1} f(x_{ij}^{b3})](u_{ij}^{a0a1} - y_{ij}^{a0a1}) \tag{7.5}$$

where H_y^{a0a1} is the gain of facilitation. This equation says that the top-down feedback signal $f(x_{ij}^{b3})$ from Field B3 facilitates transmitter mobilization in the input synapses of Field A1.

7.2.1.2 SAART's Field B1

Each cell in Field B1 receives synaptic inputs from Fields A1 and B3, as shown in Figure 7.3. The input from Field B3 is part of the feedback pathway that forms the reverberatory loop $B1 \rightarrow B2 \rightarrow B3 \rightarrow B1$. However, this reverberatory feedback does not maintain a permanent memory of the input once the input is offset. The reverberatory pattern in the loop thus decays when the input is offset, enabling new inputs to enter into the network. Since all the basic synaptic interactions were discussed previously (in section 4.3.2), below we will only write the equations that model Field B1.

Postsynaptic activation at equilibrium

$$x_{ij}^{b1} = \frac{B^{b1}(G^{a1b1} v_{ij}^{a1b1} + G^{b3b1} v_{ij}^{b3b1})}{A^{b1} + G^{a1b1} v_{ij}^{a1b1} + G^{b3b1} v_{ij}^{b3b1} + \overline{G}^{b1} \overline{v}_{ij}^{b1}} \tag{7.6}$$

Inhibitory postsynaptic potential due to lateral competition

$$\frac{d\overline{v}_{ij}^{b1}}{dt} = -\overline{A}^{b1} \overline{v}_{ij}^{b1} + \overline{B}^{b1} \sum_{(k,m) \neq (i,j)} f(x_{km}^{b1}) \tag{7.7}$$

Excitatory postsynaptic potential (pathway $A1 \rightarrow B1$)

$$\frac{dv_{ij}^{a1b1}}{dt} = -D^{a1b1} v_{ij}^{a1b1} + f(x_{ij}^{a1})[y_{ij}^{a1b1} - Y^{a1b1}]^{+}[\rho_v^{a1b1} + K_v^{a1b1} f(x_{ij}^{b1})] \tag{7.8}$$

Stored Transmitter (pathway $A1 \rightarrow B1$**)**

$$\frac{du_{ij}^{a1b1}}{dt} = \alpha_u^{a1b1}(z_{ij}^{a1b1} - u_{ij}^{a1b1}) - [\beta_u^{a1b1} + K_u^{a1b1} f(x_{ij}^{a1}) f(x_{ij}^{b1})](u_{ij}^{a1b1} - y_{ij}^{a1b1}) \tag{7.9}$$

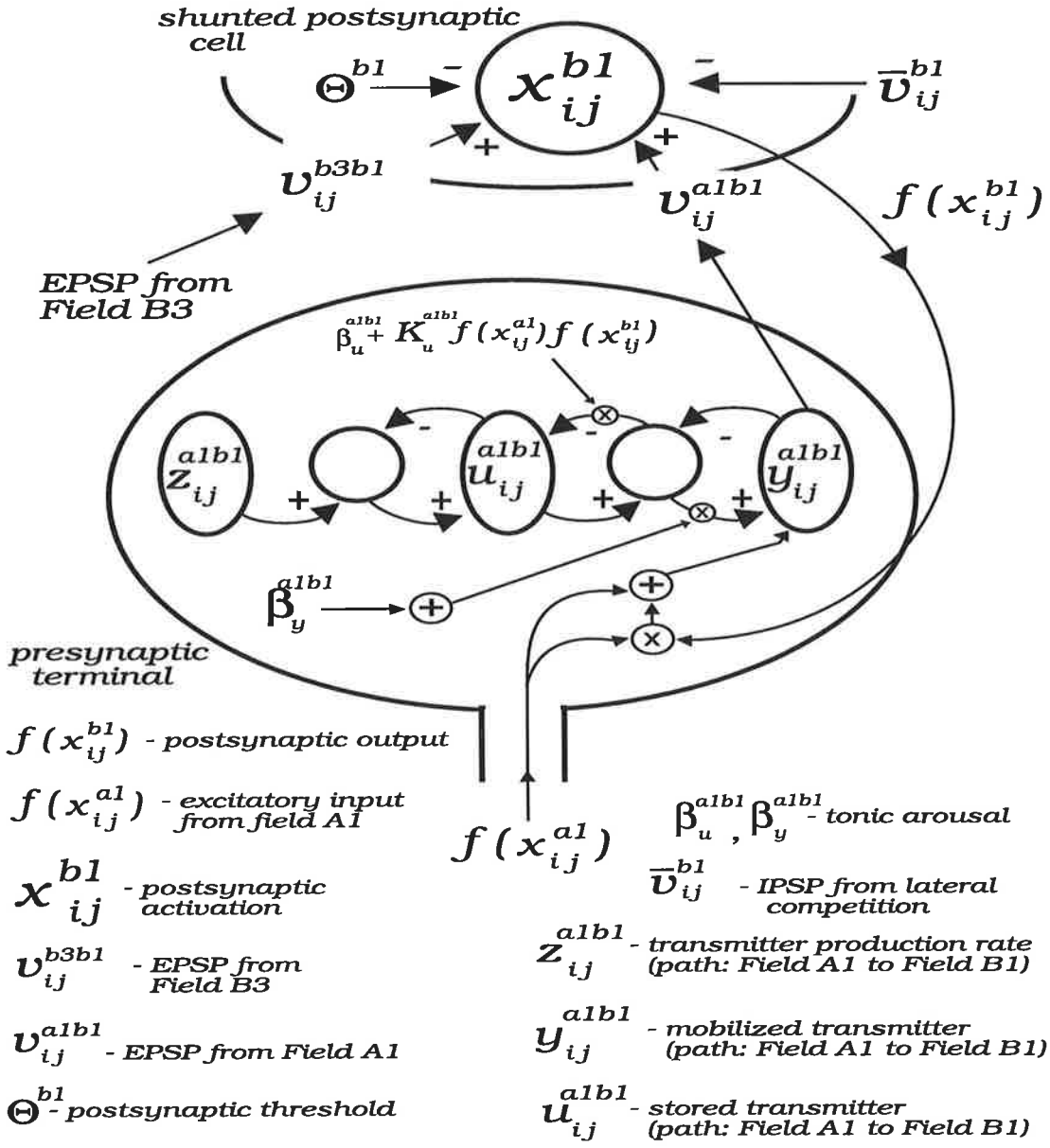


FIGURE 7.3. Pre-postsynaptic interactions in SAART's Field B1.

Mobilized Transmitter (pathway $A1 \rightarrow B1$)

$$\frac{dy_{ij}^{a1b1}}{dt} = -\gamma_y^{a1b1} y_{ij}^{a1b1} - f(x_{ij}^{a1})[y_{ij}^{a1b1} - Y^{a1b1}]^+ [\rho_y^{a1b1} + K_y^{a1b1} f(x_{ij}^{b1})]$$

$$+ \beta_y^{a1b1}(u_{ij}^{a1b1} - y_{ij}^{a1b1}) \tag{7.10}$$

The next three equations describe the synaptic pathways into Field B1 from Field B3.

Excitatory postsynaptic potential (pathway $B3 \rightarrow B1$)

$$\frac{dv_{ij}^{b3b1}}{dt} = -D^{b3b1} v_{ij}^{b3b1} + f(x_{ij}^{b3})[y_{ij}^{b3b1} - Y^{b3b1}]^+ [\rho_v^{b3b1} + K_v^{b3b1} f(x_{ij}^{b1})] \tag{7.11}$$

Stored Transmitter (pathway $B3 \rightarrow B1$)

$$\frac{du_{ij}^{b3b1}}{dt} = \alpha_u^{b3b1}(z_{ij}^{b3b1} - u_{ij}^{b3b1}) - [\beta_u^{b3b1} + K_u^{b3b1} f(x_{ij}^{b3}) f(x_{ij}^{b1})](u_{ij}^{b3b1} - y_{ij}^{b3b1}) \tag{7.12}$$

Mobilized Transmitter (pathway $B3 \rightarrow B1$)

$$\frac{dy_{ij}^{b3b1}}{dt} = -\gamma_y^{b3b1} y_{ij}^{b3b1} - f(x_{ij}^{b3})[y_{ij}^{b3b1} - Y^{b3b1}]^+ [\rho_y^{b3b1} + K_y^{b3b1} f(x_{ij}^{b1})]$$

$$+ \beta_y^{b3b1}(u_{ij}^{b3b1} - y_{ij}^{b3b1}) \tag{7.13}$$

7.2.1.3 SAART's Field B3

Field B3 receives bottom-up synaptic inputs from Field A1 and the top-down synaptic inputs from Field B2. The feedforward-feedback interactions between Fields A1 and B3 constitute the FFE-FBPF neural circuit (section 4.8, Chapter 4). The top-down input into Field B3 from Field B2 therefore has a higher postsynaptic gain than the bottom-up input from Field A1. This ensures that when the top-down memory is activated, it will mix with the bottom-up initiated activity in Field B3. If the activated top-down memory is strong, then it will completely overwrite the contents of B3, which in turn facilitates its own bottom-up input from A1. This enables the top-down memory to influence what is to enter into the network.

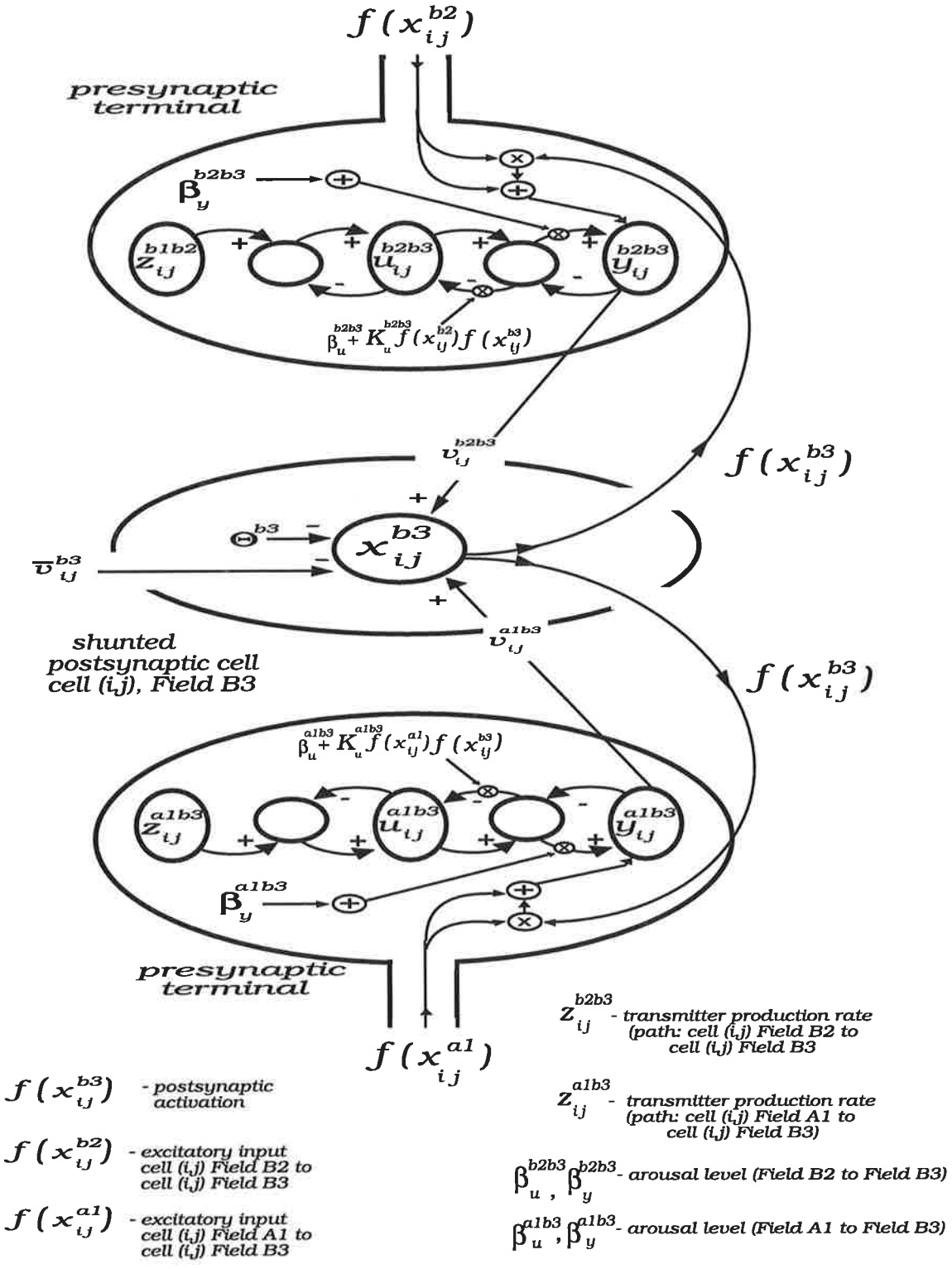


FIGURE 7.4. Pre-postsynaptic interactions in SAART's Field B3.

Postsynaptic activation at equilibrium

$$x_{ij}^{b3} = \frac{B^{b3}[G^{b2b3}v_{ij}^{b2b3} + G^{a1b3}v_{ij}^{a1b3}]}{A^{b3} + G^{b2b3}v_{ij}^{b2b3} + G^{a1b3}v_{ij}^{a1b3} + \overline{G}^{b3}\overline{v}_{ij}^{b3}} \tag{7.14}$$

Inhibitory postsynaptic potential due to lateral competition

$$\frac{d\overline{v}_{ij}^{b3}}{dt} = -\overline{A}^{b3}\overline{v}_{ij}^{b3} + \overline{B}^{b3}\sum_{(k,m)\neq(i,j)} f(x_{km}^{b3}) \tag{7.15}$$

Excitatory postsynaptic potential (pathway $A1 \rightarrow B3$)

$$\frac{dv_{ij}^{a1b3}}{dt} = -D^{a1b3}v_{ij}^{a1b3} + f(x_{ij}^{a1})[y_{ij}^{a1b3} - Y^{a1b3}]^{+}[\rho_{v}^{a1b3} + K_{v}^{a1b3}f(x_{ij}^{b3})] \tag{7.16}$$

Stored Transmitter (pathway $A1 \rightarrow B3$)

$$\frac{du_{ij}^{a1b3}}{dt} = \alpha_{u}^{a1b3}(z_{ij}^{a1b3} - u_{ij}^{a1b3}) - [\beta_{u}^{a1b3} + K_{u}^{a1b3}f(x_{ij}^{a1})f(x_{ij}^{b3})](u_{ij}^{a1b3} - y_{ij}^{a1b3}) \tag{7.17}$$

Mobilized Transmitter (pathway $A1 \rightarrow B3$)

$$\frac{dy_{ij}^{a1b3}}{dt} = -\gamma_{y}^{a1b3}y_{ij}^{a1b3} - f(x_{ij}^{a1})[y_{ij}^{a1b3} - Y^{a1b3}]^{+}[\rho_{y}^{a1b3} + K_{y}^{a1b3}f(x_{ij}^{b3})] + \beta_{y}^{a1b3}(u_{ij}^{a1b3} - y_{ij}^{a1b3}) \tag{7.18}$$

The next three equations describe the top-down synaptic dynamics from Field B2.

Excitatory postsynaptic potential (pathway $B2 \rightarrow B3$)

$$\frac{dv_{ij}^{b2b3}}{dt} = -D^{b2b3}v_{ij}^{b2b3} + f(x_{ij}^{b2})[y_{ij}^{b2b3} - Y^{b2b3}]^{+}[\rho_{v}^{b2b3} + K_{v}^{b2b3}f(x_{ij}^{b3})] \tag{7.19}$$

Stored Transmitter (pathway $B2 \rightarrow B3$)

$$\frac{du_{ij}^{b2b3}}{dt} = \alpha_u^{b2b3}(z_{ij}^{b2b3} - u_{ij}^{b2b3}) - [\beta_u^{b2b3} + K_u^{b2b3} f(x_{ij}^{b2}) f(x_{ij}^{b3})](u_{ij}^{b2b3} - y_{ij}^{b2b3}) \tag{7.20}$$

Mobilized Transmitter (pathway $B2 \rightarrow B3$)

$$\frac{dy_{ij}^{b2b3}}{dt} = -\gamma_y^{b2b3} y_{ij}^{b2b3} - f(x_{ij}^{b2})[y_{ij}^{b2b3} - Y^{b2b3}]^+ [\rho_y^{b2b3} + K_y^{b2b3} f(x_{ij}^{b3})] + \beta_y^{b2b3}(u_{ij}^{b2b3} - y_{ij}^{b2b3}) \tag{7.21}$$

7.2.1.4 SAART's Field B2

Each cell in Field B2 has a top-down presynaptic terminal originating from all the cells in Field C1, while it has only one presynaptic terminal from the corresponding cell in Field B1. The cell interacts with all the synapses that abut it, sending feedback to all the top-down synapses to provide them with a postsynaptic learning signal, while it interacts with the signals in the bottom-up synapses to modulate their transmitter depletion and release.

Postsynaptic activation at equilibrium

$$x_{ij}^{b2} = \frac{B^{b2}\left[G^{c1b2}\sum_k v_{kij}^{c1b2} + G^{b1b2} v_{ij}^{b1b2}\right]}{A^{b2} + G^{c1b2}\sum_k v_{kij}^{c1b2} + G^{b1b2} v_{ij}^{b1b2} + \overline{G}^{b2}\overline{v}_{ij}^{b2}} \tag{7.22}$$

where $\sum_k v_{kij}^{c1b2}$ is the total EPSP on the cell due to the binding of the released transmitter from all the active top-down pathways that are driven by cells in Field C1.

Inhibitory postsynaptic potential due to lateral competition (Field B2)

$$\frac{d\overline{v}_{ij}^{b2}}{dt} = -\overline{A}^{b2}\overline{v}_{ij}^{b2} + \overline{B}^{b2}\sum_{(k,m)\neq(i,j)} f(x_{km}^{b2}) \tag{7.23}$$

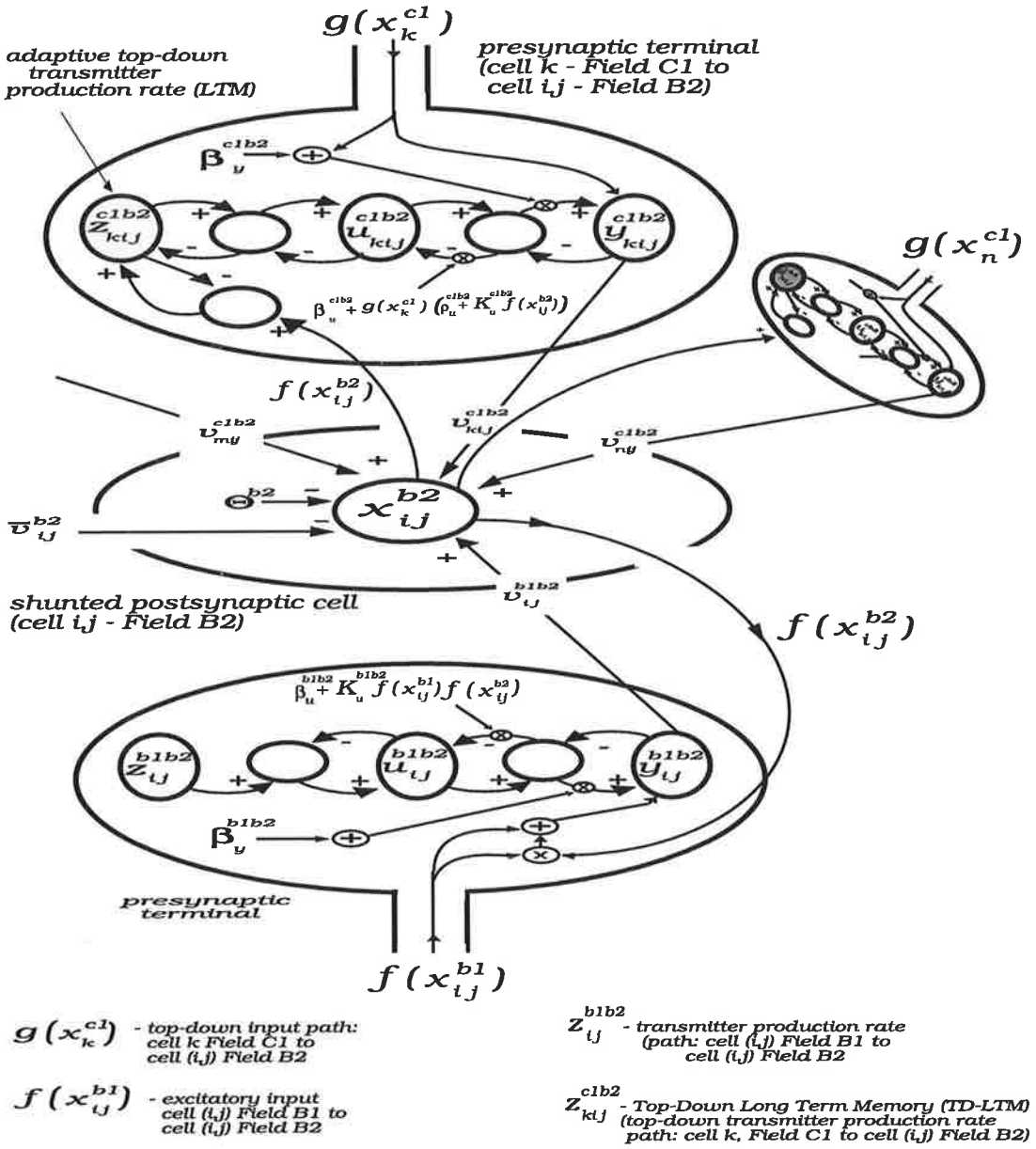


FIGURE 7.5. Pre-postsynaptic interactions in SAART's Field B2.

Excitatory postsynaptic potential due to top-down transmitter binding (pathway $C1 \rightarrow B2$)

$$\frac{dv_{kij}^{c1b2}}{dt} = -D^{c1b2} v_{kij}^{c1b2} + K_v^{c1b2} g(x_k^{c1}) [y_{kij}^{c1b2} - Y^{c1b2}]^+ \tag{7.24}$$

$$g(x_k^{c1}) = \begin{Bmatrix} 0 & if & x_k^{c1} < \overline{\theta}^{c1} \\ x_k^{c1} - \overline{\theta}^{c1} & if & x_k^{c1} \geq \overline{\theta}^{c1} \end{Bmatrix} \tag{7.25}$$

where $\overline{\theta}^{c1}$ is a defined as a "physiologically significant firing frequency" below which neurons in Field C1 do not cause transmitter release in the pathway $C1 \rightarrow B2$ ($\overline{\theta}^{c1} = 0.2$).

Stored Transmitter (pathway $C1 \rightarrow B2$)

$$\frac{du_{kij}^{c1b2}}{dt} = \alpha_u^{c1b2}(z_{kij}^{c1b2} - u_{kij}^{c1b2}) - \beta_u^{c1b2}(u_{kij}^{c1b2} - y_{kij}^{c1b2}) - g(x_k^{c1})[\rho_u^{c1b2} + K_u^{c1b2} f(x_{ij}^{b2})](u_{kij}^{c1b2} - y_{kij}^{c1b2}) \tag{7.26}$$

The term $-\beta_u^{c1b2}(u_{kij}^{c1b2} - y_{kij}^{c1b2})$ in the above equation is the passive decay of the stored transmitter (it decays towards the level of the mobilized transmitter). Since the mobilized and the stored transmitter interact, this decay slows down as the level of one approaches the other. The third term in the above equation represents a gated decay of the stored transmitter, Again, this decay is via the faster transmitter mobilization variable. However, when the pre-postsynaptic signals are simultaneously active, some top-down synaptic memory pathways will have their level of the stored transmitter selectively depleted. Hence if the top-down LTM variable (z_{kij}^{c1b2}) also decays (via the u_{kij}^{c1b2} variable), this process can lead to selective erasure of some features from a top-down memory. The need for this type of process in the top-down memory pathways will be discussed in section 7.2.2.

Mobilized Transmitter (pathway $C1 \rightarrow B2$)

$$\frac{dy_{kij}^{c1b2}}{dt} = -\gamma_y^{c1b2} y_{kij}^{c1b2} - K_y^{c1b2} g(x_k^{c1})[y_{kij}^{c1b2} - Y^{c1b2}]^+ + [\beta_y^{c1b2} + H_y^{c1b2} g(x_k^{c1})](u_{kij}^{c1b2} - y_{kij}^{c1b2}) \tag{7.27}$$

The next three equations describe the synaptic dynamics in the $B1 \rightarrow B2$ pathway.

Excitatory postsynaptic potential (pathway $B1 \rightarrow B2$)

$$\frac{dv_{ij}^{b1b2}}{dt} = -D^{b1b2} v_{ij}^{b1b2} + f(x_{ij}^{b1b2})[y_{ij}^{b1b2} - Y^{b1b2}]^+ [\rho_v^{b1b2} + K_v^{b1b2} f(x_{ij}^{b2})] \tag{7.28}$$

Stored Transmitter pathway ($B1 \rightarrow B2$)

$$\frac{du_{ij}^{b1b2}}{dt} = \alpha_u^{b1b2}(z_{ij}^{b1b2} - u_{ij}^{b1b2}) - [\beta_u^{b1b2} + K_u^{b1b2} f(x_{ij}^{b1}) f(x_{ij}^{b2})](u_{ij}^{b1b2} - y_{ij}^{b1b2}) \tag{7.29}$$

Mobilized Transmitter (pathway $B1 \rightarrow B2$)

$$\frac{dy_{ij}^{b1b2}}{dt} = -\gamma_y^{b1b2} y_{ij}^{b1b2} - f(x_{ij}^{b1})[y_{ij}^{b1b2} - Y^{b1b2}]^+ [\rho_y^{b1b2} + K_y^{b1b2} f(x_{ij}^{b2})] + \beta_y^{b1b2}(u_{ij}^{b1b2} - y_{ij}^{b1b2}) \tag{7.30}$$

Dynamics of the top-down LTM variable (z_{kij}^{c1b2}) is given in section 7.2.2.

7.2.1.5 SAART's Field C1

Field C1 differs from all the others because it is implemented as a winner-take-all layer. Furthermore, we also use a simple model of postsynaptic receptors whose desensitization provides a mechanism for attention shifting upon the network reset. Figure 7.6 shows the cellular and synaptic interactions of Field C1. In order to model fast memory search process, we also embed lateral presynaptic competition at Field C1.

Postsynaptic activation at equilibrium (cell k, Field C1)

$$x_k^{c1} = \frac{B^{c1} G^{b1c1} \sum_{(i,j)} v_{ijk}^{b1c1}}{A^{c1} + G^{b1c1} \sum_{(i,j)} v_{ijk}^{b1c1} + \overline{G}^{c1} \overline{v}_k^{c1}} \tag{7.31}$$

where $\sum_{(i,j)} v_{ijk}^{b1c1}$ is the total EPSP on the k^{th} cell that is due to the binding of the released transmitter from all the active bottom-up pathways from Field B1.

Inhibitory postsynaptic potential due to lateral competition

$$\frac{d\overline{v}_k^{c1}}{dt} = -\overline{A}^{c1} \overline{v}_k^{c1} + \overline{B}^{c1} \sum_{m \neq k} h(x_m^{c1}) \tag{7.32}$$

where $h(x_m^{c1})$ is faster than linear above a threshold, represented by equation (7.33). The faster than linear thresholding function in Field C1 models a winner-take-all competitive layer.

$$h(x_m^{c1}) = \begin{Bmatrix} 0 & if & x_m^{c1} < \theta^{c1} \\ 10(x_m^{c1} - \theta^{c1})^2 & if & x_m^{c1} \geq \theta^{c1} \end{Bmatrix} \tag{7.33}$$

Excitatory postsynaptic potential due to bottom-up transmitter binding

$$\frac{dv_{ijk}^{b1c1}}{dt} = -D^{b1c1} v_{ijk}^{b1c1} + K_v^{b1c1} w_{ijk}^{b1c1} [r_{ijk}^{b1c1} - r^{b1c1}]^+ \tag{7.34}$$

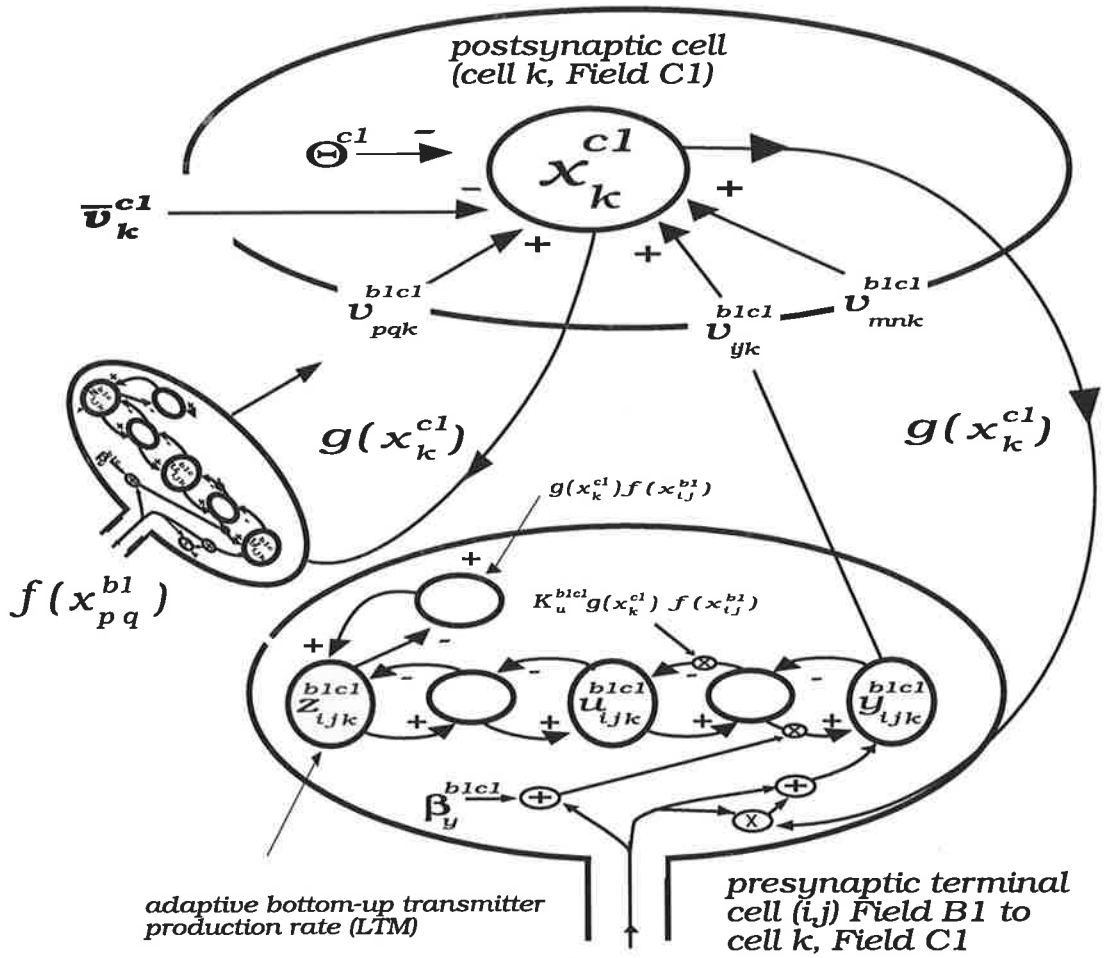
where w_{ijk}^{b1c1} is the amount of transmitter that is released from the $(i,j)^{th}$ synaptic pathway from Field B1 into k^{th} cell of Field C1 and is given by (7.35); r_{ijk}^{b1c1} is the activation level of the postsynaptic receptor in the same pathway and is given by (7.36).

Released transmitter

$$w_{ijk}^{b1c1} = f(x_{ij}^{b1}) [\beta_w^{b1c1} + K_w^{b1c1} h(x_k^{c1})] y_{ij}^{b1c1} \tag{7.35}$$

Postsynaptic receptor (on cells in Field C1)

The reason for modelling postsynaptic receptors on the cells of Field C1 is to provide a dynamic mechanism that can bias an active cell in Field C1 against continually winning the competition after the network is reset. Since the level of the postsynaptic receptor on the active cell(s) in Field C1 is reduced when the bottom-up synaptic pathways are active, the resultant EPSP on the winning cell in Field C1 is also reduced. Due to strong self-excitation, other non-active cells in Field C1 cannot overcome the lateral inhibition that is provided by the already activated cell. However, when the network is reset, other cells can subsequently win the competition since their postsynaptic receptor levels will be above the level of previously activated cells whose receptors were desensitized. The following shunting equation models postsynaptic receptors along the $(i,j)^{th}$ synaptic pathway from Field B1 to the k^{th} neuron in Field C1.



x_k^{c1} - *postsynaptic activation (cell k, Field C1)*

v_{ijk}^{b1c1} - *excitatory post-synaptic potential due to bound transmitter (path: cell (i,j) Field B1 to cell k, Field C1)*

$\overline{v}_k^{c1}$ - *inhibitory feedback from lateral interactions*

Θ^{c1} - *postsynaptic threshold*

$g(x_k^{c1})$ - *postsynaptic output (cell k, Field C1)*

$f(x_{ij}^{b1})$ - *presynaptic input (cell (i,j) Field B1 to cell k, Field C1)*

z_{ijk}^{b1c1} - *Bottom-Up Long Term Memory (BU-LTM) (adaptive transmitter production rate cell (i,j), Field B1 to cell k, Field C1)*

u_{ijk}^{b1c1} - *stored transmitter (cell (i,j), Field B1 to cell k, Field C1)*

y_{ijk}^{b1c1} - *mobilized transmitter (cell (i,j), Field B1 to cell k, Field C1)*

FIGURE 7.6. Pre-postsynaptic interactions in SAART's Field C1. For simplicity, we have not shown the postsynaptic receptors on the cells of Field C1.

$$\frac{dr_{ijk}^{b1c1}}{dt} = [\alpha_r^{b1c1} + \mu_{ijk}^{b1c1}](1 - r_{ijk}^{b1c1}) - r_{ijk}^{b1c1} w_{ijk}^{b1c1} [\beta_r^{b1c1} + K_r^{b1c1} g(x_k^{c1} - \Phi^{c1})] \qquad (7.36)$$

where μ_{ijk}^{b1c1}, given by (7.37), quickly restores (or sensitizes) postsynaptic receptors after the network reaches a steady resonant state (i.e., when the network recognizes its input). The second term in the above equation says that only those synaptic pathways whose postsynaptic cell is firing above a large threshold (Φ^{c1}) and whose bottom-up signals are active will have their level of postsynaptic receptor reduced by an amount that is proportional to the correlated firing of pre-postsynaptic signals.

$$\frac{d\mu_{ijk}^{b1c1}}{dt} = -\gamma_{\mu}\mu_{ijk}^{b1c1} + \Psi\mu_0^{b1c1} p(x_k^{c1})(1-\mu_{ijk}^{b1c1}) \tag{7.37}$$

where

$$\Psi = \begin{cases} 1 & \text{if } R^{b1b3} \geq \rho \text{ AND if } dR^{b1b3}/dt \leq \dot{\rho} \\ 0 & \text{otherwise} \end{cases} \tag{7.38}$$

In the above equation, R^{b1b3} is the match between the spatial patterns across Fields B1 and B3, while dR^{b1b3}/dt is the time rate of change of the match (ρ is the match threshold, $\dot{\rho}$ is the steady state threshold); $p(x_k^{c1}) = 1$ if the k^{th} cell in Field C1 is active above half of the maximum (otherwise $p(x_k^{c1}) = 0$). That is, the postsynaptic receptors on inactive cells in Field C1 are rapidly sensitized when the network has recognized its current input.

Mobilized Transmitter

$$\frac{dy_{ijk}^{b1c1}}{dt} = -\gamma_y^{b1c1} y_{ijk}^{b1c1} - \rho_y^{b1c1} w_{ijk}^{b1c1} + \frac{[\beta_y^{b1c1} + K_y^{b1c1} f(x_{ij}^{b1})](u_{ijk}^{b1c1} - y_{ijk}^{b1c1})}{[1 + W_k^{b1c1}]} \tag{7.39}$$

where W_k^{b1c1}, given by (7.40), is the total transmitter that is released onto the postsynaptic cell and which contributes to presynaptic inhibition of all active bottom-up synapses into k^{th} cell of Field C1.

$$W_k^{b1c1} = \sum_{(i,j)} w_{ijk}^{b1c1} \tag{7.40}$$

Stored Transmitter (pathway $B1 \rightarrow C1$)

$$\frac{du_{ijk}^{b1c1}}{dt} = \alpha_u^{b1c1}(z_{ijk}^{b1c1} - u_{ijk}^{b1c1}) - K_u^{b1c1} g(x_k^{c1}) f(x_{ij}^{b1}) \tag{7.41}$$

The second term in the above equation implies that the stored transmitter (meduim term memory or MTM) is depleted by the correlated pre-postsynaptic activityactivity o-lpiesis equation sa

Dynamics of the bottom-up long term memory variable (z_{ijk}^{b1c1}) is given in section 7.2.2.

7.2.2 Gated Decay LTM Equations

The learning equations used in the SAART neural network are based on Grossberg's gated decay long term memory (LTM) equation (Chapter 3). Through numerous computer simulations of the network and at various learning rates, we have found that during slow learning, several different shapes may be encoded in the same memory pathways. However, as learning proceeds and different spatial patterns begin to establish their own memory, it becomes necessary to selectively "forget" the memory of those features that do not belong to the individual inputs. This selective forgetting is needed and is crucial in the top-down memory pathways to ensure rapid removal of irrelevant features from memory, thus enabling the network to learn effectively in noisy inputs. Selective forgetting in the top-down memory pathways also ensures that as the top-down memory converges, the initial differences in the individual memory pathways for a particular input are levelled out and converge to the same asymptote. Based on our computer simulations, we have arrived at the following two learning equations which provide a good compromise between slow learning and selective forgetting.

Top-down gated LTM learning equation

$$\frac{dz_{Jij}^{c1b2}}{dt} = -\gamma_z^{c1b2} \overline{g}(x_J^{c1})(z_{Jij}^{c1b2} - u_{Jij}^{c1b2}) + \Psi \varepsilon_z^{c1b2}[f(x_{ij}^{b2}) - z_{Jij}^{c1b2}] \tag{7.42}$$

Bottom-up gated LTM learning equation

$$\frac{dz_{ijJ}^{b1c1}}{dt} = -\gamma_z^{b1c1} \overline{g}(x_J^{c1})(z_{ijJ}^{b1c1} - u_{ijJ}^{b1c1}) + \Psi \varepsilon_z^{b1c1}[f(x_{ij}^{b1}) - z_{ijJ}^{b1c1}] \tag{7.43}$$

where Ψ is the learn enable signal (as specified by equation (7.38)), ε_z^{c1b2} and ε_z^{b1c1} are the learning rates in the top-down and the bottom-up memory pathways respectively, while γ_z^{c1b2} and γ_z^{b1c1} are their respective memory decay rates. The thresholding function $\overline{g}(x_J^{c1})$,

given by (7.44), is a gating term that enables the decay of the LTM variables (z_{Jij}^{c1b2} and z_{ijJ}^{b1c1}) for the J^{th} cell of Field C1, but only when the activity of the cell is below a high threshold.

$$\overline{g}(x_J^{c1}) = \begin{Bmatrix} 1 & if \quad x_J^{c1} < 0.5 \\ 0 & otherwise \end{Bmatrix} \tag{7.44}$$

Note that selective forgetting is enabled only after the winning cell is deactivated by memory reset, while the LTM variables are incremented only if the top-down memory from the winning cell leads to the match of the spatial patterns across Fields A1 and B3. Thus, once the activity of the winning cell in Field C1 is reset and as long as the cell's activity level is below half of the maximum possible level, then its LTM decays. This gated LTM decay is via the medium term memory (MTM) variable u_{Jij}^{c1b2}, whose level is reduced by pre-postsynaptic signals while the cell was very active (see equation (7.26)). However, since the MTM variable tracks the LTM variable, the rate of the decay decreases with time. The effect of selective forgetting will be demonstrated in section 7.3.2.1, where we test the network's learning capability on noisy but patterned 2-D input data.

7.2.3 Simulation of the SAART Neural Network

Below we present a computer simulation of the SAART neural network on the problem of shape recognition in a cluttered bottom-up input (the network parameters are given in Appendix B.2). The simulation is designed to show that once the memory of an object's shape has stabilized, it enables the network to recognise (perceive) the familiar shape in a cluttered input, provided that the complete boundary is available.

(a) Input 1 (shape 1) (b) Input 2 (shape 2) (c) Common elements between Shapes 1 and 2. (d) Input 3 (shape 1 in cluttered background)

FIGURE 7.7. Inputs presented to the SAART neural network (each image is 32x32 elements, and are shown in reverse contrast).

The network is initially trained on two different shapes (boundaries of two ships), illustrated in Figure 7.7. Shapes 1 and 2 (inputs 1 and 2) are spatially aligned to maximise their common elements, as shown in (c), the purpose of which is to maximise the difficulty of discrimination between the two. The order of presentation is as follows:

input 1 for 100 network iterations (learns shape 1); input 2 for another 100 iterations (learns shape 2); input 3 for 30 iterations (resonance with shape 1) followed by input offset.

We now wish to observe the intermediate activities of synapses and cells at several network layers when one of the inputs is presented in a complex background, shown in Figure 7.7(d). The simulations data will be shown from the instant the new input is presented until the steady state resonance with the correct top-down memory. The purpose of presenting input 3 after input 2 is to demonstrate mismatch reset with the active memory of shape 2 followed by the activation and resonance with the correct memory.

In Figure 7.8, all shown variables are scaled by the maximum activity and are shown in reverse contrast. The simulation time (number of iterations) is shown bracketed in the far left column. The numbers below the indicated time are: the degree of match between neural activities of Fields B3 and A1 (threshold = 0.98, measured by the cosine of the angle between the two spatial patterns across B3 and A1), followed by the time rate of change of the match (measured every four iterations, steady state when $|\frac{dR^{b3a1}}{dt}| \leq 0.00005$).

The simulation data in Figure 7.8 shows that the network was reset at time t = 210 and that a stable resonant state is reached at t = 218. Comparison between x_{ij}^{b3} and x_{ij}^{a1} at t = 218 shows a match of pattern activity. Figure 7.9 below shows the dynamics for one top-down synaptic memory pathway from a cell in Field C1 during the four phases of the simulation.

The graph in Figure 7.9 shows that the transmitter production rate is increased (and converges) during the resonant period with input 1 (ship 1). These memory variables decay during the presentation of input 2 (ship 2) because the new input cannot resonate with the top-down memory of input 1 and activates a new cell in Field C1 (cell 2). When input 3 is presented to the network the top-down facilitation and the competitive interactions at Field A1 have enabled the network to resonate with shape 1.

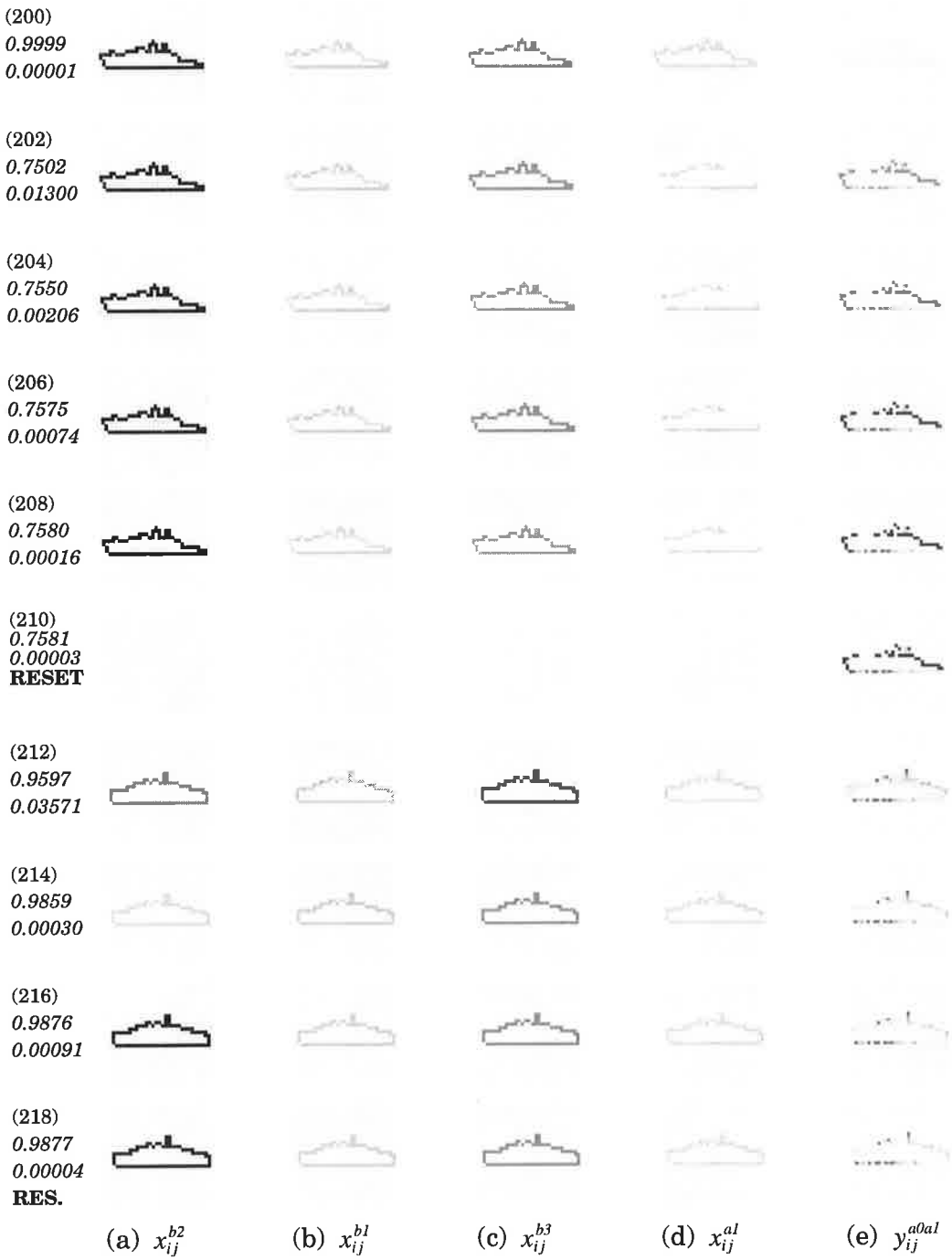


FIGURE 7.8. SAART network variables shown as a function of time from the instant input 2 of Figure 7.7 is replaced by input 3 (time flows downwards). (a) cellular activity of Field B2; (b) cellular activity of Field B1; (c) cellular activity of Field B3; (d) cellular activity of Field A1; (e) mobilized transmitter levels at the input synapses of Field A1 (note the higher synaptic gain for those input pathways whose bottom-up inputs are low in magnitude).

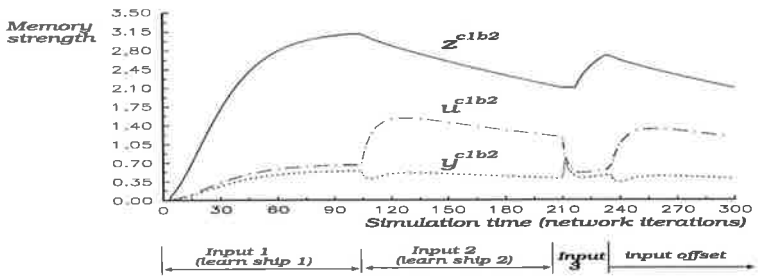


FIGURE 7.9. Dynamics of the SAART's top-down transmitter (memory) variables for one pathway.

Above results demonstrate that the network can recognize shape 1 when it appears in a cluttered background of input 3, despite the non-uniform edge intensities along its boundary. These simulations also show that although input 3 contains a familiar input (shape 1), but in a cluttered background, the network has to first undergo a reset before the memory of shape 1 is activated. In the next section we provide a simple extension to the network that enables familiar inputs do quickly activate their memory by having a direct pathway to their bottom-up LTM (thus reducing the network's reaction time, i.e., time to recognize an input).

7.3 Direct Memory Access SAART Neural Network

The SAART network as presented above has an unusual problem of not being able to recognise familiar stimuli under special circumstances. This occurs when a new (unfamiliar) stimulus is presented for a brief period of time and is learned at a slow learning rate (thus preventing its top-down LTM to grow high), followed by the presentation of a familiar stimulus. If the unfamiliar stimulus has caused a memory reset and activation of a new (previously uncommitted) cell in Field C1, then the familiar stimulus will begin to resonate with this cell. That is, if the top-down LTM of the new stimulus is too weak to cause a memory mismatch and reset when the familiar stimulus is presented, then the familiar stimulus will begin to be encoded in these new memory pathways. Hence it is possible for one stimulus to be learned by a large number of memory pathways. This is undesirable because it leads to recognition failure of a familiar stimulus. In order to solve this problem, we extend the SAART architecture of Figure 7.1 by adding another layer of neurons (Field A2) whose output has direct access to the bottom-up ($B1 \rightarrow C1$) LTM pathways, as shown in Figure 7.10.

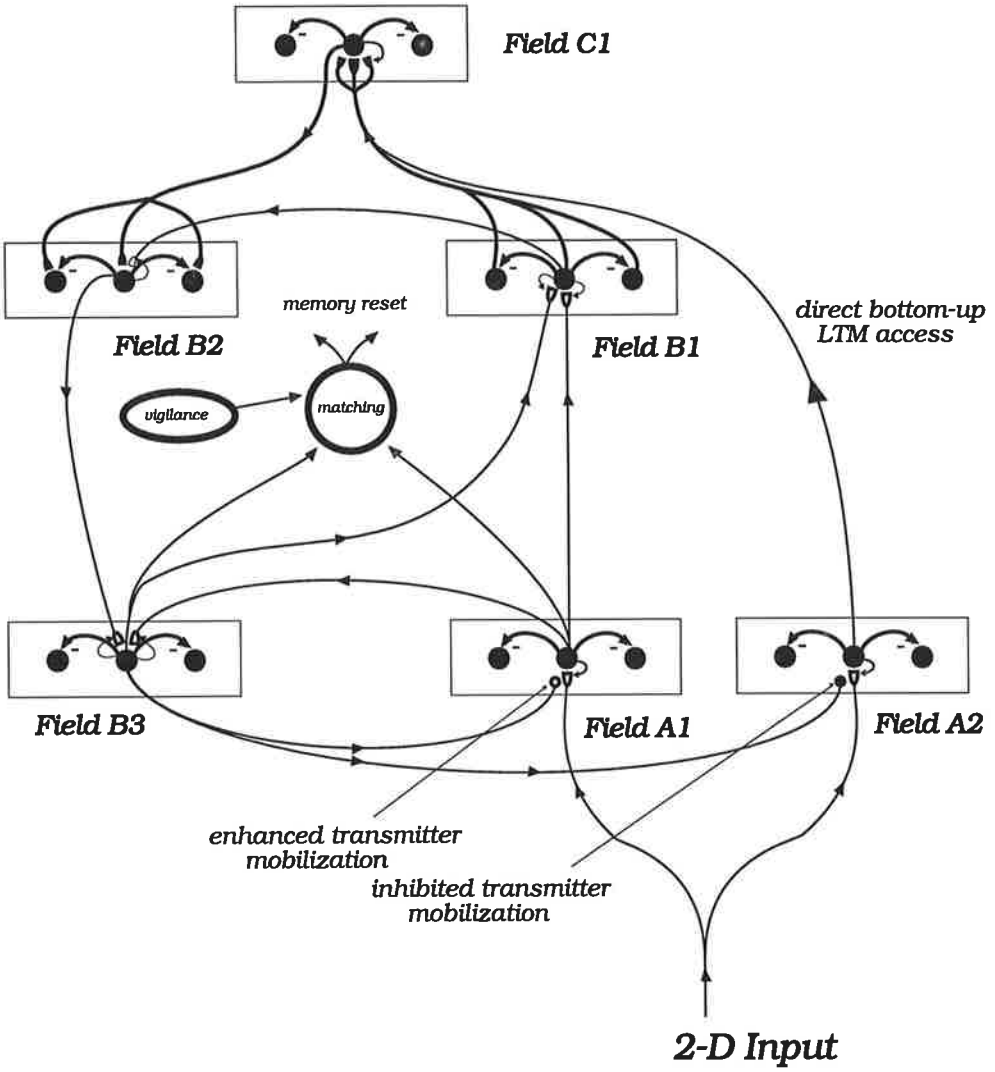


FIGURE 7.10. SAART architecture with direct bottom-up LTM access.

The following five equations describe the dynamics of Field A2. Note that as a result of adding Field A2, we also need to modify the equations in the bottom-up LTM pathways to Field C1.

Postsynaptic activity at equilibrium (Field A2)

$$x_{ij}^{a2} = \frac{B^{a2} G^{a0a2} v_{ij}^{a0a2}}{A^{a2} + G^{a0a2} v_{ij}^{a0a2} + \overline{G}^{a2} \overline{v}_{ij}^{a2}} \tag{7.45}$$

Inhibitory postsynaptic potential (IPSP) due to lateral competition (Field A2)

$$\frac{d\overline{v}_{ij}^{a2}}{dt} = -\overline{A}^{a2} \overline{v}_{ij}^{a2} + \overline{B}^{a2} \sum_{(k,m) \neq (i,j)} f(x_{km}^{a2}) \tag{7.46}$$

Excitatory postsynaptic potential

$$\frac{dv_{ij}^{a0a2}}{dt} = -D^{a0a2}v_{ij}^{a0a2} + J_{ij}[y_{ij}^{a0a2} - Y^{a0a2}]^{+}[\rho_{v}^{a0a2} + K_{v}^{a0a2}f(x_{ij}^{a2})] \tag{7.47}$$

Stored Transmitter

$$\frac{du_{ij}^{a0a2}}{dt} = \alpha_{u}^{a0a2}(z_{ij}^{a0a2} - u_{ij}^{a0a2}) - [\beta_{u}^{a0a2} + K_{u}^{a0a2}J_{ij}f(x_{ij}^{a2})](u_{ij}^{a0a2} - y_{ij}^{a0a2}) \tag{7.48}$$

Mobilized Transmitter

$$\frac{dy_{ij}^{a0a2}}{dt} = -\gamma_{y}^{a0a2}y_{ij}^{a0a2} - J_{ij}[y_{ij}^{a0a2} - Y^{a0a2}]^{+}[\rho_{y}^{a0a2} + K_{y}^{a0a2}f(x_{ij}^{a2})] + \left[\beta_{y}^{a0a2} + \frac{H_{y}^{a0a2}}{1 + \overline{H}_{y}^{a0a2}f(x_{ij}^{b3})}\right](u_{ij}^{a1b1} - y_{ij}^{a1b1}) \tag{7.49}$$

The denominator in the last term of equation (7.49) says that the transmitter mobilization rate in the $(i,j)^{th}$ input synapse of Field A2 is inhibited by the activity of the $(i,j)^{th}$ cell in Field B3. Hence if the input synapses to Field A2 have a high tonic level of transmitter mobilization ($\beta_{y}^{a0a2} + H_{y}^{a0a2}$, where $\beta_{y}^{a0a2} \ll H_{y}^{a0a2}$), the cells within the layer will be highly activated by their inputs. When the cells within Field B3 are activated above their threshold, then they inhibit the mobilization rate in the synapses of the corresponding cells in A2. This process can thus deactivate Field A2. The output of Field A2 interacts with the bottom-up LTM pathway from Field B1 to Field C1 by also being involved in the transmitter release and the excitation of neurons in Field C1. Hence equations (7.35), (7.39) and (7.41) from section 7.2.1.5 can now be rewritten to incorporate the direct bottom-up memory access by neurons of Field A2.

Released transmitter in the $B1 \rightarrow C1$ pathway

Equation (7.35) now becomes

$$w_{ijk}^{b1c1} = f(x_{ij}^{b1})\, y_{ij}^{b1c1}[\beta_{w}^{b1c1} + K_{w}^{b1c1}h(x_{k}^{c1})] + f(x_{ij}^{a2})\, y_{ij}^{b1c1}[\pi_{w}^{b1c1} + E_{w}^{b1c1}h(x_{k}^{c1})] \tag{7.50}$$

Mobilized Transmitter in the $B1 \rightarrow C1$ pathway

Equation (7.39) becomes

$$\frac{dy_{ijk}^{b1c1}}{dt} = -\gamma_y^{b1c1} y_{ijk}^{b1c1} - \rho_y^{b1c1} w_{ijk}^{b1c1} + \frac{[\beta_y^{b1c1} + K_y^{b1c1} f(x_{ij}^{b1}) + E_y^{b1c1} f(x_{ij}^{a2})](u_{ijk}^{b1c1} - y_{ijk}^{b1c1})}{[1 + W_k^{b1c1}]} \tag{7.51}$$

Stored Transmitter in the $B1 \rightarrow C1$ pathway

Equation (7.41) becomes

$$\frac{du_{ijk}^{b1c1}}{dt} = \alpha_u^{b1c1}(z_{ijk}^{b1c1} - u_{ijk}^{b1c1}) - K_u^{b1c1} g(x_k^{c1})[f(x_{ij}^{b1}) + f(x_{ij}^{a2})] \tag{7.52}$$

Field A2 can thus be involved in the direct release of the transmitter in the bottom-up LTM pathways, which under special circumstances can cause attentional switches (see computer simulations in section 7.3.2.3). In the next section we describe how Field A2 can play an important part in the network dynamics, enabling rapid access to the relevant memory pathways by the bottom-up inputs that have an established memory.

7.3.1 Processing Stages in BU-DMA SAART Network

Below we describe the processing stages in the BU-DMA SAART neural network when it is presented with input signals (waveforms or 2-D shapes) that share many common features (i.e., input pathways). Note that for the network to be applicable to general signal processing and pattern recognition applications, it is assumed that these signals are represented in a 2-D format, i.e., any signal then becomes just another shape in 2-D space. For example, the electrocardiogram (EKG) signal is a waveform with time axis running in one direction (typically the horizontal axis) while the amplitude of the signal is defined along another axis (i.e., vertical). This type of signal representation is typical of most non-imaging sensors. The amplitude and the time axes thus define a 2-D space within which the shape of the EKG signal (as well as other time dependant signals, such as a seismogram) is embedded. By a suitable division of the amplitude and the time axes into a 2-D lattice to give the desired resolution for pattern discrimination, any time dependant signal can therefore be mapped onto a 2-D array of competitive neurons and be processed and learned by 2-D neural circuits and networks such as the SAART and the BU-DMA SAART.

Now consider one such signal being fed to the BU-DMA SAART neural network as shown in Figure 7.11. The input signal that is represented in two dimensions enters the network via Fields A1 and A2. Since the tonic level of transmitter mobilization in the input synapses of Field A2 is much higher than for the input synapses of Field A1, the cells of Field A2 will initially reach a much higher level of activation. The BU-DMA pathways that are driven by the active cells in Field A2 will thus be the first to activate Field C1. At the same time, the activity from Field A1 spreads to other Fields while the competition in Field C1 establishes a winning cell.

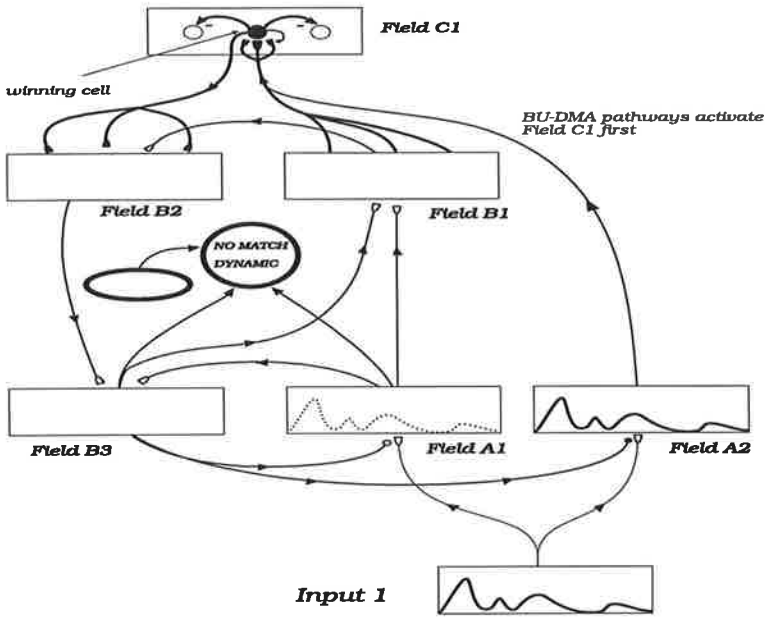


FIGURE 7.11.Initial activation of Field C1 by highly active cells in Field A2.

The reverberatory activity in the network amplifies the bottom-up inputs into Field A1 thus maintaining the activity of the winning cell in Field C1 while Field B3 presynaptically deactivates Field A2. When the steady state is reached, the reverberatory spatial pattern transfers into the bottom-up and the top-down pathways of the active cell in Field C1, Figure 7.12.

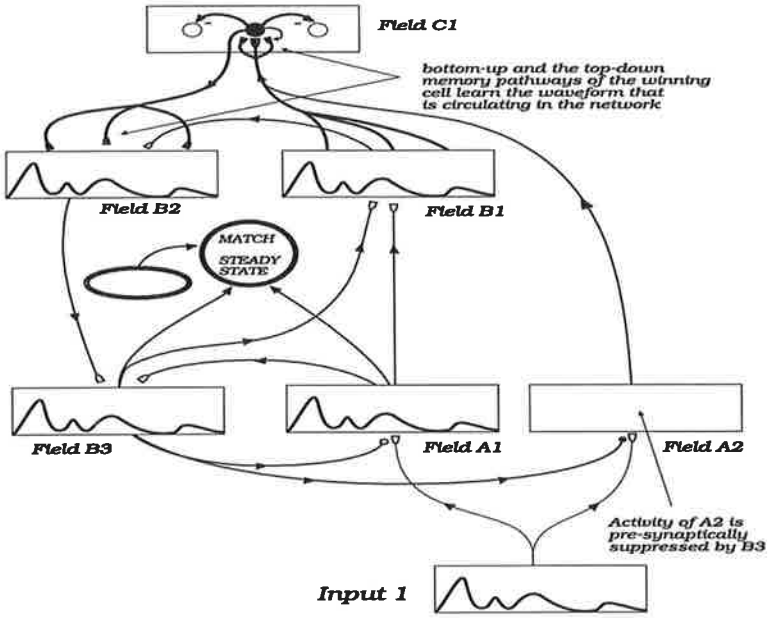


FIGURE 7.12. Establishment of a reverberatory activity and its transfer to memory.

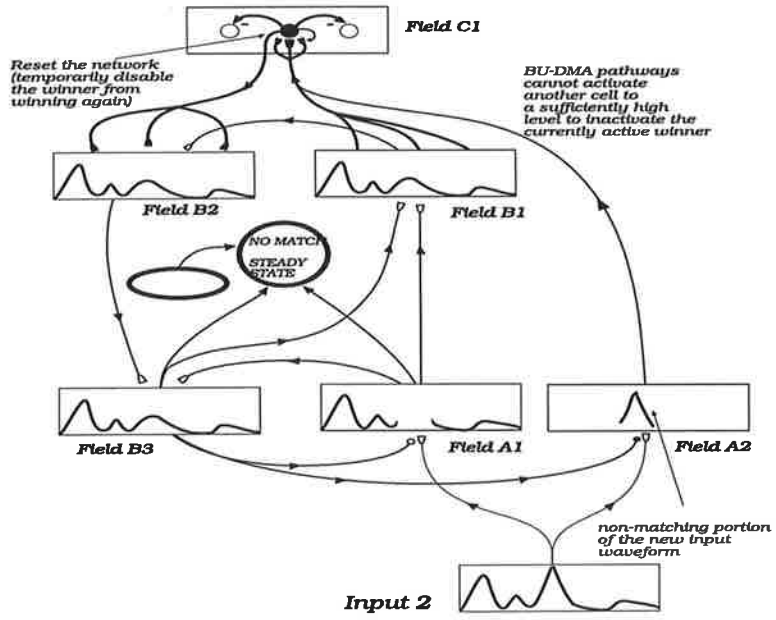


FIGURE 7.13. Non-matching portion of the next input activates A2 but cannot activate Field C1.

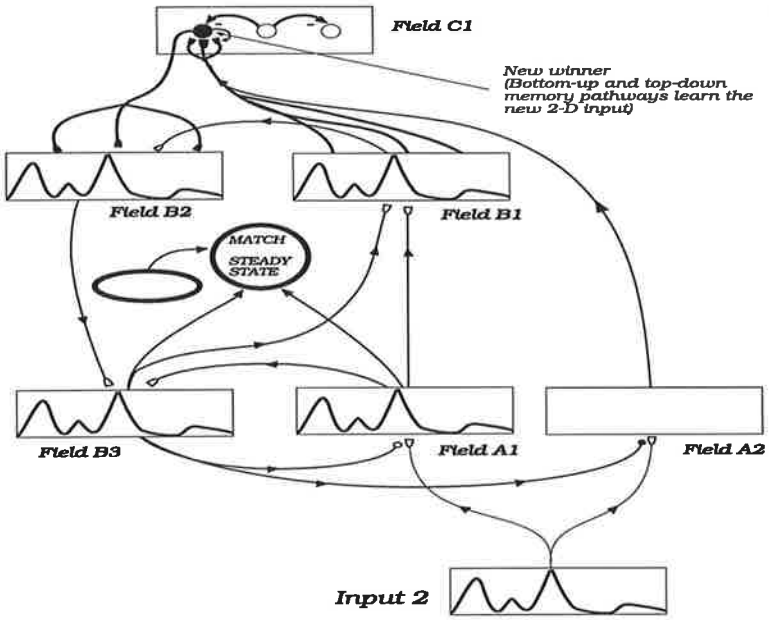


FIGURE 7.14. After network reset, the new input establishes a resonant state and gets registered into memory.

When a new input waveform (or another 2-D shape) that shares some common features (pathways) with the currently reverberating spatial pattern is presented, then the components of the new input that do not match the reverberatory pattern, as shown in Figure 7.13, will activate Field A2. Simultaneously with this, the components of the new input that do match the reverberatory activity will be amplified, thus leading to higher amplitude reverberations. The increased activity of Field A1 in the facilitated locations will suppress all other cells in Field A1 that no longer receive the input. This increased activity will also prevent the new features of the input to enter into the network via Field A1. Since the BU-DMA pathways that are driven by the activity across Field A2 do not have a prior established memory they cannot activate another cell in Field C1 to a sufficiently high level to overcome the inhibition from the already active cell. The network therefore ends up in a non-matching steady state and is then reset (i.e., all the cellular variables and the postsynaptic potentials are set to zero, leaving the synaptic variables unaltered).

Following the reset, the new input then reactivates the network and activates another memory cell in Field C1 establishes its own reverberatory activity which then gets transferred into the memory pathways of the newly activated memory cell, Figure 7.14.

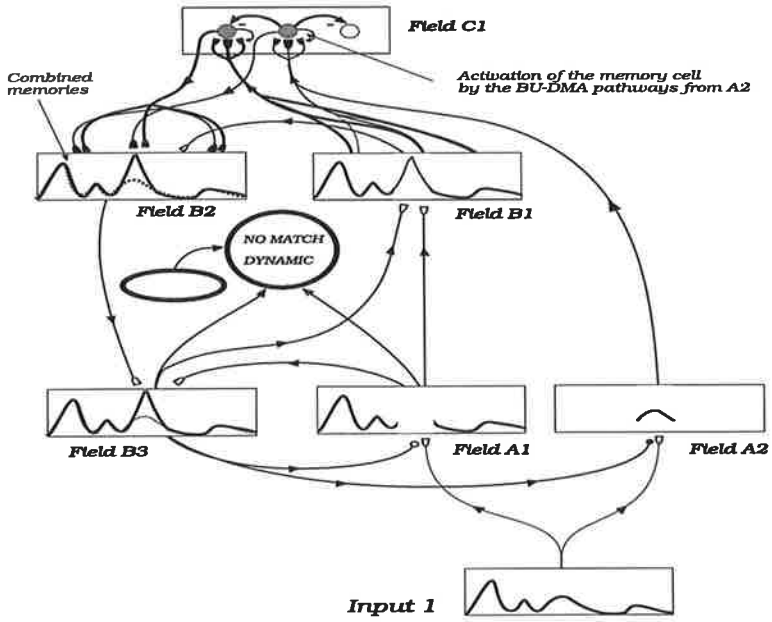


FIGURE 7.15. A familiar input activates its memory cell in Field C1 via the BU-DMA pathways.

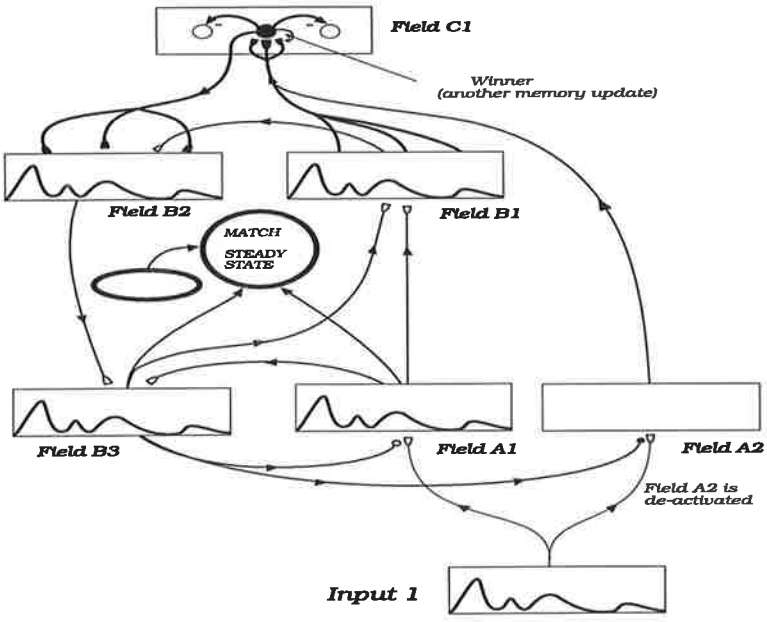


FIGURE 7.16. The familiar input has recalled its memory directly without the network reset.

When a previously learned spatial pattern is re-presented, then those components of the input that do not match the reverberatory activity and that appear across Field A2 will now be able to activate the memory cell directly via the strong bottom-up memory pathways. As shown in Fig 7.15, there is a period of time during which the already active cell and the newly activated cell in Field C1 are both sending their top-down

memory into the reverberatory loop. Although the match between the spatial patterns of activity across Fields B3 and A1 are below the required threshold level, the network is not reset because it is still in the dynamic state (i.e., the rate of change of the match is above its steady state threshold level). When the steady state is reached, the LTM of the familiar input is increased further, as indicated in Figure 7.16.

In the following section we provide several computer simulations of the BU-DMA SAART neural network.

7.3.2 Simulations of the BU-DMA SAART Neural Network

Each of the simulations in the following three sub-sections emphasises different capabilities of the network. The first simulation in section 7.3.2.1 emphasises the network's ability to self-organise its long term memory when presented with shapes that are embedded in noise. This is followed by a simulation in section 7.3.2.2 that demonstrates the network's ability to recognize objects in cluttered visual images. The last simulation in section 7.3.2.3 demonstrates the network's dynamics during perceptual reversals on a 2-D version of the Necker cube. All network parameters are given in Appendix B.2.

7.3.2.1 Learning in Noisy but Patterned Inputs

Below we show various network memories during real-time and self-organised learning in noisy but patterned input data. Because the simulated BU-DMA SAART neural network currently does not have the benefit of early visual processing, such as what can be provided by the Boundary Contour System (Grossberg and Mingolla, 1985a,b), its task is far more difficult than may be initially imagined. Since BCS is a model of early vision that emphasises neural activity of spatially aligned cells by cooperative linking, we expect that the interaction between BCS and the BU-DMA SAART neural network will provide a more powerful visual learning system. Thus, when judging the difficulty of the learning problem addressed below, we should keep in mind that because of the absence of cooperative linking between spatially aligned features, the current network initially treats all inputs equally until its memories begin to influence the processing. The difficulty of the problem can best be gauged by representing each 2-D spatial pattern as a 1-D pattern (where each row is placed side by side) as shown in Figures 7.18(a) and 7.18(b). Although the current network can (under special conditions mentioned below) learn in noisy inputs, we do not claim that this is the main feature of the network nor do we claim that it is a robust feature. However, we do claim that once the network memories are reasonably strong (have

converged), the network is robust in recognizing a familiar input in a cluttered background, as will be demonstrated in the next section where we apply the network to object recognition in cluttered visual images.

The input data consists of three shapes (boundaries of three ships) that are successively presented but each time with another random dot background. The only supervision that is provided is that after each resonance (memory update) a new input (with a new noise and shape) is presented. All other parameters of the network remain fixed. The purpose of continuously modifying the background is primarily so that the network does not overlearn any particular input (although in a full self-organising and unsupervised real-time artificial neural learning system, there is a need to automatically reduce the attentional gain when certain stimuli become familiar). Upon each memory-mismatch reset, activities of all cells in the network (and their corresponding EPSPs) are reset.

The network is thus presented with noisy 2-D inputs that contain boundaries of three objects (ships) that are shown in Figure 7.17. The binary images of the three ship boundaries are used to generate a set of noisy inputs (shown in Figure 7.18 below) using the following formula:

$$J^*(i,j) = \begin{Bmatrix} J(i,j) & \text{if} & J(i,j) & > & 0 \\ n & \text{if} & J(i,j) & = & 0 \end{Bmatrix} \tag{7.53}$$

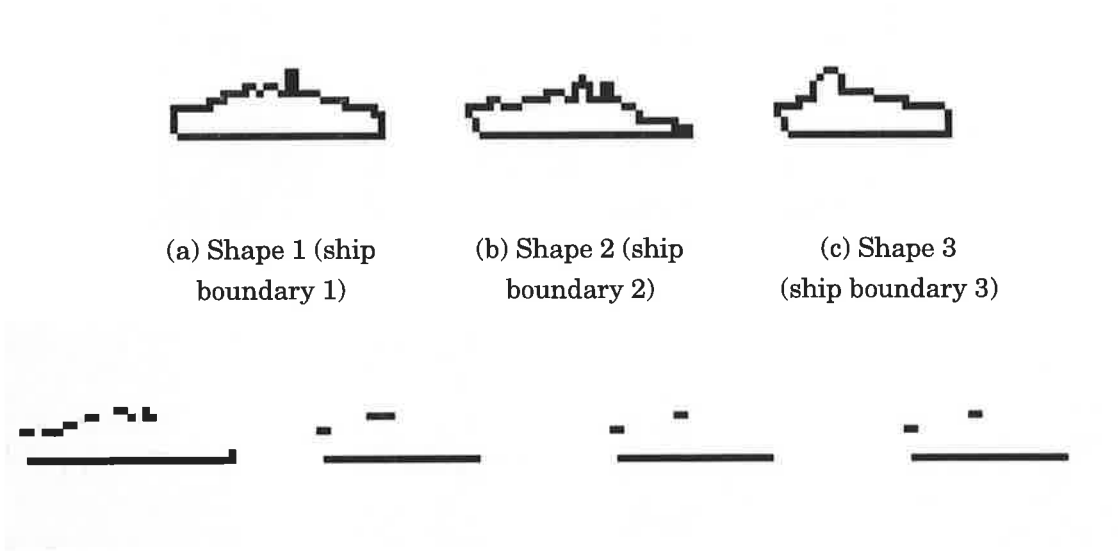
where $J^*(i,j)$ is a new input at location (i,j), $J(i,j)$ is the original input and n is a random number given by

$$n = \begin{Bmatrix} 1 & \text{if} & \text{random}(0,1) \le 0.1 \\ 0 & \text{if} & \text{random}(0,1) > 0.1 \end{Bmatrix} \tag{7.54}$$

where random(0,1) is a random number between 0 and 1. Equations (7.53) and (7.54) imply that while the original elements in the binary images of the ship boundaries are left unaltered, the background is modified by the noise. The threshold in (7.54) controls the percentage of the background elements that are modified. In the simulation shown below, this threshold is chosen to be 0.1. Hence, roughly 10% of the background elements are set to a value whose magnitude is equal to the magnitude of elements on the ship's boundary (equal to 1.0).

To enable the network to self-organise its memory and learn the stationary components in the input, the top-down learning rate is set to a small value, while the bottom-up learning rate was set to a value that is ten times larger (i.e., the parameter ε_z^{c1b2} in equation (7.42) was set to 0.02, while ε_z^{b1c1} in (7.43) was set to 0.2). The decay rates (parameters γ_z^{c1b2} and γ_z^{b1c1} in (7.42) and (7.43)) were set to 0.002. In order to enable

the network to average over several inputs before the top-down memory becomes effective, the threshold for transmitter release in the top-down memory pathways was set to 0.005 (which was predetermined by first testing the network for the first five inputs).

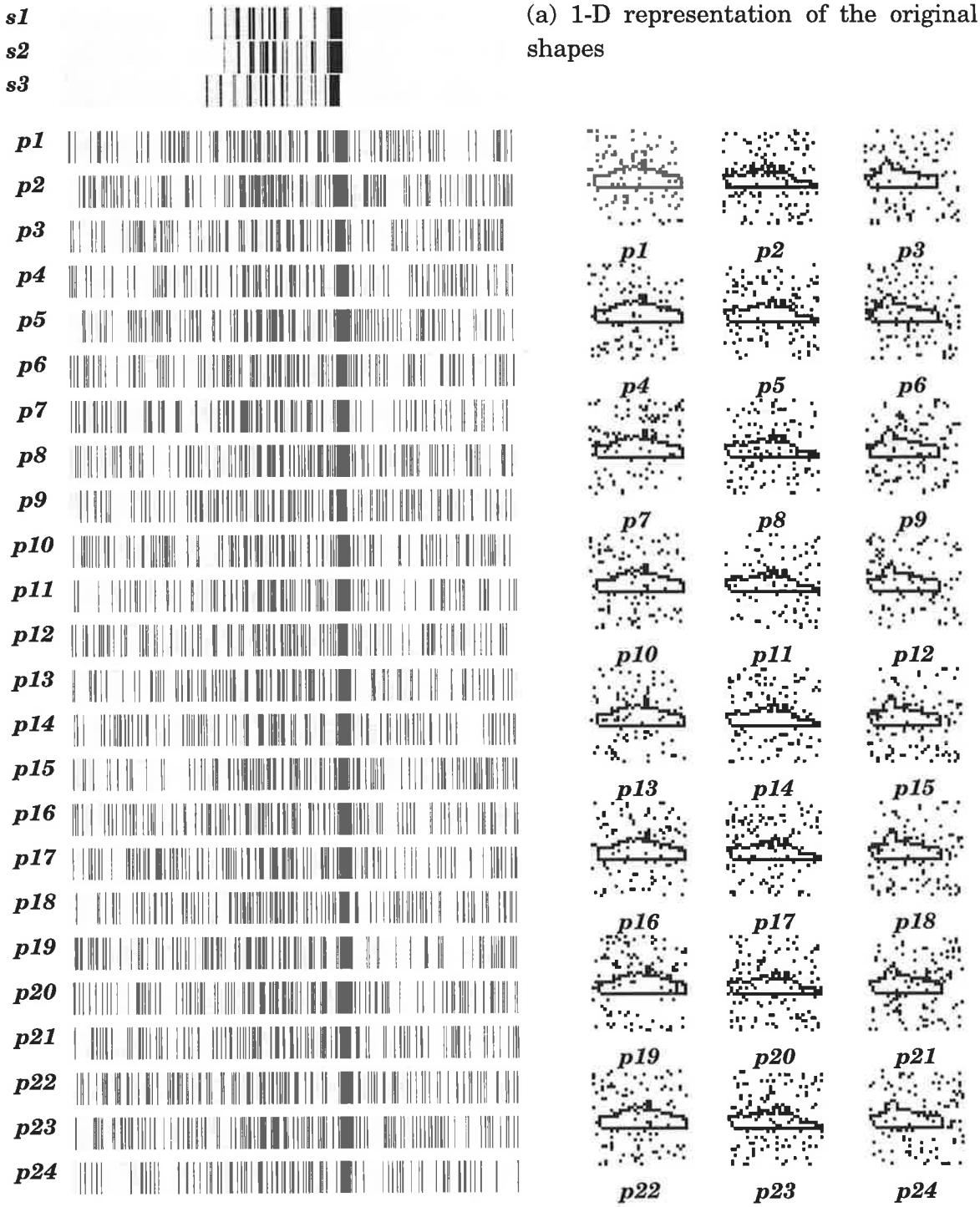


(a) Shape 1 (ship boundary 1) (b) Shape 2 (ship boundary 2) (c) Shape 3 (ship boundary 3)

(d) Common elements (shapes 1 & 2) (e) Common elements (shapes 1 & 3) (f) Common elements (shapes 2 & 3) (g) Common elements across the three shapes.

FIGURE 7.17. Binary shapes (shapes 1,2 and 3) used to generate a set of noisy inputs for the BU-DMA SAART neural network. Each input image is 32x32 elements is shown in reverse contrast. All network layers are 32x32 with the exception of Field C1, which has five cells. The three shapes are pre-aligned in the input so as to maximise their spatial overlap, as shown in (d) - (g).

The simulation data in Figure 7.19 shows that during the early phases of learning (first fifty input presentations) the memory of each LTM cell in Field C1 contains features of all three shapes. For example, cell 1 (whose memory data is shown in the left two columns) has an initially strong mixture of both shapes 1 and 2. As learning progresses and the network is presented with more examples of each noisy input, the features of shape 2 begin do decay from the LTM of cell 1. However, because of the early influence of shape 2 on the memory of cell 1 (which was finally assigned to shape 1), some features that are common between shapes 1 and 2 (e.g., the horizontal line) have a larger top-down LTM than those features that belong to shape 1 but not to shape 2.



(a) 1-D representation of the original shapes

(b) 1-D representation of the noisy inputs

(c) 2-d representation of the noisy inputs

FIGURE 7.18. Noisy inputs used to train the BU-DMA SAART neural network.

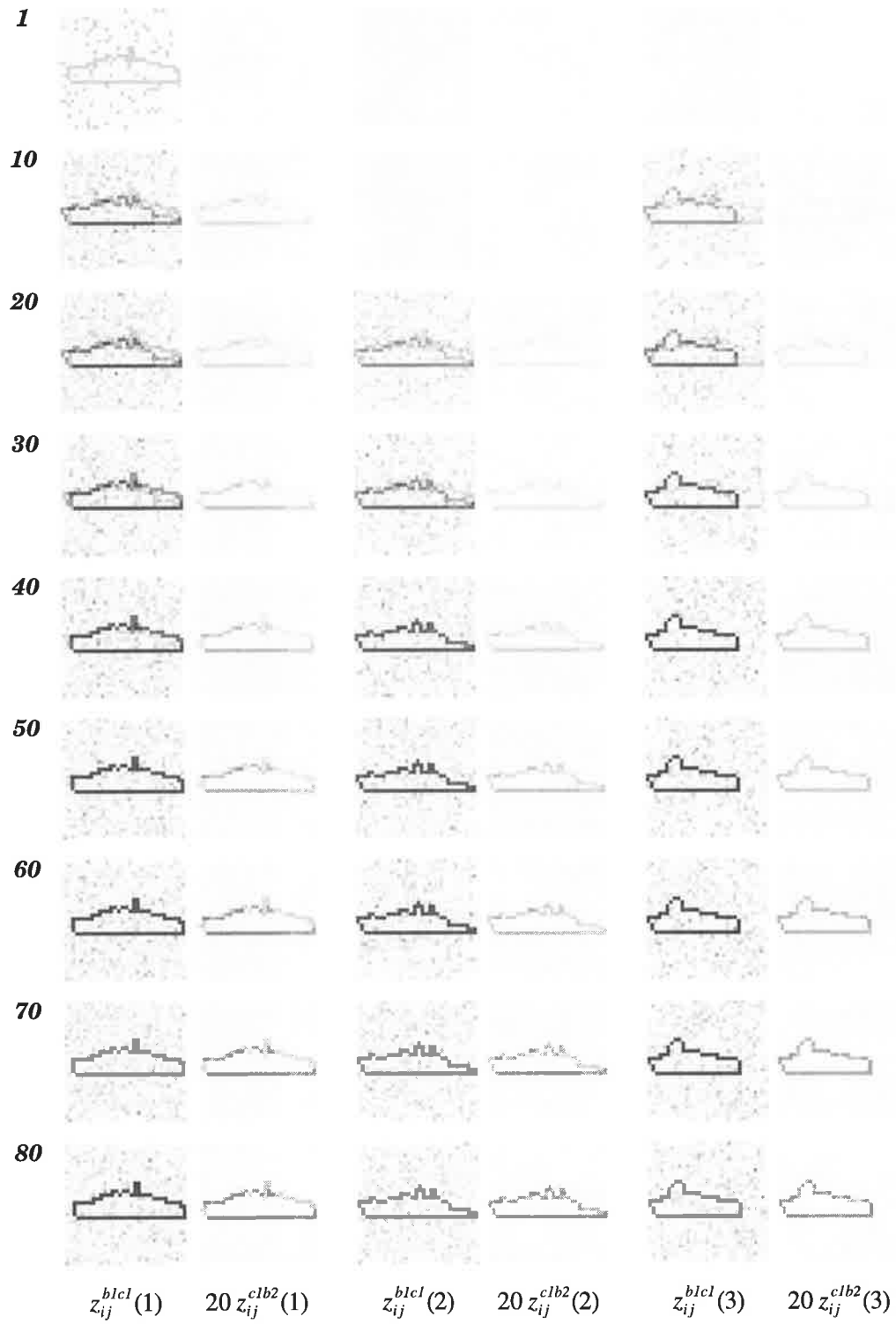


FIGURE 7.19. Bottom-up and the top-down long term memory (LTM) in the BU-DMA SAART neural network while learning in noisy but patterned input data. Shown are the bottom-up and the top-down LTMs for three cells in Field C1 whose LTM pathways were modified during learning on noisy inputs. The number in the left column indicates the input pattern since the beginning of the simulation. The memory data is shown after every 10^{th} input presentation.

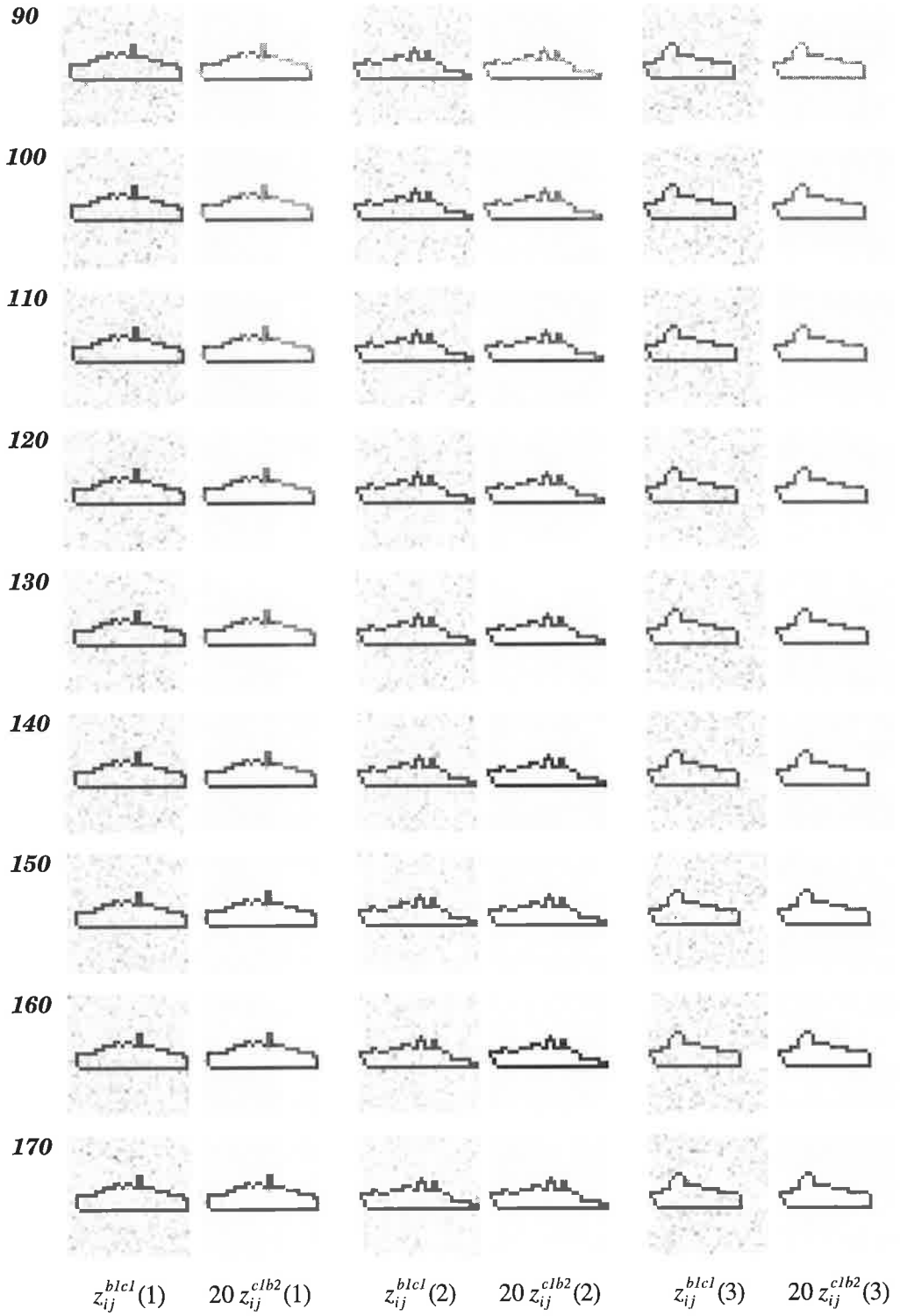


FIGURE 7.19. (Cont.) Bottom-up and the top-down long term memory (LTM) in the BU-DMA SAART neural network while learning in noisy but patterned input data. Shown are the bottom-up and the top-down LTMs for three cells in Field C1 whose LTM pathways were modified during learning on noisy inputs. The number in the left column indicates the input pattern since the beginning of the simulation. The memory data is shown after every 10^{th} input presentation.

Thus in addition to removing the irrelevant features from the top-down memory, selective memory decay in the top-down memory pathways also irons out the initial differences in the strength of the various features that are retained. As can be seen in the above data, these differences eventually become negligible. Since we have used a very low learning rate in the top-down memory pathways (= 0.02), the top-down LTM converges before it can totally suppress the non-relevant noise from Field A1. Hence the bottom-up LTM can still be modified by the noisy data. Nevertheless, this simulation demonstrates that the BU-DMA SAART neural network has powerful processing capabilities that already exceed all current self-organising artificial neural networks. This capability can be further enhanced by appropriate lower level artificial neural network architectures, such as the Boundary Contour System of Grossberg and Mingolla (1985a, b). In the next section we demonstrate the network on visual object recognition in cluttered images.

7.3.2.2 Object Recognition in Cluttered Visual Images

We now demonstrate the capability and the potential of the BU-DMA SAART neural network in real-world visual applications by providing simulation results on the recognition of 2D shapes of 3D objects in cluttered images. The images used in the simulation are generally of the type for which we currently do not have robust solutions.

Figure 7.20 shows the three target objects (toy dinosaurs) whose edges are used as the initial input to be learned by the network. The procedure for image rendering and a method for predetermining the recognition threshold are described below.

Image rendering and threshold settings

Original grey level images (8 bits, 256x256 pixels, captured by a Vidicon camera and then digitized) were preprocessed by a 3x3 Sobel edge operator to obtain object edges. The resultant edge processed images were then scaled (by a simple averaging procedure) to the size of the network layers (32x32 cells). This reduces the resolution of the target and the test images by a factor of 8. The edge processed images of the three shown target objects are initially learned by the network. The network is then tested on the cluttered visual images. Note that as a result of the competition in the network, some of the weaker edges in the edge images of the target objects did not survive and were absent in the memory, as shown in Figure 7.20. The shape of each target object was learned with a high learning rate (0.5). Note that this learning is not switched off during the test phase.

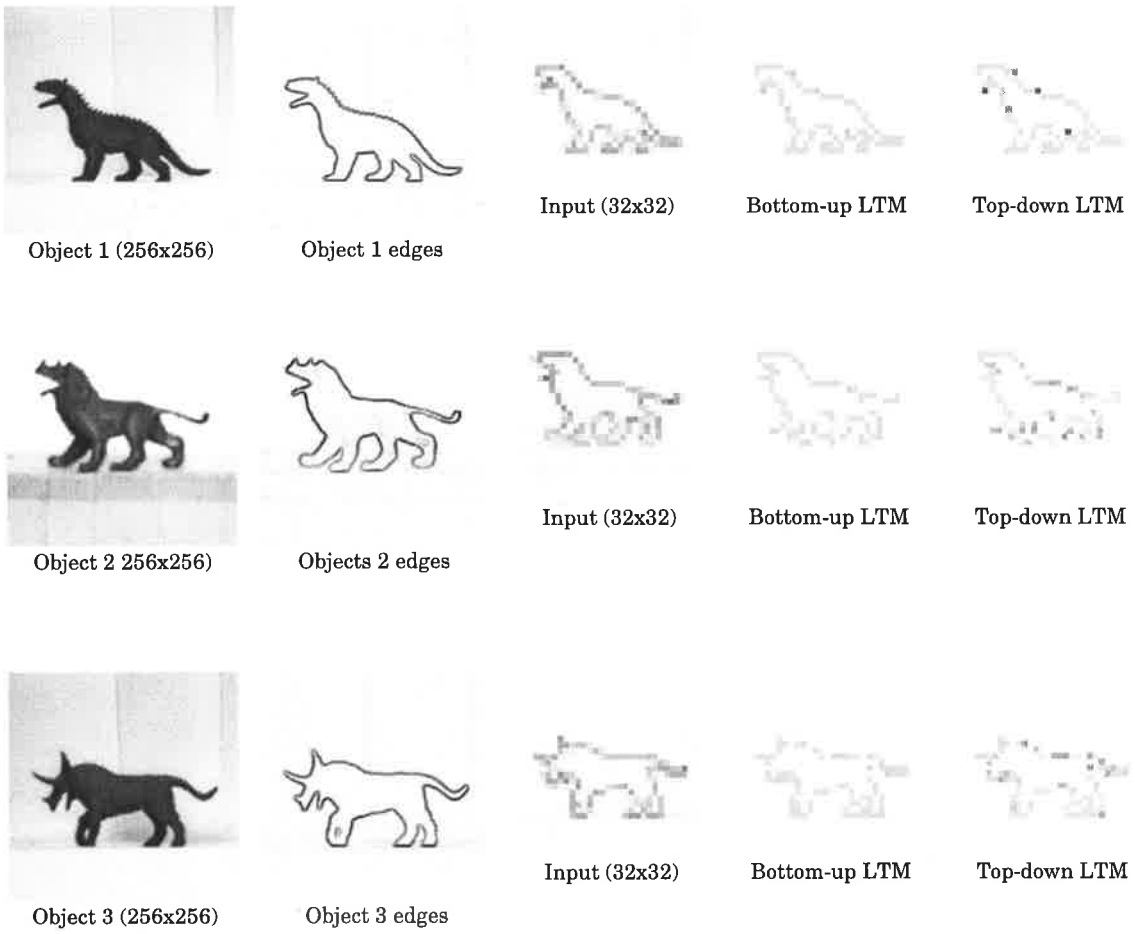


FIGURE 7.20. Objects to be learnt and recognized by the BU-DMA SAART neural network.

Recognition in the network is achieved when the match between the spatial patterns of activity across Fields B3 and A1 exceeds the pre-set threshold level of 0.90 (currently measured by the cosine of the angle between the two multidimensional vectors) and when the time-rate rate of change of the match (which is measured over four iterations of the network) is below the pre-set steady state threshold level of 0.00005 (i.e., the computational decision is taken at the steady state). The threshold for recognition is determined by first finding the highest level of the match between the three target shapes of Figure 7.20 (which was 0.851), and then setting the threshold above this value. A choice of 0.9 for the recognition threshold (or the vigilance level) was thus chosen simply on the basis of requiring a sufficiently high discriminatory power between the target shapes, while providing sufficient flexibility for the cases when a fraction of the object edges are not detected in the cluttered images. This threshold is then set and remains fixed thereafter.

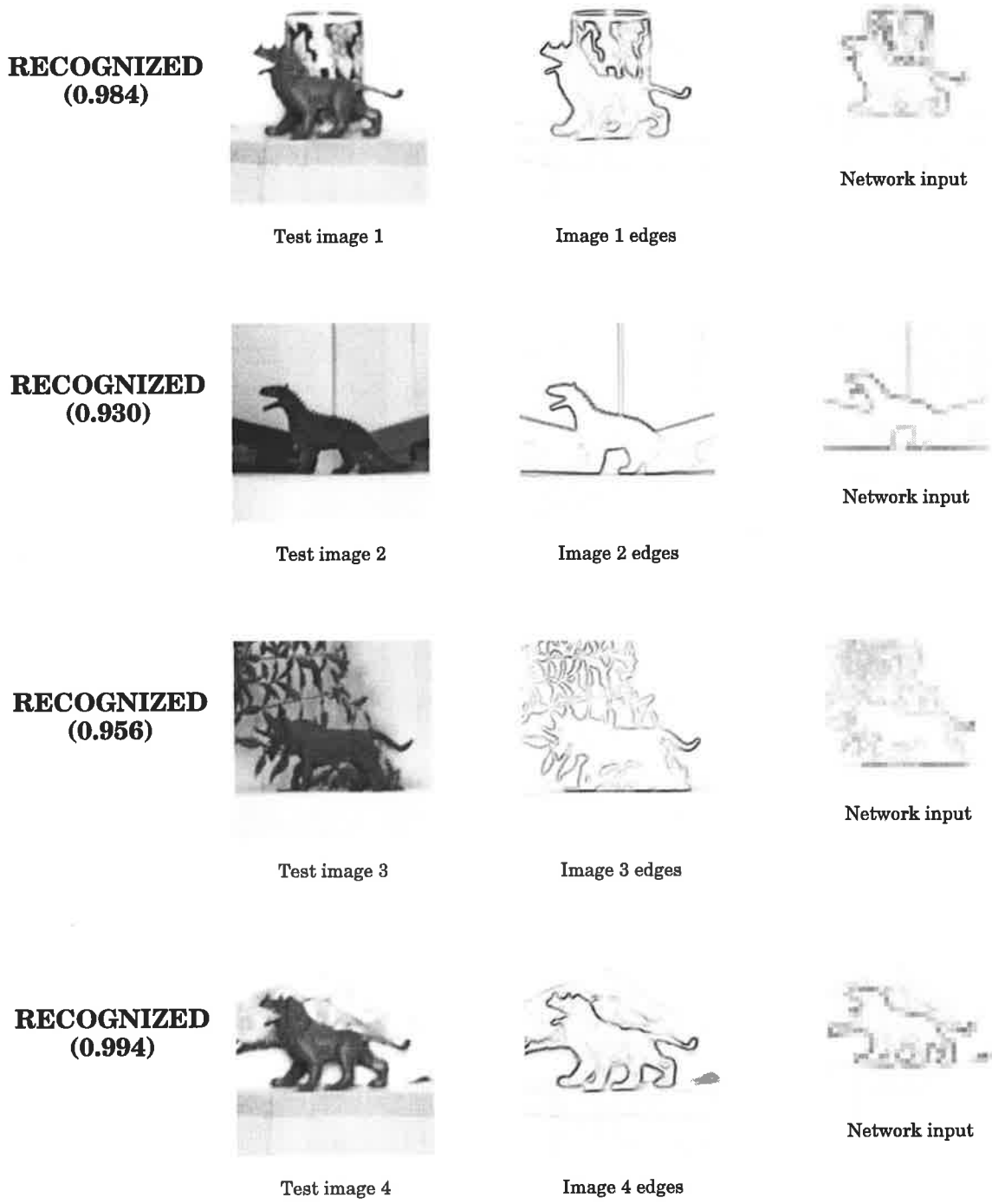


FIGURE 7.21. Cluttered visual images used to test the BU-DMA SAART neural network. In the left column we show the match level at which each individual object was recognized.

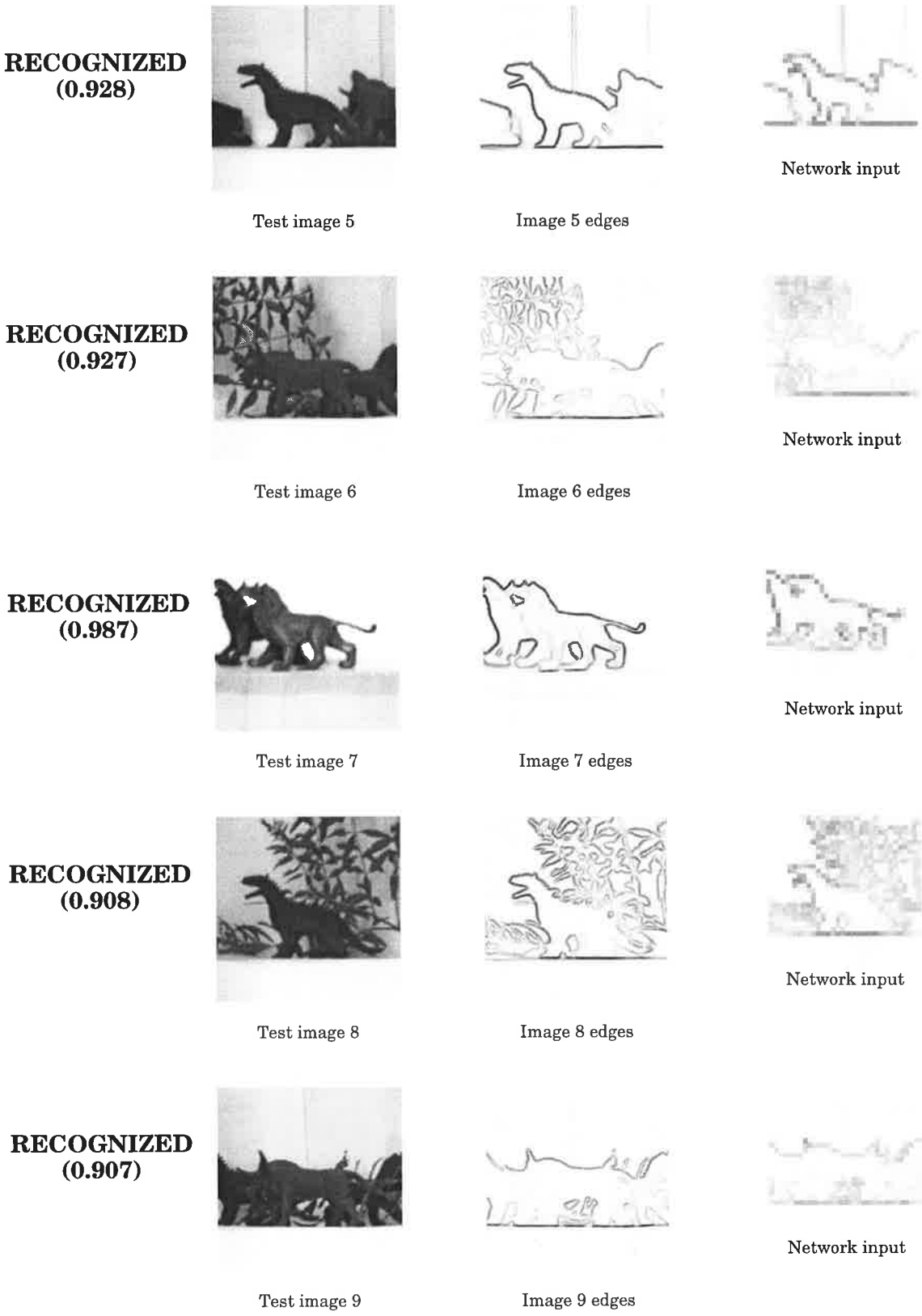


FIGURE 7.21 (Cont). Cluttered visual images used to test the BU-DMA SAART neural network. In the left column we show the match level at which each individual object was recognized.

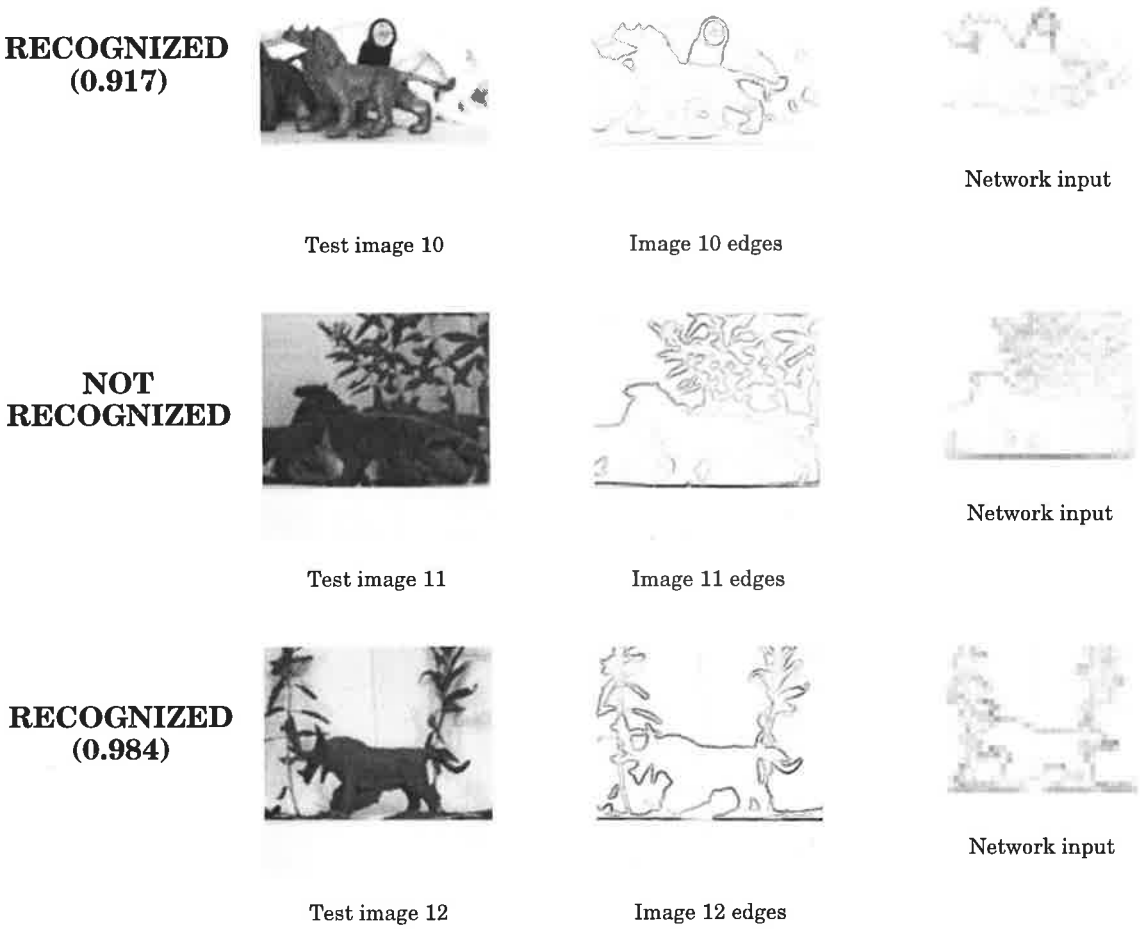


FIGURE 7.21 (Cont.). Cluttered visual images used to test the BU-DMA SAART neural network. In the left column we show the match level at which each individual object was recognized.

To summarise the above simulation, we note that only one object was not recognized (in the Test image 11). The reason for this is that a large fraction of the object's shape blends in with the cluttered background and its edges are not detected. The recognition of several objects was just above the threshold level. This is mainly due to the loss of resolution when reducing the size of the original images. However, none of the objects were misclassified. The data in Figure 7.21 indicates that had we chosen the recognition threshold to be just above 0.9 (e.g., 0.91) then the network would not have recognized two more objects. On the other hand, had the recognition threshold been close to but just above the threshold level required to discriminate the target objects in the clean background of Figure 7.20, then several objects would have been misclassified. Thus there is a tradeoff between misclassification, correct recognition and no-recognition. In Figure 7.22 we show the cellular activity of the network from the instant when the first test image (Test image 1) is presented to the network. Note

that at this instant the top-down memory of the previous object (target object 3) is still active. The cellular activity of Field C1 and the degree of match over the whole simulation run is shown in Figure 7.23.

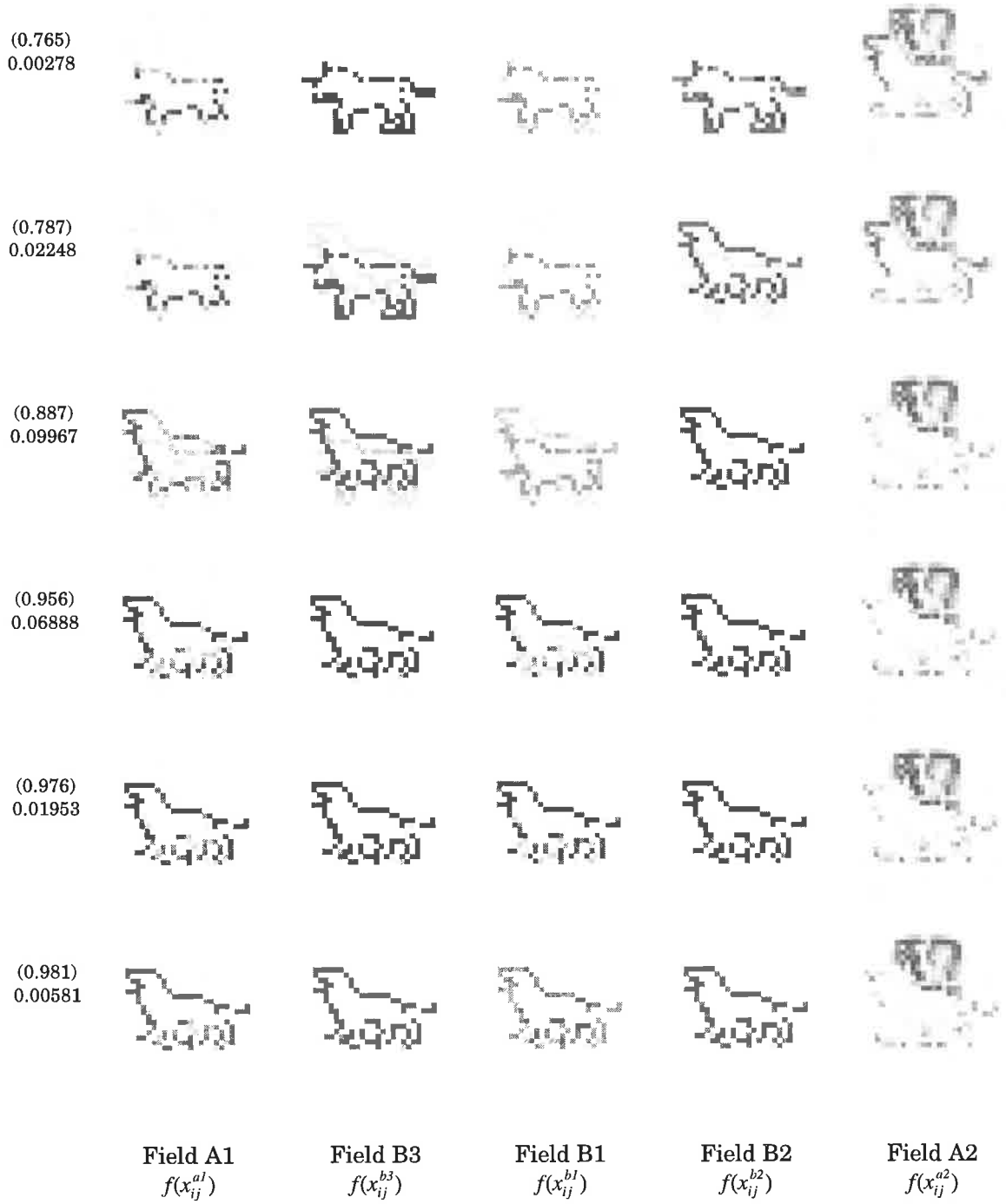


FIGURE 7.22. BU-DMA SAART neural network activity during attention switching and object recognition in a cluttered visual image. The bracketed number in the left column is the degree of the match, while the second number indicates the time rate of the match (see Figure 7.23 for more details).

Figure 7.22 shows that when the target object in the Test image 1 has activated its memory (via the DMA pathway), there is a brief period of time during which two memories are active. However, the already active memory decays soon after its LTM cell in Field C1 has been suppressed by the newly activated cell. The data shown in the far right column shows that when an object is attended (i.e., when its features appear across Field A1), then its features are removed from Field A2. The network is thus capable of quick attention switching when a familiar input is presented at the input. The graphs in Figure 7.23 show that the network was reset only 7 times. This is about half of what would have occurred had the DMA been disabled.

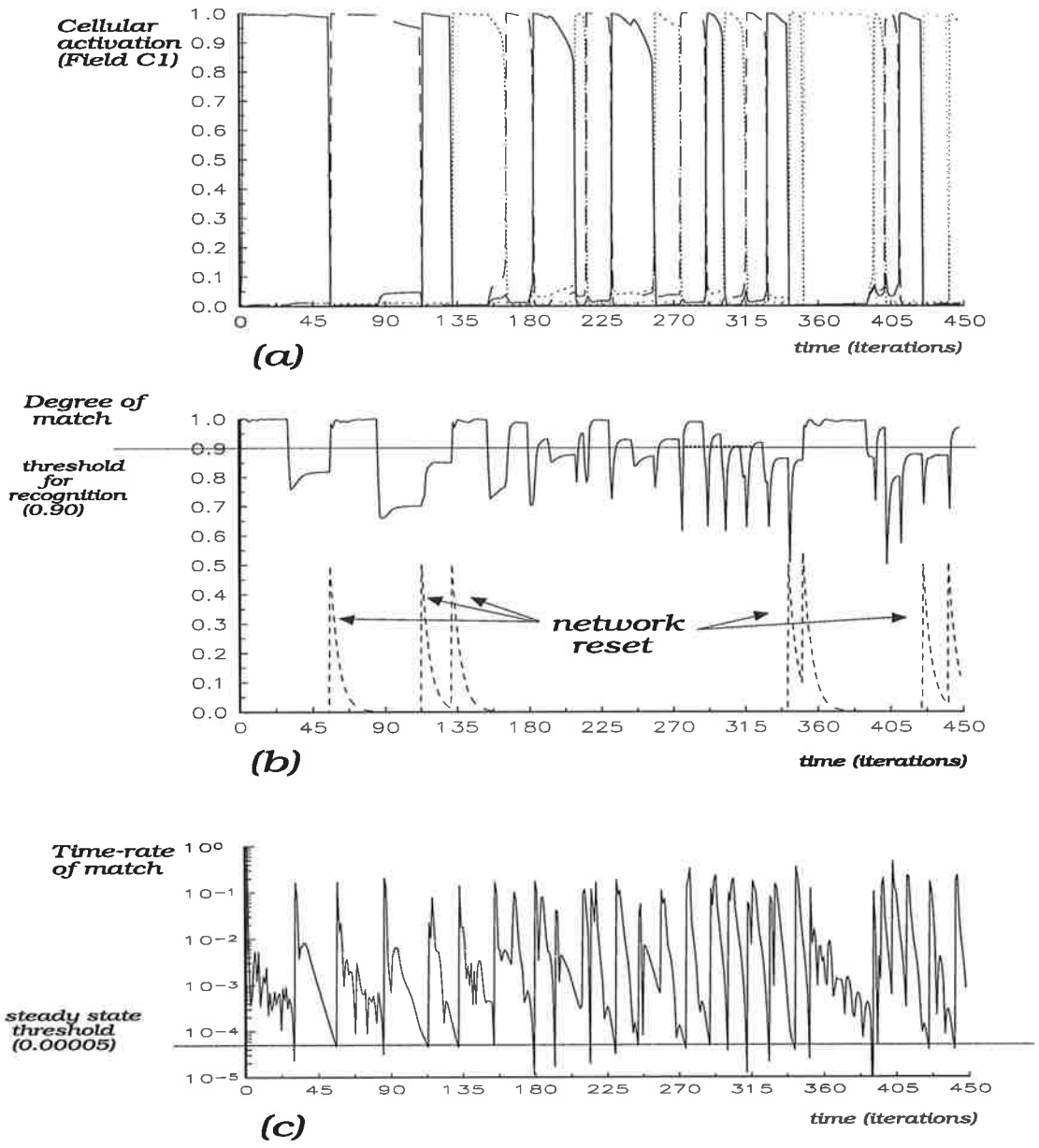


FIGURE 7.23 BU-DMA SAART neural network dynamics during real-time learning and object recognition in cluttered visual images.

Automatic attention switching without a prior memory reset is further demonstrated in the next section where we provide computer simulations of perceptual reversals on the Necker cube.

7.3.2.3 Simulation of Perceptual Reversals

The example of a visual stimulus that leads to perceptual reversals is the Necker cube, shown in Figure 7.24. Although our theory is currently aimed at 2-D visual processing, in order to simulate perceptual reversals, the cubes within the Necker cube are treated as 2-D stimuli. When viewing the Necker cube, we observe that only one of the cubes can be perceived at any one time. Because the two cubes share the exterior edges, they are always perceived (although their perceptual interpretations differ in depth). The internal edges, however, are not shared and cannot be perceived simultaneously when one or the other cubes enters perception. Thus, when one cube enters perception, the internal edges of the other cube somehow fade away. What is it that determines which set of all the possible edges should enter perception as one coherent lot? Clearly, with a large number of possibilities that can arise, there must be some mechanism that organises the edges into coherent visual precepts. Since we are quick in perceiving the alternatives, this coherence must be caused by the influence of prior visual experience, i.e., memory. If each of the cubes is independently stored in memory, then we should expect the top-down memory pathways of each stimulus to extract the inputs with which it can resonate.

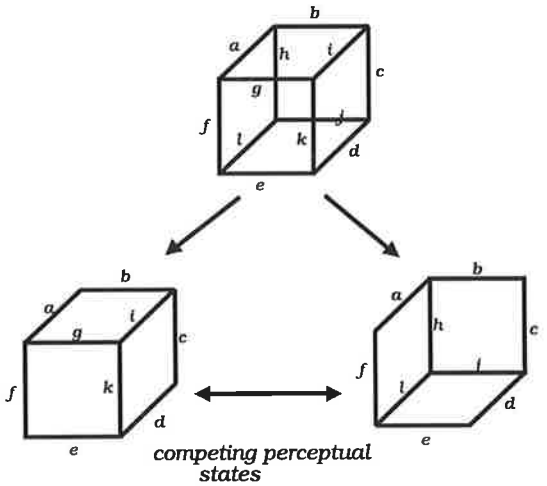


FIGURE 7.24. Necker cube and its competing perceptual states.

Since only one of the alternatives can be perceived at any one time and because perception alternates between the two, there must be a winner-take-all competition between neural cells whose bottom-up and the top-down pathways contain the visual long term memory of the stimuli. The stimulus whose memory cell has lost the

competition in one instant is eventually able to overcome its inhibition and resonate on the next cycle. This gives rise to cyclic alterations in perception of visually ambiguous stimuli.

Below we present simulation results of perceptual reversals on the Necker cube that were obtained with the BU-DMA SAART neural network. Note that initially each cube was independently presented to the network in order to establish strong top-down memory for each cube before exposing the network to the Necker cube. Because of the postsynaptic receptor desensitization, as expected, the network enters into a cyclic resonance when exposed to the Necker cube. Figure 7.25 shows the cellular activity (and the dynamics of R^{a1b3} and $| dR^{a1b3}/dt |$) for the two neurons in Field C1 whose memory pathways have encoded the individual cubes. Figure 7.26 shows the resultant levels of postsynaptic receptors, while Figure 7.27 shows the dynamics of a top-down synaptic memory trace for one of the cubes (note the dynamics of the memory decay).

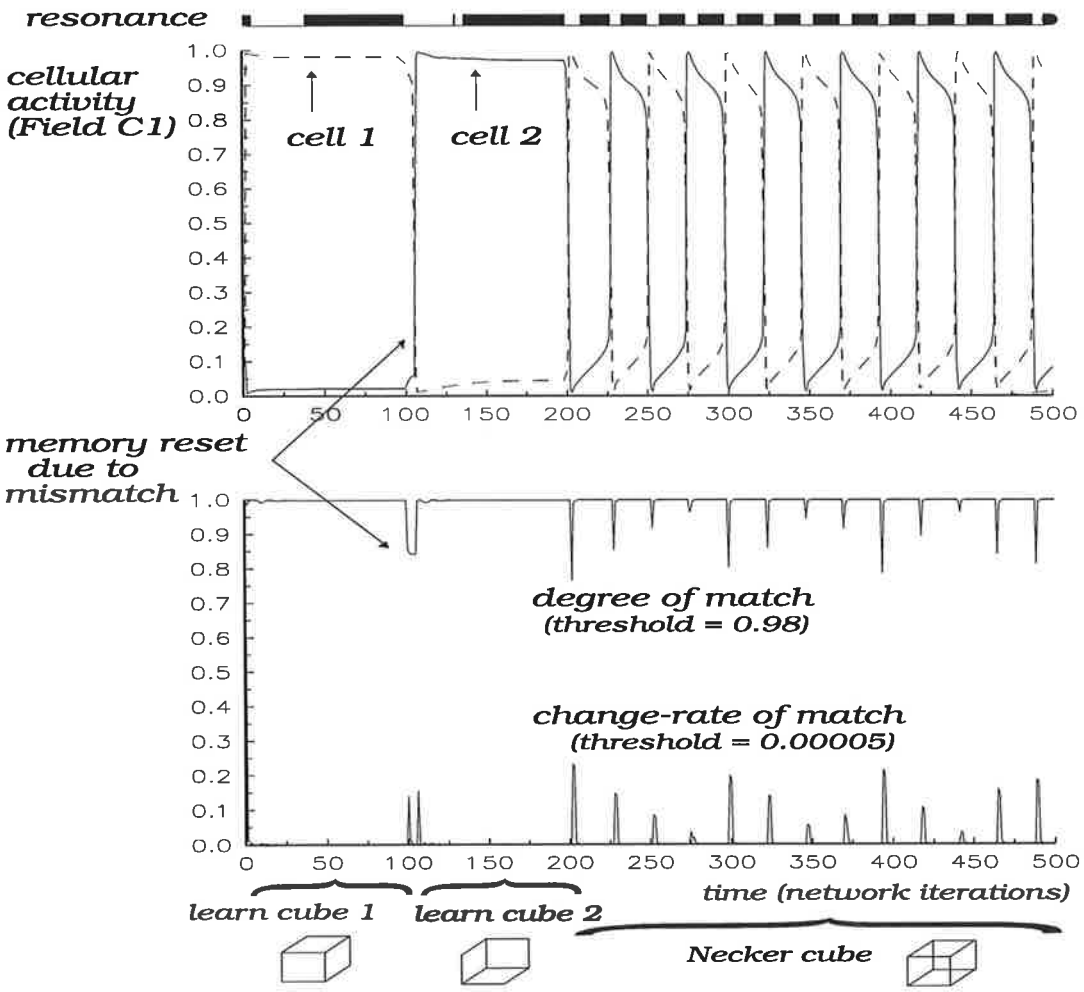


FIGURE 7.25. Cellular dynamics during learning of each cube and subsequent response to the Necker cube.

Note that in the above computer simulations we have disabled the term that sensitizes the postsynaptic receptors on Field C1 when the network is resonating (i.e., μ_{ijk}^{b1c1} in equation (6.36) was disabled). The purpose of this (although not necessary) was to allow the network to resonate with one perceptual version of the cube for several iterations.

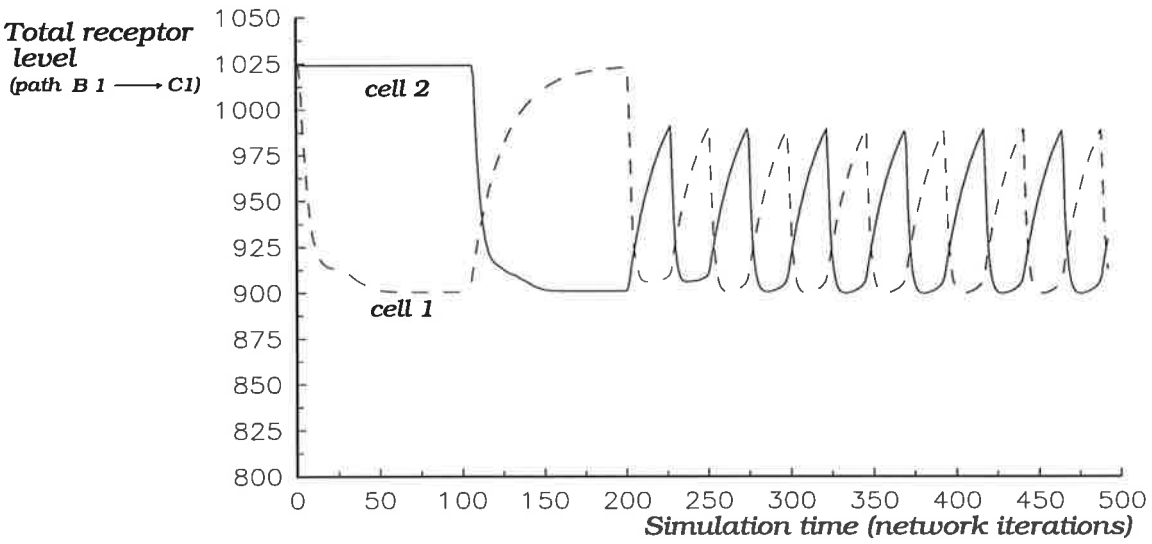


FIGURE 7.26. Postsynaptic receptor dynamics during perceptual reversal (or non-spatial attentional shifts). Shown is the total sum of all the bottom-up receptors in the $B1 \rightarrow C1$ pathway (a total of 1024 receptors) for each of the two cells of Field C1 that have encoded the memories of the individual cubes.

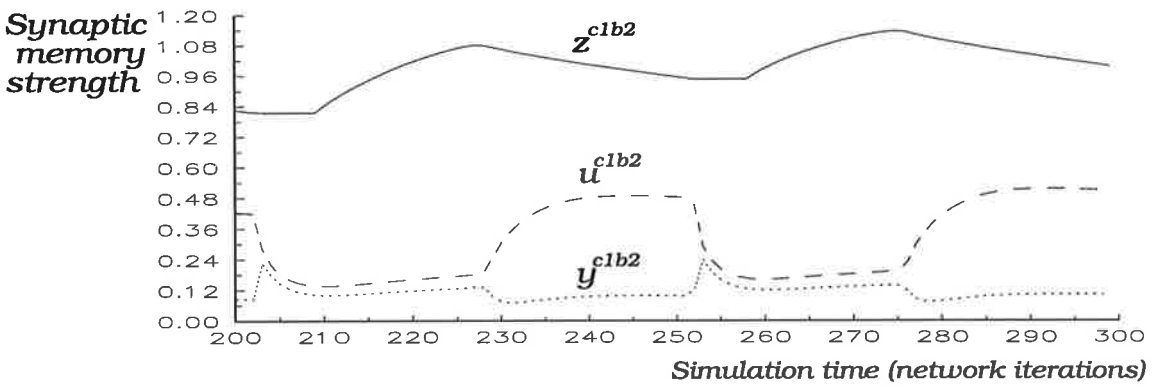


FIGURE 7.27. Top-down memory dynamics in one pathway for a cell in Field C1 whose memory represents one of the perceptual states of the Necker cube.

7.4 Conclusions

In this chapter we have described a novel real-time and self-organising artificial neural network called Selective Attention Adaptive Resonance Theory (SAART) and have provided several computer simulations that demonstrated its learning and

recognition capability. We have demonstrated that the network can 2D shapes of 3D objects in a variety of cluttered visual images, provided that most of the object's boundary is visible and that there is no spatial misalignment between the input and the relevant memory pathways. Even though the network is presently very limited in scope, it already surpasses the learning and recognition capability of all the currently known approaches on the types of images considered. The network can be easily extended to incorporate the distributed and modulated input receptive fields of Chapter 4 (which would enable the network to handle small misalignments between the input and memory that may be due to foreshortening effects) as well as the mechanism of self-regulated attentional learning of Chapter 6. However, at this stage it is of greater theoretical and practical interest to consider its extension for more general visual applications. In Chapter 8 we lay the conceptual foundations for a more advanced network and suggest how it may be generalised to translation, size and 2-D orientation invariant object recognition of object shapes.

Chapter 8

Advanced SAART Neural Network Concepts

"It has become increasingly evident in recent times, however, that nature works on a different plan. Her fundamental laws do not govern the world as it appears in our mental picture in any very direct way, but instead they control a substratum of which we cannot form a mental picture without introducing irrelevancies. The formulation of these laws requires the use of the mathematics of transformation. The important things in the world appear as the invariants (or more generally the quantities with simple transformation properties) of these transformations. The things we are immediately aware of are the relations of these nearly invariants to a certain frame of reference, usually one chosen so as to introduce special simplifying features which are unimportant from the point of view of general theory".

P.A.M. Dirac (1935)

8.1 Introduction and Overview

In this chapter we integrate the various mechanisms, design principles, circuits and networks of the previous three chapters and propose concepts for the Advanced Selective Attention Adaptive Resonance Theory (ASAART) neural network. ASAART addresses the concepts of self-organised real-time learning, memory guided search in cluttered backgrounds, perceptual constancy (ability to recognize an input stimulus despite its variation in input size, orientation or position). We first integrate the SAART neural network of Chapter 7 with the advanced neural model of visual spatial attention of Chapter 6 (sections 6.4.1 and 6.4.2) to provide a neural network

that is capable of self-organised real-time learning, translation invariant 2-D pattern and shape recognition, and memory guided search in cluttered inputs. We then consider how the concepts of the previous chapters may be used to extend the network to cater for size and orientation invariant recognition.

8.2 Advanced SAART-1 Neural Network

In Chapter 6 (section 6.4.2) we have proposed a minimum 2-D neural circuit that stores the Immediate Short Term Memory (ISTM) of the attended spatial pattern in an excitatory reverberatory loop between two competitive neural Fields of transmitter gated neurons. We have also discussed how the reverberating memory within the circuit can be used to presynaptically facilitate and hence influence the bottom-up attentional selection, thus enabling top-down memory guided selective search and recognition of 2-D spatial patterns and shapes in a wide bottom-up input field. Once the target pattern is found and selected, it is then dynamically routed to and centred in a higher neural layer such that it is transitionally invariant. We now show how that circuit can be integrated with the BU-DMA SAART neural network of Chapter 7 (section 7.3) to provide a self-organised and real-time learning artificial neural network that is capable of translation invariant 2-D pattern (shape) recognition and memory guided search in a wide and cluttered input field. The proposed neural network, called ASAART-1, is shown in Figure 8.1.

Note that the excitatory reverberations between Fields B2 and B3 maintains the ISTM of the most recent input after it is offset. Since Field B3 (as well as Field A0) backprojects a specific 2-D spatial pattern (i.e., a Facilitatory Modulating Field (FMF)) to each cell in the bottom-up attention selection layer (Field D1), the reverberatory memory can thus search for and find the same 2-D spatial pattern in the input field (Field D0) when its is embedded in the cluttered background, regardless of its location. When found, the centroid of the selected 2-D spatial pattern will be aligned (i.e., centred) with the translation invariant neural Field A0 (which then feeds the BU-DMA SAART neural network).

Since the active cells in Field D1 facilitate the synaptic pathways of Field A0 to achieve translation invariant 2-D representation of the selected stimulus, the network will be able to find and recognize a 2-D pattern wherever it appears. If the pattern had been previously stored in the reverberatory loop between Fields B3 and B2 (or in the SAART's LTM pathways), then this established memory will on subsequent input presentation influence the bottom-up visual spatial attention layer, thus enabling the network to recognize the remembered pattern or object when if it is embedded in cluttered visual images.

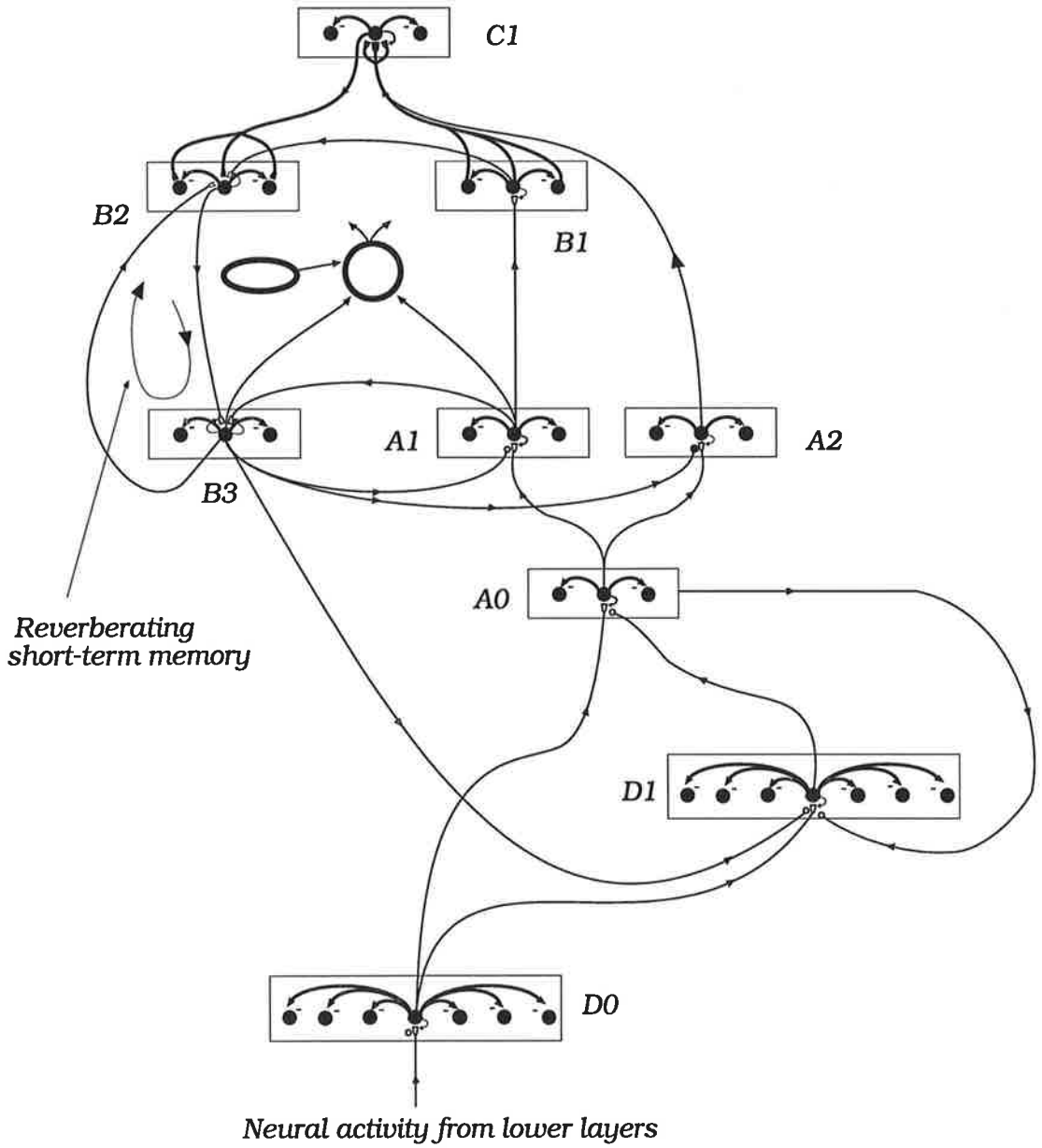


FIGURE 8.1 Advanced Selective Attention Adaptive Resonance Theory (ASAART-1) neural network for translation invariant pattern recognition and memory guided search in cluttered inputs. Each cell in the bottom-up attentional selection layer (Field D1) samples its excitatory inputs from a local cluster of cells in the input array (Field D0). The same cluster of cells from D0 also project excitatory synaptic pathways to Field A0, such that the central neuron within the cluster is centred in A0 (see Figures 6.27-6.33, Chapter 6). The input receptive fields of neurons in A0 are spatially modulated by neurons in D1. Top-down pattern specific facilitatory feedback from SAART's Field B3 to D1 constitutes a Facilitatory Modulating Field (FMF) that facilitates the input receptive fields of D1 (thus enabling a previously established reverberatory memory as well as SAART's long term memory to influence the bottom-up attentional selection).

The following equation describes the transmitter mobilization in the $(i,j)^{th}$ synapse of the $(m,n)^{th}$ neuron in the bottom-up visual spatial attention neural layer (Field D1).

Mobilized transmitter (y_{ijmn}^{D0D1})

$$\frac{dy_{ijmn}^{D0D1}}{dt} = [\beta_y^{D0D1} + H_y^{A0D1} f(x_{ij}^{A0}) + H_y^{B3D1} f(x_{ij}^{B3})](u_{ijmn}^{D0D1} - y_{ijmn}^{D0D1})$$

$$- f(x_{ij}^{D0})[\rho_y^{D0D1} + K_y^{D0D1} f(x_{mn}^{D1})][y_{ijmn}^{D0D1} - Y^{D0D1}]^+ - \gamma_y^{D0D1} y_{ijmn}^{D0D1} \tag{8.1}$$

where H_y^{A0D1} is the facilitation gain for the signals from Field A0, while H_y^{B3D1} is the facilitation gain for the signals from Field B3. If the tonic level of transmitter mobilization in these synapses is much smaller than the two facilitatory gains and if $H_y^{A0D1} = H_y^{B3D1}$, then the reverberatory ISTM across Field B3 will strongly prime all cells in Field D1. Thus when a new input scene (edges of a scene) are presented at the input Field D0, then these priming signals will enhance the activation of cells in Field D1 but only in those locations where the corresponding spatial patterns appears in the input. This enhancement is at the expense of all the other inputs.

Below we provide a computer simulation of the network to demonstrate its self-organised learning, translation invariant 2-D shape recognition and memory guided search and selection for a simple example. Because of the extremely large number of synaptic pathways between the input field (D0) and the visual spatial attention neural layer (Field D1), we are only able to provide a simulation for small input arrays. In order to relate the property of the above neural network to the recent neurobiological data, we provide a computer simulation of the visual problem that was experimentally performed by Chelazzi *et al.,* (1993) with a trained monkey. We thus first store the recognition memories of two simple shapes (a triangle and a square) by training the network on the individual shapes (triangle first and then the square). Following this period of training, we present the network with a blank input field for a period of time. During this period the cellular activity of Field C1 decays, but the ISTM of the most recent shape (the square) is retained in the reverberatory loop between Field B2 and B3. After this brief period, we re-introduce the triangle, the purpose of which is to hold the triangle in the ISTM. This is then followed by a blank period during which the triangle is being remembered in ISTM. We then introduce both shapes, but at different spatial locations. Since the triangle is also in the ISTM, we expect that the ISTM will facilitate the visual spatial attention layer to find and transmit the triangle, thus leading to its recognition.

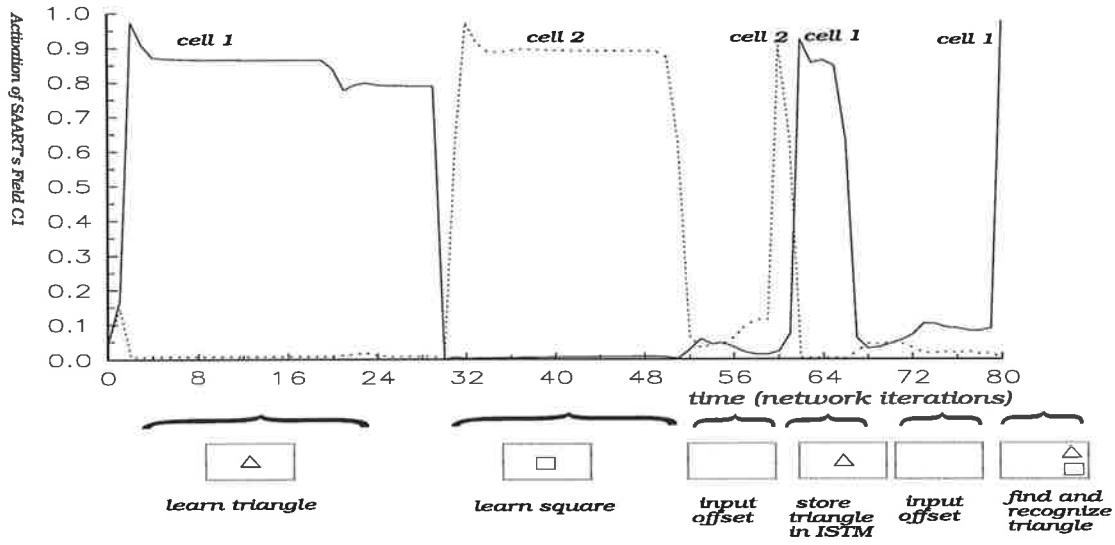


FIGURE 8.2. Simulation of Chelazzi *et al.*, (1993) experiment with the ASAART-1 neural network.

The simulation data in Figure 8.2 represents the activation level of two cells in the ASAART-1's Field C1. The data shows that when both shapes are presented to the network simultaneously, the long term memory cell of the triangle (cell 1) was activated before the other shape had time to enter into the network and activate its memory cell. This simple simulation thus demonstrates that the network is capable of translation invariant 2-D shape recognition and memory guided search in cluttered inputs. However, this simulation does not demonstrate the network on realistic visual images. Hence there is a need to simulate the network on a much larger size and with realistic inputs. Our simulation was restricted to the small size primarily because of the limitations of the computing platform on which the simulation was run.

8.3 Concepts for Size Invariant Recognition

In order to develop the neural concepts for size invariant 2-D object recognition, let us consider how the Feedforward Excitation-Feedback Presynaptic Facilitation neural circuit of Chapter 4 may recognize a stimulus when it is changed in size. Thus we initially assume that the circuit has an established top-down reference shape.

Let us suppose that the bottom-up input from Field F_0 is sampled by a number of *Size Fields* ($F_{s1}, F_{s2}F_{s3}, \ldots, F_{si}, \ldots, F_{sn}$), such that each Size Field F_{si} samples its inputs at i^{th} spatial scale. If the Size Field F_{s1} samples at the highest spatial resolution, then F_{sn} samples at the coarsest spatial resolution. In addition to there being a change in the sampling resolution of the bottom-up input as one goes from Size Field $F_{s,i}$ to $F_{s,i+1}$, there is also a gradual shift in the position of the input receptive fields (IRFs) away from the

array centroid. Hence, the highest resolution Size Field ($F_{s,1}$) will receive its information from a small input region, while the Size Field of the coarsest resolution will not receive any information from a central region.

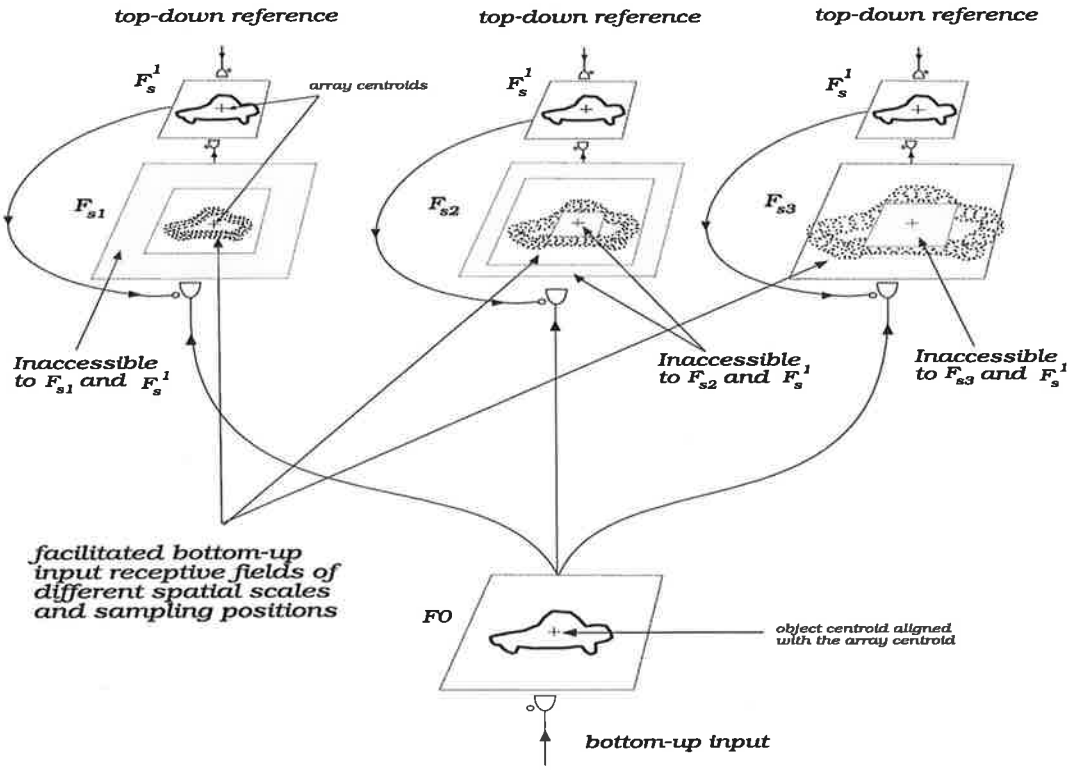


FIGURE 8.3. Bottom-up sampling strategy for achieving size invariant recognition.

Thus rather than having coincident input receptive fields, each successively coarser Size Field will have its input receptive field displaced outwards relative to the input receptive fields of smaller spatial scales. For example, Field F_{s3} will have input receptive fields whose spatial extent is larger than those of Field F_{s3} (i.e., $\| IRF_{s3} |> \| IRF_{s2} \|> \| IRF_{s1} |$).

Now let us assume that each Size Field projects its output to a common layer, as in Figure 8.4, that may also receive a top-down memory. The signal flow from each of the Size Fields needs to be regulated by layer of competitive match/mismatch neurons. The neuron that detects the highest match will facilitate the signal transmission gain from its associated Size Field.

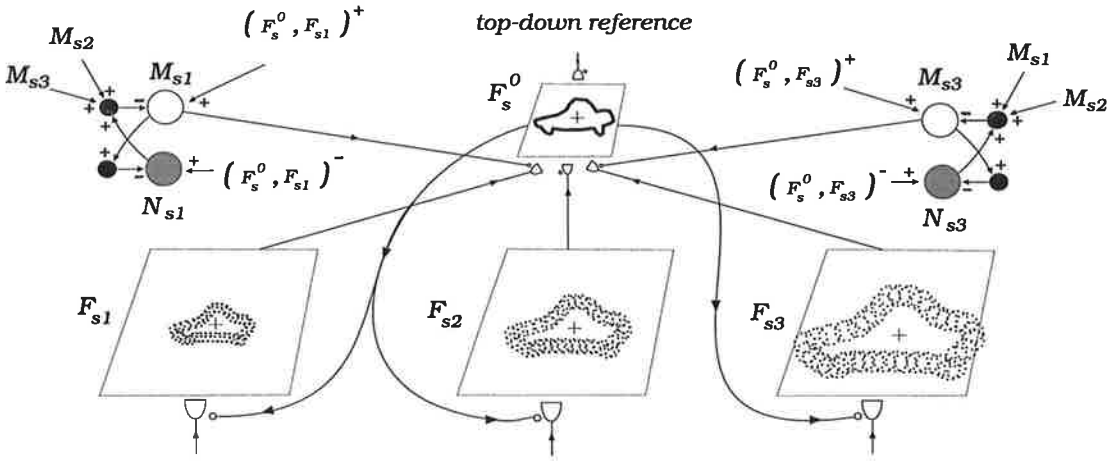


FIGURE 8.4. A competitive layer of match/mismatch neurons regulates the selection of the best matching bottom-up Size Field.

However, if there is no top-down input, then a bias is needed to initially select a Size Field for attention. We now need to consider how a particular object memory may be activated by a bottom-up input that does not match the spatial size of the stored memory. We suggest that a complimentary scheme to that shown in Figure 8.4 is required. The same match/mismatch neurons may be used to regulate the top-down and the bottom-up signal flow. Thus, in addition to the parallel Size Fields in the bottom-up direction, there is also a need for similar Size Fields in the top-down direction so that the recalled memory may be appropriately scaled to be matched with the input, as well as to regulate the bottom-up signal transmission gain.

8.4 Concepts for Orientation Invariant Recognition

In order to cater for orientation invariant recognition we suggest that a similar scheme as shown in Figure 8.4 may be used. Thus, there is a need for parallel and competing neural layers in the bottom-up and the top-down pathways, such that each succesive layer produces a slightly different oriented version of their input pattern. A competitive match/mismatch layer may be used to decide how the information is to be routed in the bottom-up and the top-down directions in order to match the top-down memory. In addition, there ought be a bias for a certain spatial orientation (e.g., vertical).

Below we describe the implementation of a massively parallel real-time artificial neural network for rotation invariant recognition of 2D shapes in visual clutter (its relationship to the psychophysical data is discussed in section D.4, Appendix D). The network shown in Figure 8.5 is based on the concepts of parallel and competing frames of reference in the bottom-up and the top-down directions and is called Rotation Invariant Selective Attention Adaptive Resonance Theory (RI-SAART) neural network.

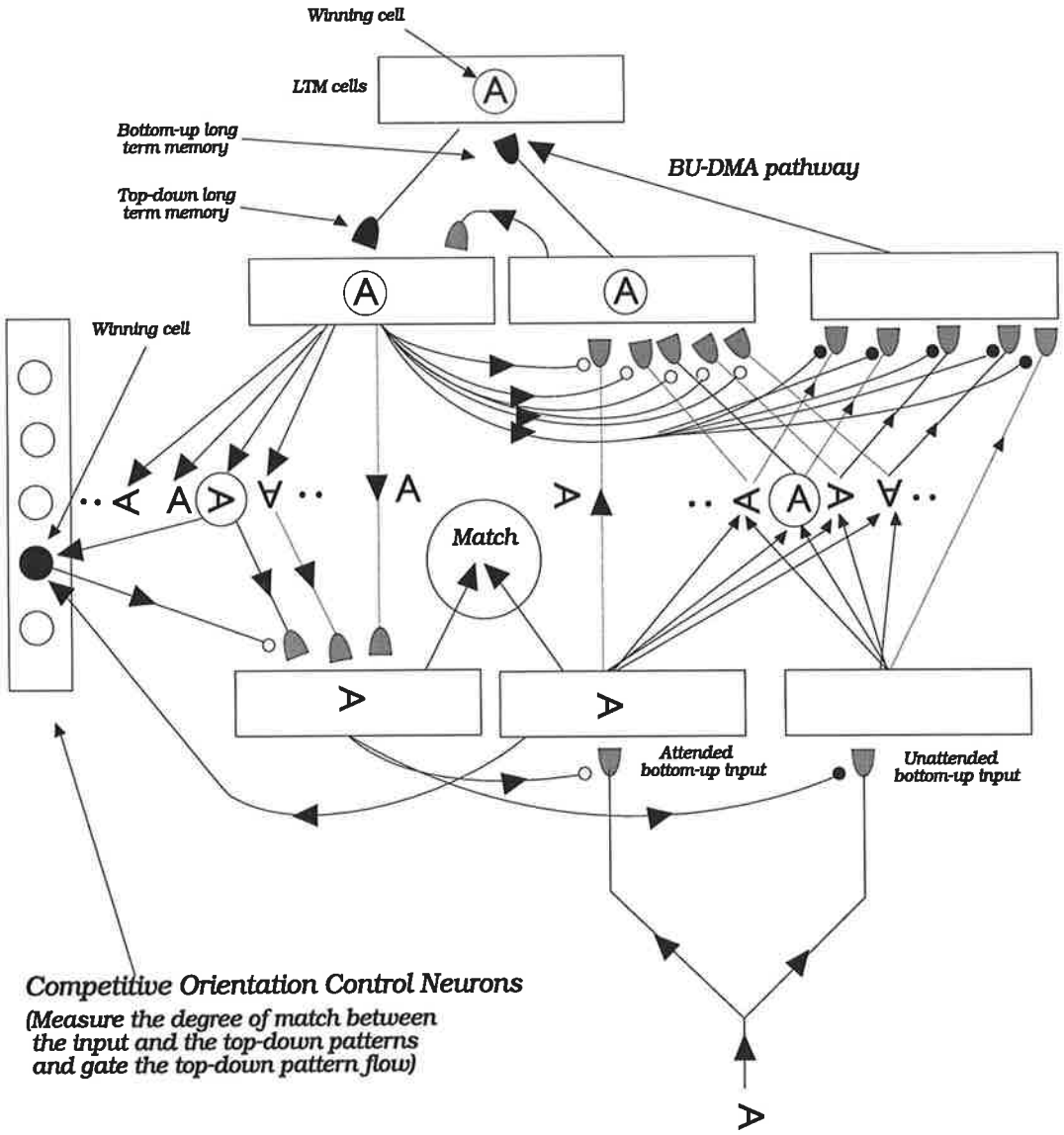


FIGURE 8.5. Rotation Invariant Selective Attention Adaptive Resonance Theory (RI-SAART) neural network.

Let us assume that a certain 2D shape (e.g., upper case character 'A') is stored in the LTM pathways such that the stored representation is that of a shape in its usual orientation (e.g., upright). Let us now consider how the above network would recognize a rotated shape. Thus if character 'A' is subsequently presented in a clockwise orientation as shown in Figure 8.5, then its bottom-up memory will be activated via the BU-DMA pathways by one of the bottom-up orientation layers whose output matches the stores BU memory pathways. This will then cause the activation of the top-down memory and its transfer into a large number of top-down orientation layers. Orientation control neurons measure the degree of match between the actual input and the top-down rotated versions of the recalled shape. The winner will further enhance the top-down pattern whose orientation best matches the input while all other top-down versions are

suppressed. This leads to a steady state where only one top-down oriented version ends in SAART Field B3. Similarly, only one oriented bottom-up version will be active across SAART's Field B2 (preventing the BU LTM to be corrupted by other orientations). Interestingly enough, the two versions have opposite orientation. The top-down version matches the orientation of the input, but the bottom-up version matches the orientation of the BU memory.

8.5 Conclusions

In this chapter we have suggested how the neural mechanisms of synaptic facilitation and the match/mismatch detection may be used in more advanced neural network designs. We have proposed an extension to the SAART neural network for translation invariant 2-D pattern recognition. We have also suggested that one approach to achieving position, size and 2-D orientation invariant object recognition is to use a large number of synaptic pathways and layers through which the bottom-up (as well as the top-down) signals may be simultaneously routed. The information flow along the various pathways is regulated by a layer of match/mismatch neurons. However, because of the large number of synaptic pathways that are required for real-world visual object recognition problems, the suggested solution is currently not practical. The implication of the proposed solution to rotation invariant 2D shape recognition is further discussed in Appendix D (section D.4) in relation to the psychophysical data from human subjects.

Chapter 9

Conclusions and Recommendations

"The foundations of science as a whole, and of physics in particular, await their next great elucidations from the side of biology, and especially from the analysis of the sensations ... psychological observation on the one side and the physical observation on the other may make such progress that they will ultimately come into contact, and that in this way new facts may be brought to light. The result of this investigation will not be a dualism but rather a science which, embracing both the organic and the inorganic, shall interpret the facts that are common to the two departments."

E. Mach (1914)

9.1 Introduction and Overview

In this chapter we summarise the main contribution that the thesis offers to the understanding of primate visual system. Although the theory presented in the thesis was developed within the context of higher cortical layers, particularly the infero-temporal cortex, the concepts are not necessarily restricted to IT cortex nor are they necessarily restricted to the visual modality.

By assuming that the neural layers in the primate visual brain may be modelled by real-time 2-D competitive artificial neural layers, we have provided a new approach to solving some of the most difficult problems in machine vision. Starting with a simple theoretical concept of top-down modulatory feedback that we have derived from Grossberg's Adaptive Resonance Theory, we have proposed novel neural layers, mechanisms and design principles for the design of complex 2-D neural circuits and networks. We have demonstrated our neural circuits and networks on difficult problems of object recognition in cluttered visual images and have shown how the mechanism of

top-down and memory guided selective attention is implemented. We have also shown how the arousal of a real-time neural system may be self-regulated by the familiarity/novelty detection and how this self-regulation may alter the degree of attention, thus altering the selectivity of cellular receptive fields.

9.2 Neurobiological Implications

The most important contribution that the theory offers to the understanding of neurobiology is the role of the massive feedback pathways. The theory proposes that, in the context of neurobiology of vision, top-down feedback pathways form a closed loop and do so by selective modulation of bottom-up neural signal transmission. The theory predicts that these modulatory signals act throughout the brain and affect the receptive fields of cortical and subcortical neurons. For example, the theory predicts that the receptive field of V4 neurons will be dramatically influenced when the animal pays attention to a visual stimulus in one part of the visual field while ignoring another, (Moran & Desimone, 1985). This is achieved by the amplification of signals from the attended source, while the unattended signals are annihilated by competitive inhibition of their target cells. In an inattentive state, each neural signal in the receptive field of the cell will be weak but their total contribution may be significant to excite the cell above its firing threshold. Because of the effects of top-down selective attention on the properties of receptive fields, we stress that one needs to be cautious when interpreting the experimental data that is obtained from anaesthetised animals. An anaesthetized monkey, for example, is unlikely to be able to exert top-down modulation of the bottom-up synapses and hence the neural response may solely be due to the simultaneous co-activation of many feedforward pathways.

The bottom-up/top-down interactions as proposed in the thesis imply that at any one instant, the vast majority of the synaptic pathways in an alert animal are not accessed due to competitive effects between them. For example, translation invariant pattern recognition can be achieved if each IT neuron has access to pathways from the whole visual field, but when selective attention is engaged, only a contiguous set of pathways from the attended area transmit their bottom-up information. Our prediction is that there is a set of complementary spatial transformations in the bottom-up and the top-down directions. Spatial transformations in the bottom-up direction are required so that a stored memory may be activated, while the top-down spatial transformations (of opposite polarity) are needed so that an appropriate top-down feedback may be matched with the input and to presynaptically facilitate the appropriate bottom-up pathways. This may explain for the large size of the inferior temporal cortex (which occupies some 7% of the monkey's neocortex) and the anatomic finding of several distinct foci in IT.

The theory has also suggested that some top-down pathways (especially those that involve memory traces) may be equated to visual imagination or visual imagery. This implies that conscious visual awareness of external stimuli relies on the integrity of the cognitive visual layers. Thus, patients whose top-down attentional pathways are lesioned will not be able to perceive a familiar object in the presence of non-relevant distractors. A small disruption to either pathway will lead to unusual effects that the subject may not be aware of, such as visual neglect. Since the top-down pathways act to force the neural system to seek a stable resonant state, their disruption will cause distortion to visual perception, such as in a patient who claims to see distorted faces (Bodamer, 1947). The equivalence between the top-down feedback and mental imagery has also been recently proposed by several researchers (Le Bihan *et al.,* 1992; Kosslyn *et al.,* 1993; Mumford 1994).

In Chapter 6 (section 6.2) we have presented a neural circuit which implies that the matching between the sensory input and the recalled top-down memory is carried out by two types of neurons. One type responds vigourously to novel stimuli while the other type responds vigourously to familiar stimuli. Since the response of these neurons may be needed to influence the signal transmission at the earliest level of the visual system, it is appropriate that these neurons be located in the part of the brain where they can also interact with similar neurons of other sensory modalities (e.g., auditory). The appropriate location for these neurons would be the thalamus. This proposal agrees with the experimental findings of a group of neurons in the medial thalamus whose properties seem to correlate well with our proposal (Rolls *et al.,* 1982; Fahy *et al.,* 1993a, 1993b).

Although the theory was developed within the context of higher level primate vision (beyond visual area V4), we propose that similar modulatory mechanisms as proposed in the thesis also exist in the lower level visual hierarchy (LGN, V1, V2 and V4). Since IT cortex does not appear to be in the retinotopic register with the visual area V4, we suggest that the observed modulatory effects on V4 cells are due to the top-down facilitatory presynaptic feedback from elsewhere, most probably the prefrontal cortex. However, since the prefrontal cortex interacts with the IT cortex, it is possible that the origin of the modulatory signals on V4 cells is from the IT cortex but are relayed via the prefrontal cortex. If that is the case, then it is tempting to suggest that, the prefrontal cortex contains cognitive neural layers where visual memories are temporarily stored, and that these layers are also used for visual cognition and imagination.

We have assumed that the bottom-up neural signals representing the edges of a visual scene and the boundaries of objects are represented in one layer. However, it is known that the early visual layers (V1, V2 and V4) are orientation selective. Hence

there is a need to extend the theory to take this into account. It should be stressed that our model has assumed distance independent competitive interactions right across the whole neural layer. Therefore, if the model is to be extended downwards (i.e., towards area V4 and below), the lateral extent of interactions should begin do decrease, such that they are restricted to small neighbourhoods of cells whose receptive fields overlap. Nevertheless we have presented a neural theory that should point us in the right direction. Perhaps the most satisfying result of our theoretical model to date is that we can now satisfactorily explain and model a recent neurobiological experiment demonstrating the memory guided visual search (Chelazzi *et al.*, 1993).

9.2.1 Visual Perception

A valid neural theory of cognitive biological vision and visual perception should go beyond explaining the role of the massive feedback pathways in the primate brain. It should also explain the computational need for and the role of neural layers within which the process of "visual imagination" is performed. Since humans (and perhaps all primates) are able to imagine a visual stimulus or a scene, it is highly likely that such process is also crucial to visual perception. A theory of visual perception should therefore be able to provide an integrated approach that relies on the interactions between the bottom-up visual processes that are responsible for "seeing" and the top-down processes that are responsible for "imagining". Numerous psychological examples demonstrate that the process of "seeing" is not sufficient for visual perception. These examples demonstrate that the brain works in closed loop fashion and achieves visual perception when the loop is successfully closed. A good example of an ambiguous visual stimulus that so potently demonstrates the closed loop nature of the brain is the Boring's wife - mother-in-law image. Although most humans eventually perceive both visual precepts, usually considerable amount of viewing time is needed before one of the gestalts is perceived. Novices to this image have a great difficulty in perceiving the second precept once their attention is focused on the first. However, if they also manage to perceive the second precept then the two will be perceived in an oscillatory fashion, i.e., the subject experiences perceptual reversals. The other observation that can be made and which has important implications to neural theories of visual perception is that although the mother-in-law image is binary, the underlying visual structure (the young/old woman's face) is perceived as three dimensional.

What happens at the instant one of the gestalts is perceived? The most obvious observation is that the perception of the alternate fades. However, this does not tell us very much other than that we are not capable of perceiving more than one stimulus at a time. We claim that, at the instant the stimulus is perceived, a closed loop steady state is achieved and is made possible by the top-down memory. Where does this visual memory end up? We certainly don't see it. It is plausible to suggest that this memory

ends-up in a "cognitive visual space" that is also used when we imagine a visual stimulus. That this may be the case is supported by visual imagery experiments. Considerable amount of findings from neuropsychology provide evidence that visual imagery and perception share common neural substrates (Kosslyn 1980; Kosslyn *et al.,* 1993). This is further support by neuropsychological experiments that show that patients who cannot imagine a visual scene (or an object) from a given vantage point also fail to perceive that scene or object when it is presented in the same part of the visual field (Bisiach and Luzzatti 1978; Bisiach *et al.,* 1981).

Because the adaptive resonance concept of ART almost proposes the closed loop functioning of the brain, ART is placed in a favourable position to lead the way to a general theory of cognitive biological vision and visual awareness. In the context of the above, the concept of adaptive resonance can be used to define visual perception, visual awareness, selective visual attention and the necessary conditions for visual learning. Below we offer a neural definition of these concepts.

(i) *visual perception and awareness is a brain state during which the external sensory stimulus elicits a top-down imagination (or memory) of a stimulus with which it resonates.*

(ii) *selective visual attention is a neural mechanism that enables visual perception by selective transmission of exogenously and endogenously generated signals.*

(iii) *only the consonant (matching) reverberation between the selected external stimulus and the imagined stimulus (or recalled memory) enters visual awareness and leads to memory adaptation.*

(iv) *the established resonance must be stable for some minimum period of time for it to be detected.*

Visual perception of an external sensory stimulus will thus fail when the external stimulus cannot activate the correct top-down imagination (or memory) of that stimulus or cannot resonate with it. This dual and complementary nature of perceptual processing arises from a need to survive and carry out intelligent behaviour in a complex sensory environment, whose sensory events need not all be analysed and interpreted. The above definitions do not differentiate between the perception of 2-D and 3-D visual objects. One may wonder how is it that a two dimensional binary visual stimulus (such as the binary image of wife - mother-in-law) can resonate with an imagined three dimensional surface (a face), unless one transforms the other. We propose that it is the bottom-up 2-D neural representation of the external stimulus that, by recalling the imagined (or memorised) 3-D representation, gets transformed into a three dimensional representation by the recalled representation (the difference between the two is

discussed in section 9.2.2). The external stimulus will thus forever be perceived as a two dimensional stimulus unless an appropriate three dimensional imagination or memory is recalled which can lead to resonance with the input. If the top-down attentional signals are defective and cannot transform the stimulus into a 3-D precept, then even direct viewing of a person's face (or viewing a picture of a real face) will also be perceived as a two dimensional stimulus. There are a number of cases in neuropsychological literature of patients who cannot perceive and identify faces. An interesting example of a patient who cannot perceive faces as three dimensional has claimed "I can see the eyes, nose, and mouth quite clearly but they just don't add up. They all seem chalked in, like on a blackboard", Pallis, 1955. Further studies of prosopagnosia (an impairment of face recognition and face memory) has led to a conclusion that faces have a special representation. However, we suggest that it is the 3-D representational system in these patients that is malfunctioning.

9.2.2 2-D versus 3-D Representation

We suggest that a 2-D representation of an input stimulus is characterized by synchronized firing of neural cells whose position in the layer corresponds to the object's boundary (or a line), whereas a 3-D surface depth is also represented by synchronized firing but across a number of neural layers. Motion sensitive layers of the visual cortex should play an important role in providing the initial modulation that can then be stored as appropriately. When a sufficiently large repertoire of modulated patterns are stored in memory, then their subsequent activation by static and binary inputs (such as the Boring's wife-mother-in-law image) should enable the appropriate 3-D perception. Often, the reflected intensity of light from real objects may be related to the surface structure, but this need not be so for the brain to interpret the correct surface structure (example is that of uneven illumination of an object).

The currently popular view is that the brain derives surface shape and depth from shading. If we are to believe that synchronized oscillations (i.e., resonances) are the key states of biological visual perception, then we must acknowledge that there are only two modes of synchronizations possible: (i) synchronization with zero phase lag; or (ii) synchronization with a constant (or modulated) phase lag. While in ideal cases the bottom-up processes may extract the correct phase modulations, in general they need to be supplemented by top-down modulations. The fact that we can all perceive depth in binary images, drawings and sketches clearly demonstrates that shading is not required to enable us to perceive surface shape and depth, although it may be a good cue in removing possible ambiguities.

9.3 Conclusions

We conclude that flexible design principles that are based on feedforward-feedback interactions in a closed-loop real-time competitive neural circuit whose modulatory mechanisms can dynamically retune the signal transmission gains and the cellular receptive field profiles at various stages of processing overcomes some of the problems and limitations that are faced by the rigid architecture of the current artificial neural networks. The neuro-engineeering design principles, mechanisms and circuits as proposed in the thesis provide a new and robust method for solving some of the most difficult problems in visual object recognition that are currently not well handled by the state-of-the-art artificial neural networks and the more conventional computer vision systems. These design principles also open new avenues for further research into more advanced modelling of cognitive and perceptual real-time artificial neural systems that use selective information processing.

From a technological viewpoint, it is unlikely that the application of the SAART/ASAART neural networks to general visual problems will be possible in the foreseeable future. The reason for this is that we currently do not have plausible hardware technology to implement extremely large scale artificial neural networks that are required for general visual applications. However, it may be possible to apply the proposed neural networks to a restricted set of problems that do not have a large degree of freedom. For such restricted problems, the implementation of the SAART neural network on a silicon wafer, coupled with time-multiplexing, may be sufficient. Multi-wafer integration may be required for more realistic visual problems. Nevertheless, the knowledge that we may finally be able to solve some extremely difficult problems in vision may provide an impetus for novel hardware processing technologies and implementations of the proposed artificial neural networks.

9.4 Recommendations

Since our computer simulations (being performed on a 486 PC) were limited by the computing platform, there is a need to integrate the proposed concepts and demonstrate it on a larger scale. We therefore recommend that:

(i) The concept of self-regulated attentional modulation of distributed cellular receptive field profiles be integrated into the SAART neural network and demonstrated on the relevant visual images;

(ii) The proposed concepts for size and orientation invariant object recognition be further developed and demonstrated. This may initially done with multiple copies of the Feedforward Excitation-Feedback Presynaptic Facilitation (FFE-FBPF) neural circuit.

From a mathematical point of view, we feel that there are still many problems that need to be addressed, such as the stability analysis (which we have ignored). Our extensive computer simulations have demonstrated that for the chosen set of parameters, the SAART neural network always ends up in a stable state. In addition to developing a formal mathematical analysis, we feel that there is also a need to develop a full mathematical theory of modulated neural layers. It is quite possible that there are many other useful neural layers that we have not considered, such as the on-surround off-centre shunting competitive neural layer (which would be very useful for the modelling of the lower level visual cortical areas, e.g., LGN and V1).

From an engineering point of view, we recommend that a study be carried out to determine the feasibility of implementing the proposed modulated competitive neural layers in VLSI.

Although we have presented our theory in the context of biological vision and have limited our discussion to the inferior temporal cortex and its immediate afferents, the concept of priming by enhanced transmitter mobilization is quite general. In addition to enhancing the strength of signal, the signals in the primed pathway will access the attention of other layers faster than a non-primed pathway. For example, if a certain stimulus is expected it will be processed faster than if it was unexpected, since in the former case all the relevant pathways will have their respective transmitter levels enhanced by the top-down priming signals. Thus, there is a need to investigate how the SAART concepts may be used in other sensory madalities where sensory clutter makes it too difficult to separate the various stimuli (e.g., in speech recognition).

Finally, we recommend that Neuro-Engineering (which we have formally defined in Chapter 3) be further evolved as a new scientific discipline. Neuro-Engineering defies the traditional departmental boundaries that separates psychologists, neurobiologists, neural network theoreticians, mathematicians and engineers. Although a researcher from one of the traditional disciplines may initially become disoriented when confronted with the wealth of the experimental data and a new mathematical language, we have found that it is only when we attempt to solve a realistic problem (that our brains need to cope with) that we can discover some underlaying mechanisms. A good knowledge of engineering design methodologies thus plays a crucial role since it enables one to see the role of the mechanism at the more global level. Through the repeated crossings of the traditional boundaries, such a researcher may eventually be able to provide an

unified theoretical understanding of neurobiological systems, from the level of single neurons and chemical synapses, to the behavioural level of whole organisms and finally to the realisation into a machine based system.

Appendix A

Basic Structure of a Modulated Competitive Neural Layer

In this Appendix we derive the basic structure of a Presynaptically Facilitated Shunting Competitive Neural Layer, which is then formally developed in Chapter 4 and analysed in Chapter 5.

In general, each neuron in a shunting competitive neural layer may have a number of excitatory and inhibitory synaptic inputs whose combined effect determines the cellular response. Each of the input synapses may also be modulated by facilitatory/inhibitory gain control signals. In order to develop a general mathematical model of a transmitter gated and presynaptically modulated shunting competitive neural layer, we first consider the qualitative properties of the model described by equation (A.1). This is a version of the shunting competitive feedback equation with a single on-centre and wide off-surround anatomy that was introduced by S. Grossberg (1973, 1988).

$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)[J_i^+ + f(x_i)] - (C + x_i)\left[J_i^- + \sum_{j=1}^{n} F_{ji} g(x_j)\right] \qquad (A.1)$$

where J_i^+ is the total excitatory input into the layer at i^{th} cell position; J_i^- is the total inhibitory input to the i^{th} cell. The term

$$\sum_{j=1}^{n} F_{ji} g(x_j) \qquad (A.2)$$

is the total lateral feedback inhibition due to cellular competition, where F_{ji} is the lateral interaction strength from the j^{th} cell to the i^{th} cell in the layer. The term $(B - x_i)[J_i^+ + f(x_i)]$ says that the total excitatory input J_i^+ and the self-excitatory feedback signal $f(x_i)$ increase the activity of x_i. The self-excitation provided by the term $(B - x_i)f(x_i)$ in the above equation ensures that a layer of neurons whose dynamics is described by (A.1) retains a contrast enhanced memory of its inputs after the input offset. The term

$$-(C+x_i)\left[J_i^+ + \sum_{j=1}^{n} F_{ji}g(x_j)\right] \qquad (A.3)$$

says that the inhibitory inputs drive the cell towards is negative saturation level of -C. The cellular activity of (A.1) is thus restricted to the range $(-C, B)$.

We now wish to qualitatively analyse the equilibrium behaviour of (A.1) when $J_i^+ = J_i^-, \forall i$. In order to do so, let us temporarily ignore the positive self-excitation term and write (A.1) in the following form:

$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)V_i^+ - (C + x_i)(V_i^- + \overline{v}_i + \Gamma) \qquad (A.4)$$

where V_i^+ is the total transmitter gated excitatory postsynaptic potential (EPSP) acting on the ith cell that is due to all the excitatory input synapses, given by (A.5); V_i^- is the total inhibitory postsynaptic potential (IPSP) acting on the cell that is due to all the inhibitory input synapses, given by (A.6); $\overline{v}_i$ is the lateral feedback inhibition that is mediated by slow inhibitory interneurons, given by the additive equation (A.7); Γ is a tonic level of inhibition. Figure A.1 illustrates a simple model of transmitter gated shunting competitive neural layers (on-centre off-surround anatomy) as modelled by (A.4) and (A.7).

$$V_i^+ = \sum_{j}^{M} v_{ji}^+ \qquad (A.5)$$

$$V_i^- = \sum_{j}^{N} v_{ji}^- \qquad (A.6)$$

where M represents the number of excitatory input synapses into the i^{th} cell and N represents the number of inhibitory input synapses into the cell. For simplicity we will assume that M = N.

$$\frac{d\overline{v}_i}{dt} = -\overline{A}\overline{v}_i + \overline{B}\sum_{j \neq i} f(x_j) \qquad (A.7)$$

where $\overline{A} < A$ and $\overline{B} \ll B$. The thresholding function $f(x_j) = \max(x_j - \Theta, 0) = [x_j - \Theta]^+$ in the above equation ensures that a cell contributes to lateral inhibition of other cells only when its activity is above a threshold of Θ. The excitatory postsynaptic potential due to one input synapse is given by

$$\frac{dv_{ji}^+}{dt} = -D^+ v_{ji}^+ + E^+ S_{ji}^+ y_{ji}^+ \qquad (A.8)$$

where D^+ and E^+ are constants that define the decay and the rise time of EPSP respectively; S_{ji}^+ is the synaptic input signal; y_{ij}^+ is the amount of mobilized transmitter in the synapse. Since we are initially ignoring any self-excitation that may be provided by postsynaptic feedback, we can write the synaptic dynamics by the following two equations:

$$\frac{du_{ji}^+}{dt} = \alpha_0(z_{ji}^+ - u_{ji}^+) - \beta_0(u_{ji}^+ - y_{ji}^+) \tag{A.9}$$

$$\frac{dy_{ji}^+}{dt} = \beta_0(u_{ji}^+ - y_{ji}^+) - \rho S_{ji}^+ y_{ji}^+ - \gamma y_{ji}^+ \tag{A.10}$$

where z_{ji}^+ is the transmitter production level (presently assumed to be constant); u_{ji}^+ is the level of stored transmitter; and y_{ji}^+ is the mobilized transmitter (available for release). Similar set of equations may be written for inhibitory synapses.

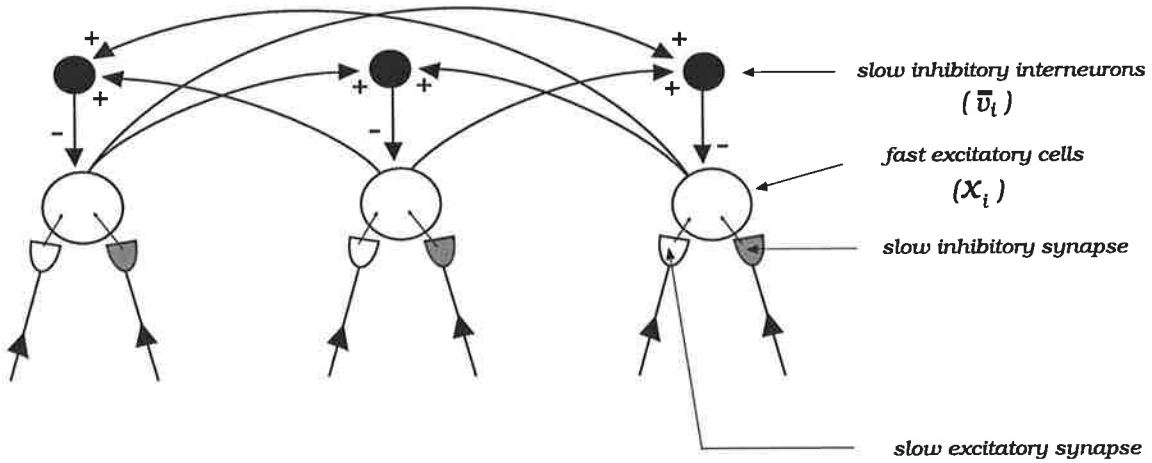


FIGURE A.1. Simple model of transmitter gated shunting competitive neural layers. External inputs to each cell are gated by slow transmitter based synapses. Lateral competition is mediated by slow inhibitory interneurons.

If lateral feedback inhibition ($\bar{v}_i$) is assumed to be due to slowly charging inhibitory interneurons and if the input synapses are also assumed to have slow dynamics relative to the fast excitatory cells, then we can represent (A.4) by its equilibrium solution given by (A.11).

$$x_i = \frac{BV_i^+ - C(V_i^- + \bar{v}_i + \Gamma)}{A + V_i^+ + V_i^- + \bar{v}_i + \Gamma} \tag{A.11}$$

Since B and C are positive constants that determine the upper and lower saturation levels respectively, we can let $C = aB$, where $a \geq 0$. Then equation (A.11) becomes

$$x_i = \frac{BV_i^+ - aB(V_i^- + \overline{v}_i + \Gamma)}{A + V_i^+ + V^- + \overline{v}_i + \Gamma}$$

$$= \frac{B(V_i^+ - a(V_i^- + \overline{v}_i + \Gamma))}{A + V_i^+ + V^- + \overline{v}_i + \Gamma} \qquad (A.12)$$

whose resting potential (no synaptic inputs) is given by

$$x_i = \frac{-aB(\overline{v}_i + \Gamma)}{A + \overline{v}_i + \Gamma} \qquad (A.13)$$

Because the resting potential is negative (provided that $a \neq 0$ and $\Theta \geq 0$), then, by (A.7), the lateral feedback inhibition ($\overline{v}_i$) will decay to zero. Therefore, the resting cellular potential is given by

$$x_i = \frac{-aB\Gamma}{A + \Gamma} \qquad (A.14)$$

Now let as consider what happens to (A.12) as we parametrically increase both V_i^+ and V_i^- while keeping them equal (i.e., $V_i^+ = V_i^-, \forall i$). Then

$$\lim_{\substack{V_i^+ \to \infty \\ (V_i^- = V_i^+)}} x_i = \frac{BV_i^+ - aB(V_i^+ + \overline{v}_i + \Gamma)}{A + 2V_i^+ + \overline{v}_i + \Gamma}$$

$$= \frac{(1-a)B}{2} \qquad (A.15)$$

which monotonically approaches the asymptote of $(1-a)B/2$. When $0 < a < 1$ ($C < B$), $(1-a)B/2$ is positive. Thus, even though the cell is receiving net inhibition via the term Γ, its equilibrium activity keeps increasing as we parametrically increase both the excitatory and the inhibitory inputs. Similarly, when $a > 1$ ($C > B$), the equilibrium monotonically approaches a negative asymptote. The problem does not disappear when $a = 1$ (i.e., when $C = B$), even though the asymptote is shifted to zero (it is above the negative resting state). When $a = 0$ (i.e., $C = 0$), the resting potential is zero and the asymptote is shifted to $B/2$.

In Figure A.2 we show the simulation results of a transmitter gated shunting competitive neural layer defined by equations (A.7)-(A.11). The excitatory and the

inhibitory synaptic inputs were equal in magnitude and progressively increased. Note that each point on the graphs was obtained after 100 iterations of the above equations and that the network variables were reset upon the beginning of each simulation run.

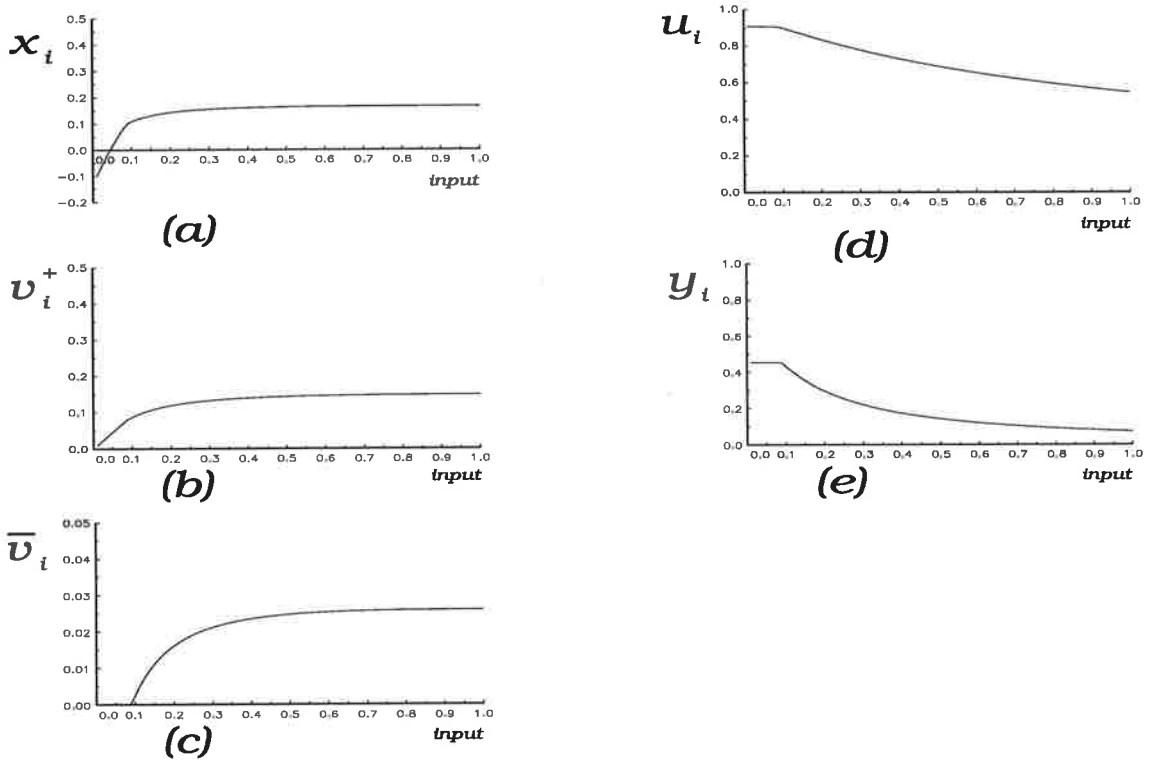


FIGURE A.2. (a) Steady state synaptic excitatory postsynaptic potential as a function of the input signal strength (equal in strength to the synaptic inhibitory postsynaptic potential).

As can be seen in the graphs of Figure A.2, the steady state cellular activation level converges to a level that is positive and well above the resting state. The graphs clearly demonstrate that even though the cell is receiving higher inhibition than excitation (the difference being $\overline{v}_i + \Gamma$), the steady state cellular activity of the system described by the above equations still keeps increasing and converges at a positive value well above the negative resting potential (the actual asymptote depends on the magnitude of the tonic inhibition term Γ). This undesirable characteristic of the system where the inhibitory synaptic inputs do not effectively regulate the cellular response is not very useful and would make it awkward to engineer a neural network whose layers use both excitatory and inhibitory input synapses.

In order to overcome this undesirable characteristic of the system described by equation (A.1), we propose the following equation

$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)[G(V_i^+ - V_i^-)]^+ - (C + x_i)\left([G(V_i^- - V_i^+)]^+ + \overline{G}\overline{v}_i + \Gamma\right) \tag{A.16}$$

where $[G(V_i^+ - V_i^-)]^+ = \max(G(V_i^+ - V_i^-), 0)$ is the net excitatory postsynaptic potential that drives the cell towards its positive saturation limit of B; $[G(V_i^- - V_i^-)]^+ + \overline{G}\overline{v}_i + \Gamma$ is the net inhibitory postsynaptic potential that drives the cell towards its negative saturation point of -C; G is the gain of excitatory and inhibitory synaptic inputs (assumed to be equal); Γ is the tonic level of inhibition. The equilibrium solution of (A.16) is given by

$$x_i = \frac{B[G(V_i^+ - V_i^-)]^+ - C\left([G(V_i^- - V_i^+)]^+ + \overline{G}\overline{v}_i + \Gamma\right)}{A + [G(V_i^+ - V_i^-)]^+ + [G(V_i^- - V_i^+)]^+ + \overline{G}\overline{v}_i + \Gamma} \tag{A.17}$$

Note that when $V_i^+ = V_i^-$, the equilibrium state is independent of the synaptic inputs and is given by

$$\begin{aligned} x_i &= -\frac{C(\overline{G}\overline{v}_i + \Gamma)}{A + (\overline{G}\overline{v}_i + \Gamma)} \\ &= -\frac{C\Gamma}{A + \Gamma} \qquad (\text{since} \quad \overline{v}_i \to 0 \quad \text{when } x_i < 0 \quad \forall i) \end{aligned} \tag{A.18}$$

which is zero when C = 0. The system described by (A.16) still has the useful properties of automatic gain control provided by the $(B - x_i)$ term. All neural circuits in the thesis are based on a version of (A.16) with C = 0, shown below.

$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)[G(V_i^+ - V_i^-)]^+ - x_i\left([G(V_i^- - V_i^+)]^+ + \overline{G}\overline{v}_i + \Gamma\right) \tag{A.19}$$

whose equilibrium is given by

$$x_i = \frac{B[G(V_i^+ - V_i^-)]^+}{A + [G(V_i^+ - V_i^-)]^+ + [G(V_i^- - V_i^+)]^+ + \overline{G}\overline{v}_i + \Gamma} \tag{A.20}$$

In order to generate the model described by (A.7) - (A.10) and (A.16), we have ignored the positive self-excitation term that appears in (A.1). In the above we have also considered an idealised model of chemical synapses whose internal dynamics is

independent of postsynaptic cellular activity. The full mathematical model of a Presynaptically Facilitated Shunting Competitive Neural Layer, whose self-excitation is provided by postsynaptic feedback, is developed in Chapter 4 (section 4.3.2).

Appendix B

Parameters of the Simulated Circuits and Networks

This appendix lists the equations of the oscillatory neural layers and the parameters of the SAART neural network. Most of the listed parameters are based on the parameter design procedure of Chapter 5, with the exception of the oscillatory neural layers. We will therefore list the equations that were simulated for the oscillatory layers.

B.1 Equations for the Oscillatory Neural Layers

The following are the equations that were simulated for oscillatory neural layers of section 4.5 (Chapter 4). Note that all the parameters in the following four simulations are almost identical (the main difference being in the coupling strength).

(I) Equations for Figure 4.20 (Chapter 4)

The following is a list of equations that were used to obtain the data in Figure 4.20 of Chapter 4. The input is constant across the layer ($J_i = 1, \forall i$).

Excitatory postsynaptic potential

$$v_i(t+1) = v_i(t) - 0.5v_i(t) + 10J_i[y_i(t) - 0.2]^+[1 + 80f(x_i(t))] \qquad (B.1)$$

Stored transmitter

$$u_i(t+1) = u_i(t) + 0.05(1 - u_i(t)) - 0.004[2 + J_i + 300J_i f(x_i(t)](u_i(t) - y_i(t)) \qquad (B.2)$$

Mobilized transmitter

$$y_i(t+1) = y_i(t) - 0.01y_i(t) - 0.001J_i[y_i(t) - 0.2]^+[1 + 20f(x_i(t))] + 0.002[2 + J_i + 30s_i(t)](u_i(t) - y_i(t)) \quad (B.3)$$

where $s_i(t) = \sum_{j \neq i} f(x_j(t))$ is the total lateral presynaptic facilitation from nearest neighbours.

Postsynaptic cellular activation

$$x_i(t+1) = x_i(t) - 0.01x_i(t) + 0.01[1 - x_i(t)]v_i(t) - 5x_i(t)\bar{v}_i(t) \quad (B.4)$$

Lateral inhibition

$$\bar{v}_i(t+1) = \bar{v}_i(t) - 0.001\bar{v}_i(t) + 0.000025\bar{I}_i(t) \quad (B.5)$$

where $\bar{I}_i(t) = \sum_{j \neq i} f(x_j(t))$ and is summed over the whole layer of 20 neurons. Function $f(x_i(t)) = \max(x_i(t) - 0.3, 0)$.

(II) Equations for Figure 4.23 (Chapter 4)

The following is a list of equations that were used to obtain the data in Figure 4.23 of Chapter 4 (all inputs $J_i = 1, \forall i$).

Excitatory postsynaptic potential

$$v_i(t+1) = v_i(t) - 0.5v_i(t) + 10J_i[y_i(t) - 0.2]^+[1 + 80f(x_i(t)) + 40s_i(t)] \quad (B.6)$$

where $s_i(t) = \sum_{j \neq i} f(x_j(t))$ is the total lateral presynaptic excitation from nearest neighbours.

Stored transmitter

$$u_i(t+1) = u_i(t) + 0.05(1 - u_i(t)) - 0.004[2 + J_i + 300J_i f(x_i(t)](u_i(t) - y_i(t)) \quad (B.7)$$

Mobilized transmitter

$$y_i(t+1) = y_i(t) - 0.01y_i(t) - 0.001J_i[y_i(t) - 0.2]^+[1 + 20f(x_i(t))] + 0.002[2 + J_i](u_i(t) - y_i(t)) \quad (B.8)$$

Postsynaptic cellular activation

$$x_i(t+1) = x_i(t) - 0.01x_i(t) + 0.01[1 - x_i(t)]v_i(t) - 5x_i(t)\overline{v}_i(t) \quad (B.9)$$

Lateral inhibition

$$\overline{v}_i(t+1) = \overline{v}_i(t) - 0.001\overline{v}_i(t) + 0.000025\overline{I}_i(t) \quad (B.10)$$

where $\overline{I}_i(t) = \sum_{j \neq i} f(x_j(t))$ and is summed over the whole layer of 20 neurons. Function $f(x_i(t)) = \max(x_i(t) - 0.3, 0)$.

(III) Equations for Figure 4.25 (Chapter 4)

The following is a list of equations that were used to obtain the data in Figure 4.25 of Chapter 4 (random inputs into the layer).

Excitatory postsynaptic potential

$$v_i(t+1) = v_i(t) - 0.5v_i(t) + 10J_i[y_i(t) - 0.2]^+[1 + 80f(x_i(t)) + 20s_i(t)] \quad (B.11)$$

where $s_i(t) = \sum_{j \neq i} f(x_j(t))$ is the total lateral presynaptic excitation.

Stored transmitter

$$u_i(t+1) = u_i(t) + 0.05(1 - u_i(t)) - 0.004[2 + J_i + 300J_i f(x_i(t)](u_i(t) - y_i(t)) \quad (B.12)$$

Mobilized transmitter

$$y_i(t+1) = y_i(t) - 0.01y_i(t) - 0.001J_i[y_i(t) - 0.2]^+[1 + 20f(x_i(t))] + 0.002[2 + J_i](u_i(t) - y_i(t)) \quad (B.13)$$

Postsynaptic cellular activation

$$x_i(t+1) = x_i(t) - 0.01x_i(t) + 0.01[1 - x_i(t)]v_i(t) - 5x_i(t)\bar{v}_i(t) \qquad (B.14)$$

Lateral inhibition

$$\bar{v}_i(t+1) = \bar{v}_i(t) - 0.001\bar{v}_i(t) + 0.000025\bar{I}_i(t) \qquad (B.15)$$

where $\bar{I}_i(t) = \sum_{j \neq i} f(x_j(t))$ and is summed over the whole layer of 20 neurons. Function $f(x_i(t)) = \max(x_i(t) - 0.3, 0)$.

(IV) Equations for Figure 4.27 (Chapter 4)

Excitatory postsynaptic potential

$$v_i(t+1) = v_i(t) - 0.5v_i(t) + 10[y_i(t) - 0.2]^+ [J_i + 80J_i f(x_i(t)) + 40s_i(t)] \qquad (B.16)$$

where $s_i(t) = \sum_{j \neq i} f(x_j(t))$ is the total lateral presynaptic excitation from nearest neighbours.

Stored transmitter

$$u_i(t+1) = u_i(t) + 0.05(1 - u_i(t)) - 0.004[2 + J_i + 300J_i f(x_i(t)]\,(u_i(t) - y_i(t)) \qquad (B.17)$$

Mobilized transmitter

$$y_i(t+1) = y_i(t) - 0.01y_i(t) - 0.001J_i[y_i(t) - 0.2]^+ [1 + 20f(x_i(t))] + 0.002[2 + J_i]\,(u_i(t) - y_i(t)) \qquad (B.18)$$

Postsynaptic cellular activation

$$x_i(t+1) = x_i(t) - 0.01x_i(t) + 0.01[1 - x_i(t)]v_i(t) - 5x_i(t)\bar{v}_i(t) \qquad (B.19)$$

Lateral inhibition

$$\bar{v}_i(t+1) = \bar{v}_i(t) - 0.001\bar{v}_i(t) + 0.000025\bar{I}_i(t) \qquad (B.20)$$

where $\bar{I}_i(t) = \sum_{j \neq i} f(x_j(t))$ and is summed over the whole layer of 20 neurons. Function $f(x_i(t)) = \max(x_i(t) - 0.3, 0)$.

B.2 Parameters for the SAART Neural Network

The following are all the SAART neural network parameters that were used in the computer simulations of Chapter 7. Each differential equation was iterated using the Euler's first order approximation method, with $\Delta t = 1$. The parameters for all the processing Fields, with the exception of Field C1, are very similar and are mostly as prescribed in Chapter 5. The main difference between the parameters across all the simulations discussed in the chapter was the learning rate (very slow in the simulation on noisy inputs, very fast in other cases). The other main difference from the prescribed parameters was that we have set the threshold level for postsynaptic activation to be 0.01 (instead of 0.1). However, the parameters that were designed for the higher threshold level are also valid for lower thresholds. The main reason for setting a lower threshold was to enable new inputs to register into the reverberatory loop more effectively (particularly when the network needs to learn slowly, as in the case of noisy input data).

(i) Field A1

$A^{a1} = 1$; $B^{a1} = 1$; $G^{a0a1} = 1000$; $\overline{G}^{a1} = 5000$; $\overline{A}^{a1} = 0.1$; $\overline{B}^{a1} = 0.1/n$ where n is the number of neurons in the layer (n = 1024); $D^{a0a1} = 0.5$; $Y^{a0a1} = 0.005$; $\rho_v^{a0a1} = 0.5$; $K_v^{a0a1} = 5$; $\alpha_u^{a0a1} = 0.05$; $\beta_u^{a0a1} = 0.01$; $K_u^{a0a1} = 0.99$; $\gamma_y^{a0a1} = 0.5$; $\rho_y^{a0a1} = 0.05$; $K_y^{a0a1} = 0.5$; $\beta_y^{a0a1} = 0.01$; $H_y^{a0a1} = 0.5$; $\Theta^{a1} = 0.01$; $z_{ij}^{a0a1} = 1, \forall(i,j)$.

(ii) Field A2

$A^{a2} = 1$; $B^{a2} = 1$; $G^{a0a2} = 1000$; $\overline{G}^{a2} = 5000$; $\overline{A}^{a2} = 0.1$; $\overline{B}^{a2} = 0.1/n$; $D^{a0a2} = 0.5$; $Y^{a0a2} = 0.005$; $\rho_v^{a0a2} = 0.5$; $K_v^{a0a2} = 5$; $\alpha_u^{a0a2} = 0.05$; $\beta_u^{a0a2} = 0.001$; $K_u^{a0a2} = 0.01$; $\gamma_y^{a0a2} = 0.5$; $\rho_y^{a0a2} = 0.05$; $K_y^{a0a2} = 0.5$; $\beta_y^{a0a2} = 0.001$; $H_y^{a0a2} = 100$; $\overline{H}_y^{a0a2} = 10000$; $\Theta^{a2} = 0.01$; $z_{ij}^{a0a2} = 1, \forall(i,j)$.

(iii) Field B1

$A^{b1} = 1$; $B^{b1} = 1$; $G^{a1b1} = G^{b3b1} = 1000$; $\overline{G}^{b1} = 5000$; $\overline{A}^{b1} = 0.1$; $\overline{B}^{b1} = 0.1/n$; $D^{a1b1} = D^{b3b1} = 0.5$; $Y^{a1b1} = Y^{b3b1} = 0.005$; $\rho_v^{a1b1} = \rho_v^{b3b1} = 0.5$; $K_v^{a1b1} = K_v^{b3b1} = 5$; $\alpha_u^{a1b1} = \alpha_u^{b3b1} = 0.05$; $\beta_u^{a1b1} = \beta_u^{b3b1} = 0.01$; $K_u^{a1b1} = K_u^{b3b1} = 0.5$; $\gamma_y^{a1b1} = \gamma_y^{b3b1} = 0.5$; $\rho_y^{a1b1} = \rho_y^{b3b1} = 0.05$; $K_y^{a1b1} = K_y^{b3b1} = 0.5$; $\beta_y^{a1b1} = 0.1$; $\beta_y^{b3b1} = 0.01$; $\Theta^{b1} = 0.01$; $z_{ij}^{a1b1} = z_{ij}^{b3b1} = 1, \forall(i,j)$.

(iv) Field B3

$A^{b3} = 1$; $B^{b3} = 1$; $G^{b2b3} = 10000$; $G^{a1b3} = 500$; $\overline{G}^{b3} = 5000$; $\overline{A}^{b3} = 0.1$; $\overline{B}^{b3} = 0.1/n$; $D^{a1b3} = D^{b2b3} = 0.5$; $Y^{a1b3} = Y^{b2b3} = 0.005$; $\rho_v^{a1b3} = \rho_v^{b2b3} = 0.5$; $K_v^{a1b3} = K_v^{b2b3} = 5$; $\gamma_y^{a1b3} = \gamma_y^{b2b3} = 0.5$; $\alpha_u^{a1b3} = \alpha_u^{b2b3} = 0.05$; $\beta_u^{b2b3} = 0.001$; $K_u^{b2b3} = 0.02$; $\beta_y^{b2b3} = 0.01$; $\rho_y^{b2b3} = 0.01$; $K_y^{b2b3} = 0.1$; $\beta_u^{a1b3} = 0.001$; $K_u^{a1b3} = 0.05$; $\beta_y^{a1b3} = 0.05$; $\rho_y^{a1b3} = 0.05$; $K_y^{a1b3} = 0.5$; $\Theta^{b3} = 0.01$; $z_{ij}^{a1b3} = z_{ij}^{b2b3} = 1, \forall(i,j)$;

(v) Field B2

$A^{b2} = 1$; $B^{b2} = 10$; $G^{b1b2} = 100$; $G^{c1b2} = 10000$; $\overline{G}^{b2} = 5000$; $\overline{A}^{b2} = 0.1$; $\overline{B}^{b2} = 0.1/n$; $D^{b1b2} = 0.5$; $Y^{b1b2} = Y^{c1b2} = 0.005$; $\rho_v^{b1b2} = 0.5$; $K_v^{b1b2} = 5$; $\alpha_u^{b1b2} = 0.05$; $\beta_u^{b1b2} = 0.01$; $\beta_y^{b1b2} = 0.05$; $\rho_y^{b1b2} = 0.01$; $K_y^{b1b2} = 0.1$; $z_{ij}^{b1b2} = 1, \forall(i,j)$.

(vi) Field C1

$A^{c1} = 1$; $B^{c1} = 1$; $G^{b1c1} = 1000$; $\overline{G}^{c1} = 5000$; $\overline{A}^{c1} = 0.1$; $\overline{B}^{c1} = 0.005$; $D^{c1} = 0.5$; $K_v^{b1c1} = 0.05$; $r^{b1c1} = 0.005$; $\beta_w^{b1c1} = 0.01$; $K_w^{b1c1} = 1$; $\alpha_r^{b1c1} = 0.001$; $\beta_r^{b1c1} = 0.002$; $K_r^{b1c1} = 0.2$; $\Psi^{c1} = 0.5$; $\gamma_\mu = 0.5$; $\mu_0^{b1c1} = 0.05$; $\gamma_y^{b1c1} = 0.01$; $\rho_y^{b1c1} = 0.01$; $\beta_y^{b1c1} = 0.01$; $K_y^{b1c1} = 0.5$; $\alpha_u^{b1c1} = 0.1$; $\beta_u^{b1c1} = 0.01$;

(vii) Gated LTM learning equations

$\gamma_z^{c1b2} = \gamma_z^{b1c1} = 0.002$; $\varepsilon^{b1c1} = 0.2$ for slow learning (0.5 for fast learning); $\varepsilon^{c1b2} = 0.02$ for slow learning (0.5 for fast learning); all bottom-up and the top-down LTM variables were randomly initialized to small values.

Appendix C

Adding Selective Attention to ART-3 Neural Network

Below we suggest how ART-3 (and ART-2) neural network of Carpenter and Grossberg (1990) may be modified to model top-down memory guided selective attention. The modification shown is very easy to add to ART-3 and should enable it to recognize familiar inputs when embedded in cluttered and noisy background.

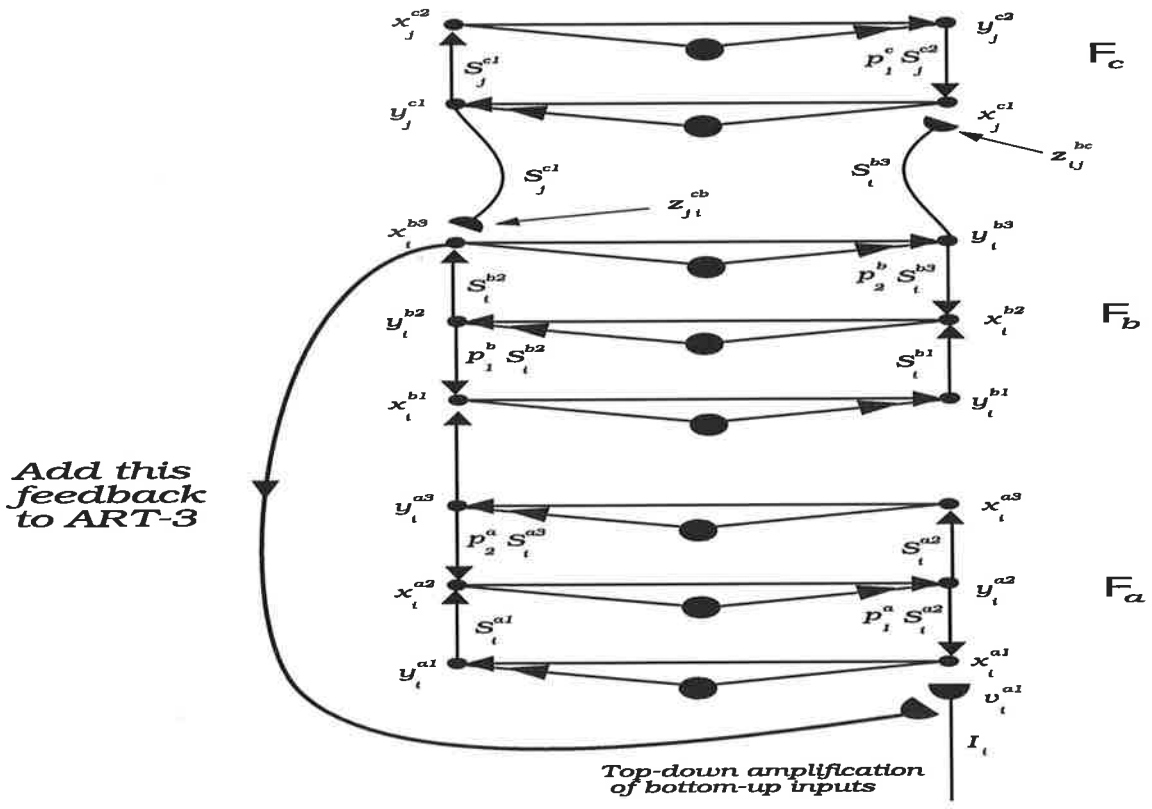


FIGURE C.1. Simple extension to ART-3 neural network to model selective attention. Project signals x_i^{b3} to amplify the bottom-up signals into Field F_a as indicated. Also reduce the top-down gain in F_a. This simple extension enables ART-3 to recognize previously learned 2-D shapes when they are subsequently embedded in clutter. However, the network is unable to learn in noisy inputs.

The following three equations specify the dynamics at the bottom-up input node.

$$x_i^{a1} = p_1^a + v_i^{a1} \tag{C.1}$$

$$\frac{dv_i^{a1}}{dt} = -v_i^{a1} + I_i u_i^{a1}(0.1 + x_i^{a1}) \tag{C.2}$$

$$\frac{du_i^{a1}}{dt} = (1 - u_i^{a1}) - u_i^{a1}(0.1 + I_i)(1 + x_i^{b3}) \tag{C.3}$$

Since Carpenter and Grossberg have not provided the learning laws for ART-3 (see Carpenter and Grossberg, 1990), below we specify the learning equations that were originally used by us. Note that ART-3 is the first of the ART based neural networks to use simple models of chemical synapses (in the bottom-up and the top-down adaptive memory pathways), but the model uses only two variables. Rather than using the published version (which did not model learning) we have used a modified version with an adaptive threshold applied to cells in Field F_c (Lozo, 1993a). The learning law for the top-down adaptive memory pathways, from cell j in F_c to cell i in F_b was

$$\frac{dz_{ji}^{cb}}{dt} = r(x_i^{xb3} - z_{ji}^{cb}) \tag{C.4}$$

where r is the learning rate (r = 0.5). Similar equation was used in the bottom-up memory pathway. Below we present our initial results that were obtained with the above modification.

(a) ship 1 (b) ship 2 (c) ship 3

FIGURE C.2. Binary silhouettes of three ships used to train the extended ART-3 neural network. These were scaled and aligned to approximately occupy the same region of the input image so that the shapes, rather than the sizes or locations, become the distinguishing features that the network has to learn. After the network was trained on the above three examples it was then exposed to the noisy images (shown below).

The network was initially trained on binary silhouettes of three different ships, after which it was tested on severely degraded images of these silhouettes, shown below. Note that the level of noise is high (probably higher than what may be expected in applications) and extends throughout the whole image region. The purpose of degrading these images by a severe amount of noise was to demonstrate that the network is as good as many well trained observers (and better than novices) in differentiating the ship classes in the test images.

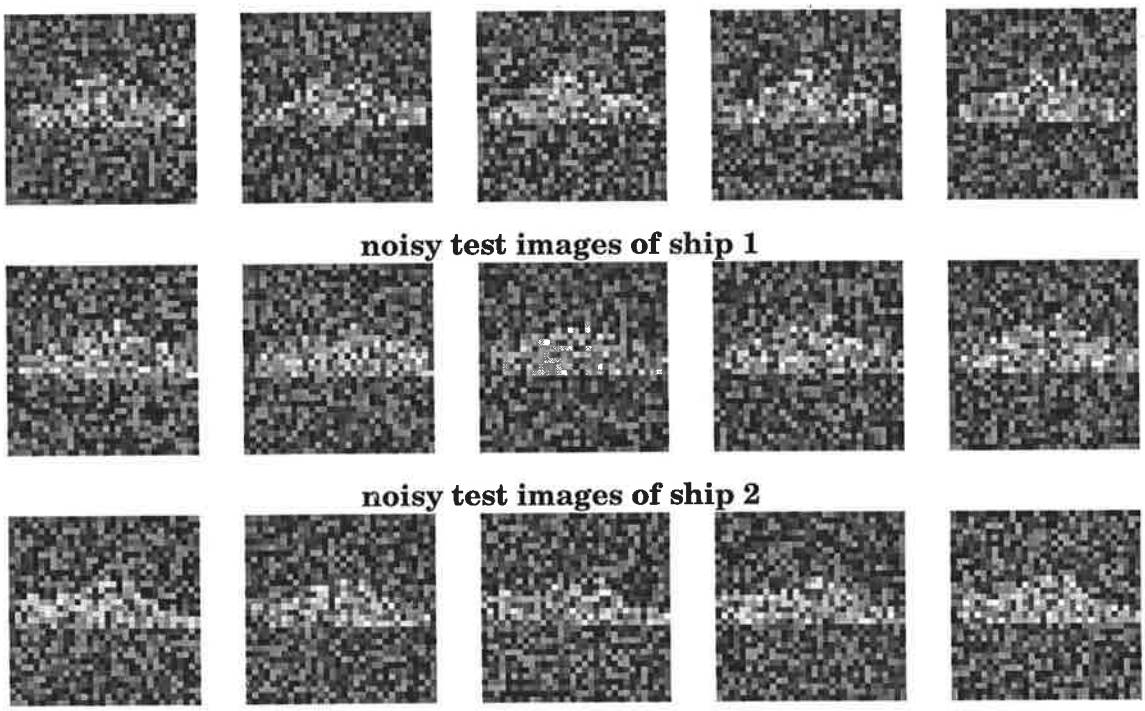


FIGURE C.3. Examples of noisy images (silhouettes of three different ships) used to test the extended ART-3 neural network. These images were generated from those of Figure C2 using the following equation: $I^* = r_1 I + r_2$ where I^* is a new pixel intensity, I is the original value,$0 \le r_1, r_2 \le 1$ and r_1, r_2 are random numbers generated by the random number generator. One hundred such images were generated for each of the three ship silhouettes.

The signal-to-noise ratio of the test images lies in range 2.8 to 3.2 DB. Below we provide the classification results (note that when the network unsuccessfully activated the same memory at least three times it was deemed that the input was unknown). The network's learning is never terminated, but once the memories of objects are well established (have converged) they cannot be corrupted by noise. Although the network was given up to 5 cells in field F_c, it only activated the nodes that contained the memory of the three silhouettes.

Recognition results (across a total of 300 images)

Correct	Incorrect	Unknown
82%	10%	8%

Although the above results are very impressive, we have since been able to establish that the concepts of top-down selective attention as discussed in this thesis should not be applied to grey level directly but to edge processed images (as it is possible to introduce "illusionary" states). Similarly, the concept of top-down modulatory feedback is valid only when the inputs are treated as 2-D spatial patterns.

Appendix D

Log-polar Transform

vs.

Parallel Frames of Reference

"We trace the oceans of hyperbole, controversy, and rediscovery which still flood our science of the inability of some investigators to fully let go of unappropriate technological metaphors and nineteenth century mathematical concepts. Although initially attractive because of their simplicity and accessibility, these approaches have regularly shown their impotence when they are confronted by a nontrivial set of the phenomena that they have set out to explain. A unified theoretical understanding cannot be achieved without an appropriate mathematical language in our science any more than in any other science."

S. Grossberg (in preface to **The Adaptive Brain**, volumes I & II, 1987)

D.1 Background

In this Appendix we review and offer a brief critique of the well known log-polar transform which was ignored in the theoretical approach presented in the thesis, even though this important transform is often invoked as a model of the geometrical transformations imposed by the connectivity from LGN to V1, as noted in the often cited work of Schwartz (1980). In addition to its apparent relevance in early visual pathways, log-polar transform has found a useful application in several neural networks for 2D shape and 3D object recognition (reviewed in Chapter 1).

As a comparison with the log-polar transform method, below we also discuss the alternate solution to size and rotation invariant recognition in the primate visual brain, i.e., the concept of parallel frames of reference in the bottom-up and the top-down pathways that was introduced in Chapter 8.

D.2 Non-linear retino-cortical transformation

It is known that there is an intimate connection between the non-uniform sampling of the image on the retina and the resulting cortical magnification factor as measured directly on the primary visual cortex (Virsu and Rovamo, 1979. Thus the representation of the fovea on the cortex is larger than that of the periphery. It was found that retino-cortical projection can be approximated by a logarithmic transformation (Schwartz, 1980). It can therefore be argued that because this non-linearity is a non-affine transformation about a unique origin (the optical line of sight) that the projected 2D shape of a 3D object will very much depend on its location with respect to the line of sight. This would then imply that there should be a gross deformations of the shape when it is viewed in the periphery. Gross deformations of the shape in the periphery would also invalidate the biological plausibility of the neural model of translation invariant 2D shape recognition as proposed in Chapter 6 (and a large number of related and published models). The model of translation invariant 2D shape recognition as proposed in this thesis, although largely based on the psychophysical data that was reviewed in Chapter 2, has ignored the problem that is associated with the non-linear retinal sampling strategy and the curvature of the sensory surface, primarily because it is aimed for artificial neural vision and 2D shape recognition systems that use a planar sensory surface, such as a standard TV camera.

Since the afore-mentioned non-linearity should introduce distortions of the visual world in our periphery, it is natural to question why is it that other than reduced resolution, we don't see such distortions in our peripheral vision?

Furthermore, the various psychophysical experiments on visual spatial attention and attentional cuing (reviewed in Chapter 2) has clearly shown that humans are able to recognize shapes and objects away from their periphery without having to consciously perform any complex transformations on the visual data. The reader may verify this by directing their fovea to the dot shown in Figure D.1 while attempting to recognize the shown character from a normal reading distance (without eye movements away from the dot).

We are therefore prompted to ask whether the observed non-linear transformation in the retino-cortical pathways is there for any other useful reason than to facilitate the application of the log-polar transform at higher levels of vision.

. **A**

FIGURE D.1. A simple test problem designed to show that we can recognize a shape in our periphery without noticing any significant distortions of the shape (although the recognition accuracy does depend on the size of the shape and its location; the further it is in the periphery the larger it needs to be for successful recognition).

In order to obtain a large field of view (about 200 degrees in human vision), the evolution of the retina and the cortex had to overcome the major problem of how to deal (at higher levels of vision) with visual information that is sensed through a curved sensory surface. It is therefore valid to assume that the nonlinear mapping from the retina to the cortex is there primarily to minimise (or annul) the distortions that would otherwise be definitely seen because of the curved nature of the retina. This would then allow higher brain regions to store internal neural representation of the external visual world in such a manner as to allow a one-to-one correspondence between the visual input and its memory, as well as enabling the cognitive neural centres to covertly shift attention to recognize visual stimuli in peripheral locations.

Fast covert attentional shifts without eye movements to detect and recognize a stimulus away from direct fixation would definitely provide a survival advantage to those organisms that are endowed with it, compared to those that constantly had to direct their fovea to all possible objects and predators in the scene. While eye motions may often be required, for example to provide a more detailed information though the central part of vision (the fovea) they also come at a cost. Motion of an eye is a physical motion of a sensor which requires motor control and generally takes much longer than covert attentional shifts. The latter does not involve any physical movement since it is a dynamic redirection of the neural signals, but it comes at the cost of providing the recognition system with information that is of lower quality (i.e., of lower resolution). This most primitive invariance surely had to be established before other more complex and demanding operations such as mental rotations.

Considering the benefit of a large field of view, which therefore imposes curvature on the retina; the need for fast covert attentional shifts to peripheral locations; the psychophysical data on visual spatial attention and the observation that we don't see a distorted world and objects in our periphery leads us to conclude that the non-linear transformation in the retino-cortical pathways has evolved to enable stimulus recognition in the periphery, i.e., translational invariance.

The purpose of the non-linear retino-cortical transformation (and the validity of the log-polar transform in neurobiology) has also been questioned recently by several neuroscientists (Oram and Perrett, 1994). These authors also note that translational invariance is a very primitive operation which needed to be established very early in the visual system (and before the size and the rotational invariance). Direct application of a non-affine transformation such as the log-polar transform (which is not the same as the application of a non-uniform sampling strategy of the retina) to a TV image would definitely prevent translational invariance. However, if one were to design a camera with a curved sensory surface, then one would definitely need to first eliminate distortions by applying the relevant non-affine transformation (to cancel out the effect of a curved sensory surface) before translational invariance can be achieved.

In the next section we present a case against the biological plausibility of the log-polar transform for size and orientation invariant 2D shape recognition.

D.3 The case against the log-polar transform

Although the psychophysical and neurophysiological data does not reveal the exact nature of processing by which the primate brain achieves the capability of size and orientation invariant 2D shape recognition, it certainly provides sufficient evidence contrary to what the log-polar transform implies. Below we present two simple examples that when combined should expose the inadequacy and biological implausibility of the log-polar transform, particularly in dealing with the rotation invariant 2D shape recognition. Since the psychophysical data on rotation invariant recognition is far more extensive than that on size invariant recognition, the arguments presented below will therefore concentrate on this problem (although similar arguments can be used for size invariant recognition).

Let us consider a familiar and well studied problem of character recognition. Consider the following four lower case characters: 'b', 'd', 'p', and 'q':

b d p q

Figure D2. Example of characters that cannot be distinguished by phase insensitive pattern recognition systems.

Each of the above shapes represents a distinct entity in many written languages, including English. Each thus has a unique interpretation and therefore distinct visual representation in our memory. The shape of each of the above characters not only has the same centroid, but each shape is related to others via a simple geometrical transformation of rotation, reflection or a combination of both. Therefore, if one were to apply

the log-polar transform to each shape, followed by the centering in the transformed feature domain, then each of the shapes will have an exactly the same representation. This implies that any 2D shape recognition system that uses the log-polar transform (or any other invariant feature set) will not be able to distinguish between them. Thus, while providing a method of achieving rotation invariant recognition, the log-polar transform discards some non-invariant information since phase information is discarded, as was also noted by others (Carpenter, Grossberg and Lesher, 1992). Hence any approach that achieves orientation invariant recognition of 2D shapes by ignoring the phase information (i.e. orientation) by definition, cannot store different representations for the above character set (as well as many other examples).

The second argument against the log-polar transform is a bit more involved and requires knowledge of the psychophysical data (from humans). The data that we wish to invoke concerns the influence of shape familiarity on the recognition speed. This data (Koriat and Norman, 1985; Jolicoeur, Snow and Murray, 1987) shows that the strength of memory strongly influences how fast a disoriented shape will be recognized. Of particular interest here is also the data of Cooper and Shepard (1973) which shows that the recognition time for disoriented characters is symmetrical about 180 degrees. This data poses an intriguing paradox. If our recognition time depends on the angle of orientation of the presented shape and if our brain performs mental rotations, then how does the brain know which direction of mental rotation (clock-wise or anti-clock-wise) will produce the faster recognition response? Psychophysical experiments (Jolicoeur, 1990) also indicate that while in a sequential presentation of disoriented characters, the reaction time to the very first character is longer than it is for its upright version but the reaction time to subsequent characters (even if they are different to the ones preceding it) is reduced as long as their orientation is congruent to the ones preceding it. When all of this data is combined, then a clear picture emerges: the visual **brain does perform mental rotations**; the brain **does know in which direction** to perform the mental rotation so as to achieve the fastest recognition time; and the **magnitude and the direction of mental rotation does depend on the stored representation**. This is in direct conflict with the log-polar approach (or any other approach that does not preserve the phase of orientation) because such methods can predict the orientation of the shape (with respect to the chosen frame of reference) regardless of whether that shape has been previously stored in memory or not.

In the next and the final section of the appendix we consider the implication of the massively parallel solution that was briefly outlined in Chapter 8.

D.4 Massively parallel and competing frames of reference

One way (and possibly the only way) that our visual brain can store different representations for shapes that are related via a geometrical transformation of rotation and/or reflection and which would also be consistent with the available psychophysical data, is for our visual brain to have multiple and competing frames of reference through which the visual input can be analysed and interpreted. Although the concept of parallel frames of reference has been proposed previously (Hinton, 1981; Hinton and Parsons, 1981) it has never been considered seriously because: (i) it implies an extremely massive implementation; and (ii) the solution that was proposed does not solve the problems entirely nor does it fully explain the psychophysical data, primarily because it only suggests massively parallel frames in the bottom-up direction. It is not sufficient to have parallel frames of reference in the bottom-up direction because that creates a problem as to how is it that we are able to recognize any of the shapes when they appear in an orientation other than their usual upright, while still being able to differentiate between them in their upright version. However, the novelty of the proposal in Chapter 8 is that there be a complementary frames of reference in the top-down pathways, including a layer of competitive orientation (and size) control neurons, each of which is responsible for the gating of the information flow though a pair of complementary bottom-up and top-down neural layers. Such an approach has the potential to explain and unify more data than any of the alternatives proposed in the literature so far, as will be revealed in the following paragraphs.

If we assume that there are many possible interpretations that may be used (as implied by the multiple frames of reference in the bottom-up and the top-down pathways), then the context within which one sees various shapes and characters should play an important part in selecting the most relevant interpretation. With reference to Figure D.2 and without any other prior knowledge of the context, then the first of the shapes ('b') can be interpreted as follows; (i) character 'b' in its upright orientation; (ii) character 'd' reflected about a vertical axis; (iii) character 'p' upside down and reflected about a vertical axis; or (iv) character 'q' rotated by 180 degrees. Similar transformation links the shapes of the other three characters. The correct interpretation would then be influenced by other information. For example, if the word "darallel" is to make any sense in English, then what should have been typed is "parallel". Hence the correct frame of reference through which the first character in the word is to be interpreted is the one that takes "d" and produces "p". The simple mental rotation of "d" by 180 degrees would then be the correct operation. The orientation of the other characters and their combined meaning thus adds context which quickly narrows the options as to how the viewed character should be interpreted. However, if one sees a partial word (such as

"dar ") without knowing how many other characters are to follow, then the search space for the correct frame of reference is much larger (i.e., one cannot conclude from the available information whether the first character is 'p' typed as 'd', or whether it should be a 'd', etc.).

Some predictions and experimental support

We now make two predictions that arise from the hypothesis of multiple BU-TD (bottom-up and top-down) frames of reference, and relate these to the experimental data. We then follow by a theoretical explanation of what is it that we are aware of during the process of a mental orientation and why.

1. Massive size of the visual brain

The first prediction of the massively parallel bottom-up and top-down frames of reference theory is that whatever region in the brain is responsible for size and orientation invariant 2D shape recognition, it is necessarily very massive in size, it should be composed of several distinct modules and should be easy to locate anatomically. The lesion and other experimental data on the area TEO indicate not only that this area is involved in perceptual constancy but that is reasonably large and consists of a number of distinct anatomical foci.

2. Difficulty in learning some shapes without supervision

The second prediction is that humans (and animals) should have a relative difficulty in learning those 2D shapes that are related via simple geometrical transformations of rotation and/or reflection (which implies that there should be a need for external supervision to initially enable the learning of some shapes). There is some evidence that young children initially confuse characters whose shapes are related by a simple geometrical transformation, most commonly the left-right inversion (the author has observed this with his own children during their early education in reading and writing). The experimental data (Gross, 1978; Holmes and Gross, 1984) also shows that TEO lesioned animals find some pattern discrimination tasks, such as being able to discriminate between two identical patterns (one being the reflected version of the other) relatively easier than non-operated animals. These results indicate that non-operated animals tend to treat mirror-reflected pair of patterns as being the same (and therefore are confused when the task context requires that they be interpreted as being different). On the other hand, operated animals whose area TEO is destroyed are no longer able to recognize that the two mirror-reflected patterns are the same and therefore treat them as being different and hence store in their memory a different representation for each.

Why are we aware of the top-down but not the bottom-up mental rotation?

The theoretical proposal of multiple and competing frames in both the bottom-up and the top-down directions (whose selection may be influenced by a prior memory) invites us to ask why is it that when recognizing an oriented shape we are only aware of the top-down and not the bottom-up mental rotation. Consider the following example: suppose that we are presented with a familiar character that is oriented clockwise from its usual upright orientation. It can be easily verified that during the process of recognising such a character we are only aware of the clockwise mental rotation (a top-down rotation), although an anti-clockwise bottom-up rotation of the character is required to activate the stored memory. How can we explain that? The theory of massive competing frames of reference, when combined with ART's principle of resonance can provide a simple insight into what is going on. The principle of resonance, which in its most glorious application, can be used to argue that we are immediately aware of what is it that we recognize through our vision (or any other sensory modality) corresponds to the very instant that a detectable resonant state is established between the sensory input and the past memories. The theoretical explanation of what happens during the process of recognising a clockwise oriented shape can then be summarised in the following few steps:

(i) Although all possible orientations are initially available to activate the stored memory in the bottom-up pathways, the memory can only be activated by those bottom-up representations whose active pathways are sufficiently well aligned with the stored bottom-up memory pathways;

(ii) When the stored memory is activated via the bottom-up pathways, various top-down representations become available and are tested with the input;

(iii) Since the best matching top-down representation is the one that is in the opposite orientation to the best matched bottom-up representation, the two complimentary orientational representations mutually support each other and competitively suppress all other active representations via their control neurons;

(iv) Because the competition strength between different orientational representations is biased (such that it grows with the angular difference), the order of suppression (being also dependant on the degree of mismatch) will be such that highly mismatched orientations will be the first to lose the competition, followed by the next highest mismatched orientation, etc., until the winner is selected; in the context of the above example, the worst matching top-down representations are the ones that are oriented in the opposite direction to the input, followed by their nearest clockwise oriented neighbours and will therefore be the first to be shut off, followed by the next nearest clockwise oriented representation and so on, until the winner emerges whose clockwise

top-down orientational representation provides a best match to the bottom-up input and whose bottom-up representation best matches the orientation of the stored bottom-up memory;

(v) The gradual shutting off of the top-down representations in the clockwise direction will be sensed as a internal clockwise rotation by other neural circuits within the brain that are specialised for detecting rotations; such a circuit will be activated by the order in which the neural activity spreads across the neurons that are involved in the gating of the various orientational representations; if this activity can resonate for a period long enough for it to be detected by other circuits, then the brain becomes aware of the internal clockwise rotation of its activated memory.

The above explanation also implies that there should be an interference between the viewing of an external rotation of an object and an internal mental rotation. That is, viewing a clockwise rotating object should also activate an imagined (top-down mental) rotation, thus skewing the symmetry of the reaction times. The psychophysical data does show such an effect. Character rotation in the direction of mental rotation facilitates mental rotation, whereas rotation in the opposite direction inhibited it (Jolicouer and Cavanagh, 1992). The data from a related experiment, where the subjects were first exposed to a rotating disk and then had to judge whether the displayed characters were oriented normally or not, shows a strong interaction between the perceived motion aftereffect and the orientation of the character (Corbalis and McLaren, 1982). The perceived motion aftereffect (which was opposite to that of the observed rotation) did skew the latency of the response about 180 degrees, the direction of which depended on whether the rotational aftereffect was in the same or the opposite direction to the orientation of the presented character.

The interaction between observed external motion such as a rotation, scale change. etc., and imagined motion is understandable considering that the learning of the internal processes and the establishment of the necessary synaptic connections between the various neural layers does depend on first being able to see (or feel?) such motions.

D.5 Conclusion

Despite the lack of sufficient and conclusive data from neurophysiology, our own experience and capability in the visual world of real objects and written text, combined with the extensive psychophysical data, provides a very strong base for reasons to believe the existence of multiple frames of reference in our brain through which we analyse and interpret what we see. The intelligence and the survival of a species does not only depend on it being able to recognize an object from various orientations but also on it being able to recognize the pose of the object. Humans are particularly good

at adopting to new frames of reference and in combining past knowledge and experience with context to decide upon the most suitable frame of reference through which the visual (or any other) sensory data should be interpreted. It is possible that a theoretical solution to size and orientation invariant recognition has eluded various theoreticians because they did not view the available data from visual psychophysics through the most relevant paradigm or frame of reference. This thesis has attempted to show that when the same data is analysed with Adaptive Resonance Theory, then it does resolve a number of paradoxes as it immediately invites us to consider massively parallel frames of reference. Although the initial theory did not propose the existence of a vast number of parallel pathways through which the data can be analysed simultaneously in both the bottom-up and the top-down directions, it certainly did not exclude such a possibility.

The ultimate goal of any theoretical approach to the understanding of brain mechanisms is to compress a large body of experimental data into a manageable set of laws while at the same time establishing itself as a new theoretical framework or a window through which the future experimental data may be understood. The success of the theory is then the measure of the volume of the data it can unify, explain and predict. Although the theoretical solution implied in this thesis will need to be further developed and proven in a simulated artificial neural network, the fact that its massive neural network implementation poses a technological problem should not prevent us from seeking out those theoretical solutions that in the longer term will prove to be more robust, more fault tolerant and more flexible.

Bibliography

Anderson, C.H. and Van Essen, D.C. (1987). Shifter circuits: A computational strategy for dynamic aspects of visual processing. *Proc. Natl. Acad. Sci. USA* **84**, pp. 6297-6301.

Bachelder, I.A. and Waxman, A.M. (1994). Mobile robot visual mapping and localization: A view-based neurocomputational architecture that emulates hippocampal place learning. *Neural Networks*, **7**, 1083-1099.

Ballard, D.H. and Brown, C.M. (1982). **Computer Vision**. Englewood Cliffs, NJ: Prentice Hall.

Baloch, A.A. and Waxman, A.M. (1991). Visual learning, adaptive expectations, and behavioural conditioning of the mobile robot MAVIN. *Neural Networks*, **4**, 271-302.

Banquet, J-P. and Grossberg, S. (1987). Probing cognitive processes through the structure of even-related potentials during learning: An experimental and theoretical analysis. *Applied Optics*, **26**, 4931-4946.

Barchas, J.D., Akil, H., Elliot, G.R., Holman, R.B. and Watson, S.J. (1978). Behavioural neurochemistry: Neuroregulators and Behavioural States. *Science*, **200**, 964-973.

Barnard, E. and Casasent, D. (1990). Shift invariance in neocognitron. *Neural Networks*, **3**, 403-410.

Bauer, R.H. and Fuster, J.M. (1976). Delayed-matching and delayed-response deficit from cooling dorsolateral prefrontal cortex in monkeys. *Journal of Comparative Physiological Psychology*, **90**, 293-302.

Baylis, G.C. and Rolls, E.T. (1987). Responses of neurons in the inferior temporal cortex in short term and serial recognition memory task. *Experimental Brain Research*, **65**, 614-622.

Bekkers, J.M. and Stevens, C.F. (1990). Computational implications of NMDA receptor channels. *Cold Spring Harbour Symposium in Quantitative Biology*, **55**, 131-135.

Bender, D.B. and Butter, C.M. (1987). Comparison of the effects of superior colliculus and pulvinar lesions on visual search and tachistoscopic pattern discrimination in monkeys. *Experimental Brain Research*, **69**, 140-154.

Bhara, O., Dale, N., Hochner, B., Klein, M., Abrams, T.W. and Kandel, E.R. (1990). Second messengers involved in the two processes of presynaptic facilitation that contribute to sensitization and dishabituation in *Aplysia* sensory neurons. *Proceedings of the National Academy of Sciences, U.S.A.*, **87**, 2040.

Bisiach, E., Capitani, E., Luzzatti, C., and Perani, D. (1981). Brain and conscious representation of outside reality. *Neuropsychologia*, **19**, 543-551.

Bisiach, E., and Luzzatti, C. (1978). Unilateral neglect of representational space. *Cortex*, **14**, 129-133.

Bodamer, J. (1947). Die Prosop-Agnosie. *Archiv für Psychiatrie und Nervenkrankheiten*, **179**, 6-53.

Bohme, G.A., Bon, C., Stutzman, J.M., Doble, A. and Blanchard, J.C. (1991). Possible involvement of nitric oxide in long-term potentiation. *Eur. J. Pharmacol.*, **199**, 379-381.

Boussaoud, D., Desimone, R. and Ungerleider, L.G. (1991). Visul topography of area TEO in the macaque. *Journal of Comparative Neurology*, **306**, 554-575.

Bradsky, G. and Grossberg, S. (1995). Fast-learning VIEWNET architecture for recognizing three-dimensional objects from multiple two-dimensional views. *Neural Networks*, **Vol. 8**, No. 7/8, pp. 1053-1080.

Braun, J. (1994). Visual search among objects of unequal salience: Removal of selective attention mimics a lesion in extrastriate area V4. *Journal of Neuroscience*, **14**, 554-567.

Bredt, D.S., Hwang, P.H. and Snyder, S.H. (1990). Localization of nitric oxide synthase indicating a neural role for nitric oxide. *Nature*, **347**, 768-770.

Breitmeyer, B.G. (1978). Disinhibition in metacontrast masking of vernier acuity targets: Sustained channels inhibit transient channels. *Vision Research*, **18**, 1401-1405.

Breitmeyer, B.G. (1980). Unmasking visual masking: A look at the "why" behind the veil of the "how". *Psychological Review,* **87**, 52-69.

Breitmeyer, B.G. and Ganz, L. (1976). Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing. *Psychological Review,* **83**, 1-36.

Breitmeyer, B.G. and Valberg, A. (1979). Local, foveal inhibitory effects of global, peripheral excitation. *Science,* **203**, 463-465.

Broadbent, D.E. (1958). **Perception and Communication**. London: Pergamon Press.

Carew, T.J. (1987). Cellular and molecular advances in the study of learning in Aplysia. In: Changeux, J.P. and Konishi, M. (Eds.), **The Neural and Molecular Bases of Learning**. John-Wiley & Sons Limited, 177-204.

Carpenter, G.A. and Grossberg, S. (1981). Adaptation and transmitter gating in vertebrate photoreceptors. *Journal of Theoretical Neurobiology*, **1**, 1-42.

Carpenter, G.A. and Grossberg, S. (1987a). A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics, and Image Processing*, **37**, 54-115.

Carpenter, G.A. and Grossberg, S. (1987b). ART-2: Self-organization of stable pattern recognition codes for analog input patterns. *Applied Optics*, **26**, 4919-4930.

Carpenter, G.A. and Grossberg, S. (1989). Search mechanisms for Adaptive Resonance Theory (ART) architectures. *Proceedings of the International Joint Conference on Neural Networks*, June 18-22, Washington, DC, pp. I 201-205.

Carpenter, G.A. and Grossberg, S. (1990). ART-3: Hierarchical search using chemical transmitters in self-organizing pattern recognition architectures. *Neural Networks*, **3**, 129-152.

Carpenter, G.A. and Grossberg, S. (1993). Normal and amnesic learning, recognition and memory by a neural model of cortico-hippocampal interactions. *Trends in Neuroscience*, **16**, 131-137.

Carpenter, G.A., Grossberg, S. and Lesher, G.W. (1992). A What-and-Where Neural Network for Invariant Image Processing. *IJCNN International Joint Conference on Neural Networks*, Baltimore, **Vol III**, pp. 303-308.

Carpenter, G.A., Grossberg, S. and Reynolds, J.H. (1991). ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network. *Neural Networks*, **4**, 565-588.

Chalupa, L.M. (1991). Visual function of the pulvinar. In: Leventhal, A.G. (Ed.), **Vision and Visual Dysfunction: The Neural Basis of Visual Function**, **4**, 140-159.

Changeux, J-P. (1993). Chemical signalling in the brain. *Scientific American*, Nov., 30-37.

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

Chow, K.L. (1954). Effects of temporal neocortical ablation on visual discrimination learning sets in monkeys. *Journal of Comparative and Physiological Psychology*, **47**, 194-198.

Churchland, P.S. (1986). **Neurophilosophy: Toward a Unified Science of the Mind/Brain**. Cambridge, MA: MIT Press.

Churchland, P.S., Ramachandran, V.S. and Sejnowski, T.J. (1994). A critique of pure vision. In: Koch, C. and Davis, J.L. (Eds.), **Large-Scale Neuronal Theories of the Brain**, The MIT Press, pp. 23-60.

Clark, M., Bovik, A.C. and Geisler, W.S. (1987). Texture segmentation using Gabor modulation/demodulation. *Pattern Recognition Letters*, **6**, 261-267.

Cohen, M.A. and Grossberg, S. (1983). Absolute stability of global pattern formation and parallel memory storage by competitive neural networks. *IEEE Transactions on Systems, Man, and Cybernetics*, **13**, 815-826.

Cohen, M.A. and Grossberg, S. (1984). Neural dynamics of brightness perception: Features, boundaries, diffusion, and resonance. *Perception and Psychophysics*, **36**, 428-456.

Collingridge, G.L. (1987). The role of NMDA receptors in learning and memory (news and views). *Nature*, **330**, 604-605.

Colombo, M. and Gross, C.G. (1994). Responses of inferior temporal cortex and hippocampal neurons during delayed matching to sample in monkeys (Macaca fascicularis). *Behavioural Neuroscience,* **108**, 443-455.

Constantine-Paton, M. (1990). NMDA receptor as a mediator of activity-dependent synaptogenesis in the developing brain. *Cold Spring Harbour Symposium in Quantitative Biology*. **55**, 431-443.

Cooper, L.A. and Shepard, R.N. (1973). Chronometric studies of the rotation of mental images. In W.G. Chase (Ed.), *Visual information processing*, pp. 75-176. New York: Academic Press.

Corbalis, M. (1988). Recognition of Disoriented Shapes. *Psychological Review*. **Vol. 95**, No. 1, 115-123.

Corbalis, M. and McLaren, R. (1982). Interaction Between Perceived and Imagined Rotation. *Journal of Experimental Psychology: Human Perception and Performance*. **Vol. 8**, No. 2, 215-224.

Cowey, A. and Gross, C.G. (1970). Effects of foveal prestriate and inferotemporal lesions on visual discrimination by rhesus monkeys. *Experimental Brain Research*, **11**,128-144.

Crick, F. and Koch, C. (1990a). Some reflections on visual awareness. *Cold Spring Harbour Symposium in Quantitative Biology*. **55**, 953-962.

Crick, F. and Koch, C. (1990b). Towards a neurobiological theory of consciousness. *Seminars in Neurosciences*, **2**, 263-275.

Crick, F. and Koch, C. (1992). The problem of consciousness. *Scientific American*, **267**, 110-117.

Daugmen, J.G. (1985). Uncertainty relation for resolution in space, spatial frequency and orientation optimized by two-dimensional visual cortical filters. *Journal of Optical Society of America*, **2**, 1160-1169.

Daugman, J.G, and Kammen, D.M. (1987). Image statistics, gases, and visual primitives. *ICNN 1987*, **Vol. IV**, 163-175.

Daugman, J.G. (1988). Complete discrete 2-D Gabor transform by neural networks for image analysis and compression. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **36 (7)**, 1169-1179.

Delacour, J. (1977). Role of temporal lobe structures in visual short-term memory, using a new test. *Neuropsychologia*, **15**, 681-684.

De Renzi, E. (1982). **Disorders of space exploration and cogntion.** Wiley, Chichester.

Desimone, R. (1992). Neural circuits for visual attention in the primate brain. In: Carpenter, G. and Grossberg, S. (Eds.), **Neural Networks for Vision and Image processing**. A Bradford Book, The MIT Press, pp. 343-364.

Desimone, R., Albright, T. D., Gross, C.G. and Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience,* **4**, 2051-2062.

Desimone, R. and Gross, C.G. (1979). Visual areas in the temporal cortex of the macaque. *Brain Research*, **178**, 363-380.

Desimone, R. and Ungerleider, L.G. (1989). Neural mechanisms of visual processing in monkeys. In: Boller, F. and Grafman,J. (Eds.), **Handbook of Neuropsychology, II,** Elsevier, Amsterdam.

Desimone, R., Wessinger, M., Thomas, L. and Schneider, W. (1990). Attentional control of visual perception: Cortical, and subcortical mechanisms. *Cold Spring Harbour Symposium in Quantitative Biology*, **55**, 963-971.

Deutsch, J.A. and Deutsch, D. (1963). Attention: some theoretical considerations. *Psychological Review*, **70**, 80-90.

DeYoe, E.A. and Van Essen, D.C. (1987). Concurrent processing streams in monkey visual cortex. *Trends in Neurosciences*, **11**,219-226.

Dirac, P.A.M. (1935). **The Principles of Quantum Mechanics**. Oxford.

Du Buf, J.M.H. (1990). Gabor phase in texture discrimination. *Signal Processing*, **21**, 221-240.

Eccles, J.C. (1964). **The Physiology of Synapses**, Springer-Verlag, Berlin.

Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M. and Reitbock, H.J. (1988). Coherent oscillations: A mechanism of feature linking in the visual cortex? *Biological Cybernetics*, **60**, 121-130.

Eckhorn, R., Reitboeck, H.J., Arndt, M. and Dicke, P. (1990). Feature linking via synchronization among distributed assemblies: The simulation of results from cat visual cortex. *Neural Computation*, **2**, 293-307.

Ellias, S.A. and Grossberg, S. (1975). Pattern formation, contrast control, and oscillations in the short term memory of shunting on-center off-surround networks. *Biological Cybernetics*, **20**, 69-98.

Enroth-Cugell, C. and Robson, J.G. (1966). The contrast sensitivity of retinal ganglion cells of the cat. *Journal of Physiology*, **187**, 517-552.

Epstein, W. and Lovitts, B.E. (1985). Automatic and attentional components in perception of shape-at-a-slant. *Journal of Experimental Psychology: Human Perception and Performance*, **1**, 355-366.

Eriksen, C.W. (1988). Attentional search of the visual field. In D. Brogan (Ed.)., **Visual Search**, *Proceedings of the First International Conference on Visual Search*, University of Durham, England, September 5-9, Taylor & Francis, London.

Eriksen, B.A. and Eriksen, C.W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception and Psychophysics*, **16**, 143-149.

Eriksen, C.W. and Schultz, D. (1979). Information processing in visual search: A continuous flow conception and experimental results. *Perception and Psychophysics,* **25**, 249-263.

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

Eriksen, C.W. and Yeh Y-Y. (1985). Allocation of attention in the visual field: *Journal of Experimental Psychology: Human Perception and Performance*, **11**, 583-597.

Eriksen, C.W. and Murphy, T.D. (1987). Movement of attentional focus across the visual field: A critical look at the evidence. *Perception and Psychophysics*, **42**, 299-305.

Eskandar, E.N., Richmond, B.J. and Optican L.M. (1991). Inferior temporal neurons convey information about stimulus patterns and their behavioral significance. *Society for Neuroscience Abstracts*, **17**, 443

Eskandar, E.N., Richmond, B.J. and Optican L.M. (1992a). Role of inferior temporal neurons in visual memory I: Temporal encoding of information about visual images, recalled images, and behavioural context. *Journal of Neuorphysiology*, **68**, 1277-1295.

Eskandar, E.N., Optican L.M. and Richmond, B.J. (1992b). Role of inferior temporal neurons in visual memory II: Multiplying temporal waveforms related to vision and memory. *Journal of Neurophysiology*, **68**, 1296-1306.

Fahy, F.L., Riches, I.P. and Brown, M.W. (1993a). Neuronal activity related to visual recognition memory: long-term memory and the encoding of recency and familiarity information in the primate anterior and medial inferior temporal and rhinal cortex. *Experimental Brain Research,* **96**, 457-472.

Fahy, F.L., Riches, I.P. and Brown, M.W. (1993b). Neuronal signals of importance to the performance of visual recognition memory tasks: evidence from recordings of single neurons in the medial thalamus of primates. In: Hicks, T.P., Molotchnikoff, S. and Ono, T. (Eds.), **Progress in Brain Research**, **95**, Elsevier Science Publishers B.V., 401-416.

Felleman, D.J. and Van Essen, D.C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, **1**, 1-47.

Flaton, K. and Toborg, S. (1989). An approach to image recognition using sparse filtered graphs. *International Joint Conference on Neural Networks (IJCNN-89),* I, 313-320.

Fogel, I. and Sagi, D. (1989). Gabor filters as texture discriminator. *Biological Cybernetics*, **61**, 103-113.

Friedlander, M.J., Fregnac, Y. and Burke, J.P. (1993). Temporal covariance of postsynaptic membrane potential and synaptic input - role in synaptic efficacy in visual cortex. In: T.P. Hicks, S. Molotchnikoff and T. Ono (Eds.), **Progress in Brain Research, 95,** Elsevier Science Publishers B.V., 207-223.

Friedman, D.P. (1983). Laminar patterns of termination of corticocortical afferents in the somatosensory system. *Brain Research*, **273**, 147-151.

Fukushima, K. (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, **36**, 193-202.

Fukushima, K. (1987). Neural network model for selective attention in visual pattern recognition and associative recall. *Applied Optics*, **26** (23), 4985-4992.

Fukushima, K. (1988). Neocognitron: A hierarchical neural network capable of visual pattern recognition. *Neural Networks*, **1**, 119-130.

Fukushima, K. and Miyake, S. (1984). Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position. *Pattern Recognition*, **15**, 455-469.

Fuster, J.M. (1973). Unit activity in prefrontal cortex during delayed-response performance: neuronal correlates of transient memory. *Journal of Neurophysiology*, **36**, 61-78.

Fuster, J.M. (1990). Inferotemporal units in selective visual attention and short-term memory. *Journal of Neurophysiology,* **64**, 681-697.

Fuster, J.M. and Alexander, G.E. (1970). Delayed response deficit by cyrogenic depression of frontal cortex. *Brain Research*, **20**, 85-90.

Fuster, J.M., Bauer, R.H. and Jervey, J.P. (1985). Functional interactions between inferotemporal and prefrontal cortex in a cognitive task. *Brain Research*, **330**, 299-307.

Fuster, J.M. and Jervey, J.P. (1981). Inferotemporal neurons distinguish and retain behaviorally relevant features of visual stimuli. *Science,* **212**, 952-955.

Gabor, D.J. (1946). Theory of Communication. *J. Inst. Electr. Eng.* **93,** 429-457.

Gaffan, D. and Harrison, S. (1986). Visual identification following inferotemporal ablation in the monkey. *The Quarterly Journal of Experimental Psychology*, **38B**, 5-30.

Gally, J.A., Montague, P.R., Reeke, G.N., and Edelman, G.M. (1990). The NO hypothesis: Possible effects of a short-lived, rapidly diffusable signal in the development and function of the nervous system. *Proc. Natl. Acad. Sci. USA*, **87**, 3547-3551.

Gingrich, K.J., Baxter, D.A. and Byrne, J.H. (1988). Mathematical model of cellular mechanisms contributing to presynaptic facilitation. *Brain Res. Bull.*, **21:** 513.

Gingrich, K.J. and Byrne, J.H. (1985). Simulation of synaptic depression, posttetanic potentiation and presynaptic facilitation of synaptic potentials from sensory neurons mediating gill-withdrawal reflex in *Aplysia. Journal of Neurophysiology*, **53**, 652-669.

Goldman-Rakic, P.S. (1987). Circuitry of pre-frontal cortex and the regulation of behaviour by representational memory. In: Blum, F. and Mountcastle, V. (Eds.), **Higher Cortical Function: Handbook of Physiology**. Washington, DC: American Physiological Society, 373-417.

Goldman-Rakic, P.S., Chafee, M., and Friedman, H. (1993). Allocation of function in distributed circuits. In: Ono, T., Squire, L.R., Raichle, M.E., Perrett, D.I. and Fukuda, M. (Eds)., **Brain Mechanisms of Perception and Memory: From Neuron to Behaviour.** Oxford University Press, 445-456.

Goldman-Rakic, P.S., Funahashi, S. and Bruce, C.J. (1990). Neocortical memory circuits. *Cold Spring Harbour Symposium in Quantitative Biology*. **55**, 1025-1038.

Gopal, N., Bovik, A.C. and Ghosh, J. (1990). Multiple channel surface orientation from texture. *Human Vision and Electronic Imaging: Models, Methods, and Applications, SPIE*, **1249**, 366-375.

Gove, A., Grossberg, S. and Mingolla, E. (1995). Brithtness perception, illusionary contours, and corticogeniculate feedback. *Visual Neuroscience*, **12**, 1027-1052.

Gray, C.M. and Singer, W. (1989). Stimulus specific neuronal oscillations in orientation columns of cat visual cortex. *Proceedings of the National Academy of Sciences USA,* **86**, 1698-1702.

Gray, C.M., Konig, P., Engel, A. and Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, **338**, 334-337.

Grenander, U. (1976-81). **Lectures in Pattern Theory I, II and III: Pattern Analysis, Pattern Synthesis and Regular Structures.** Springer-Verlag, Berlin.

Gross, C.G. (1973). Inferotemporal cortex and vision. In: Stellar, E. and Sprague, J.M. (Eds.), **Progress in Physiological Psychology**. New York: Academic, **5**, 77-123.

Gross, C.G. (1973a). Visual functions of inferotemporal cortex. In: Jung, R. (editor), **Handbook of Sensory Physiology, 7: Central Processing of Visual Information**. Berlin: Springer-Verlag.

Gross, C.G. (1978). Inferior temporal lesions do not impair discrimination of rotated patterns in monkeys. *Journal of Comparative Physiological Psychology*, **Vol. 92.**, 1095-1109.

Gross, C.G., Bender, D.B. and Gerstein, G.L. (1979). Activity of inferior temporal neurons in behaving monkeys. *Neuropsychologia*, **17**, 215-229.

Gross, C.G., Rocha-Miranda, E.C. and Bender, D.B. (1972). Visual properties of neurons in inferotemporal cortex of the macaque. *Journal of Neurophysiology*, **35**, 96-111.

Gross, C. G. and Mishkin, M. (1977). The neural basis of stimulus equivalence across retinal translation. In: Harnard, S., Doty, R., Jaynes, J., Goldstein, L. and Krauthamer, G. (Eds.), **Lateralization in the Nervous System**. Academic Press, New York, pp. 109-122.

Grossberg, S. (1968). Some physiological and biochemical consequences of psychological postulates. *Proceedings of the National Academy of Sciences*, **60**, 758-765.

Grossberg, S. (1968b). Some nonlinear networks capable of learning a spatial pattern of arbitrary complexity. *Proceedings of the National Academy of Sciences*, **59**, 368-372.

Grossberg, S. (1969). On the production and release of chemical transmitters and related topics in cellular control. *Journal of Theoretical Biology*, **22**, 325-364.

Grossberg, S. (1972). Neural expectation: Cerebellar and retinal analogs of cells fired by learnable or unlearned pattern classes. *Kybernetik*, **10**, 49-57.

Grossberg, S. (1973). Contour enhancement, short-term memory, and constancies in reverberating neural networks. *Studies in Applied Mathematics*, **52**, 217-257.

Grossberg, S. (1976). Adaptive pattern classification and universal recoding, II: Feedback, expectation, olfaction, and illusions. *Biological Cybernetics*, **23**, 187-202.

Grossberg, S. (1978). A theory of visual coding, memory, and development. In: Leeuwenberg, E. and Buffart, H. (Eds.), **Formal theories of visual perception**. New York: Wiley.

Grossberg, S. (1980). How does a brain build a cognitive code? *Psychological Review*, **87**, 1-51.

Grossberg, S. (1982). **Studies of mind and brain: Neural Principles of learning, perception, development, cognition, and motor control.** Boston: Reidel Press.

Grossberg, S. (1984). Some psychophysiological and pharmacological correlates of a developmental, cognitive, and motivational theory. In: Karrer, R., Cohen, J. and Tueting. P. (Eds.), **Brain and Information: Event Related Potentials, 425**. New York: New York Academy of Sciences, 58-151.

Grossberg, S. (1987). Competitive learning: From Interactive Activation to Adaptive Resonance, *Cognitive Science,* **11**,23-63.

Grossberg, S. (1988). Nonlinear neural networks: Principles, mechanisms, and architectures. *Neural Networks*, **1**, 17-61.

Grossberg, S. and Levine, D.S. (1987). Neural dynamics of attentional modulated pavlovian conditioning: Blocking, inter-stimulus interval, and secondary reinforcement. *Applied Optics*, **26**, 5015-5030.

Grossberg, S. and Mingolla, E. (1985a). Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations. *Perception and Psychophysics*, **38**, 141-171.

Grossberg, S. and Mingolla, E. (1985b). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review*, **92**, 173-211.

Grossberg, S. and Somers, D. (1991). Synchronized oscillations during cooperative feature linking in a cortical model of visual perception. *Neural Networks*, **4**, 453-466.

Haenny, P.E., Maunsell, J.H.R. and Schiller, P.H. (1988). State dependent activity in monkey visual cortex. II. Extraretinal factors in V4. *Experimental Brain Research*, **69**, 245-259.

Haenny, P.E. and Schiller, P.H. (1988). State dependent activity in monkey visual cortex. I. Single cell activity in V1 and V4 on visual tasks. *Experimental Brain Research*, **69**, 225-244.

Haley, J.E., Wilcox, G.L. and Chapman, P.F. (1992). The role of nitric oxide in long-term potentiation. *Neuron*, **8**, 211-216.

Hasselmo, M.E., Rolls, E.T., Baylis, G.C. and Nalwa, V. (1989). Object-centered encoding by face-selective neurons in the cortex in the superior temporal sulcus of the monkey. *Experimental Brain Research*, **75**, 417-429.

Hawking, S.W. (1988). **A brief history of time**. Bantham Books.

Hawkins, R.D., Abrams, T.W., Carew, T.J. and Kandel, E.R. (1983). A cellular mechanism of classical conditioning in *Aplysia*: Activity-dependent amplification of presynaptic facilitation. *Science*, **91**, 400-405.

Hebb, D.O. (1949). **The Organization of Behaviour**. New York: Wiley.

Henderson, J.M. (1991). Stimulus discrimination following covert attentional orienting to an exogenous cue. *Journal of Experimental Psychology: Human Perception and Performance*, **17**, 91-106.

Hinton, G.E. (1981). A parallel computation that assigns canonical object-based frame of reference. In *Proceedings of the Seventh International Joint Conference on Artificial Intelligence*. **Vol. 2**. pp. 683-685. Los Altos, CA: Kaufman.

Hinton, G.E. and Parsons, L.M (1981). Frames of reference and mental imagery. In J. Long and A.D. Baddeley (Eds.), *Attention and performance*, **Vol. 9**, pp. 261-277. Hillsdale, NJ: Erlbaum.

Hochner, B., Klein, M., Schacher, S. and Kandel E.R. (1986). Action-potential duration and the modulation of transmitter release from sensory neurons of *Aplysia* in presynaptic facilitation and behavioural sensitization. *Proceedings of the National Academy of Sciences, U.S.A.*, **83**, 8410-8414.

Hodgkin, A.L. and Huxley, A.F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology*, **117**, 500-544.

Holmes, G. (1919). Disturbances of visual space perception. *British Medical Journal*, **2**, 230-233.

Holmes, E.J. and Gross, C.G. (1984). Effects of inferior temporal lesions on discrimination of stimuli differing in orientation. *The Journal of Neuroscience*, **Vol. 4**, No. 12, 3063-3068.

Hopfield, J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences USA*, **79**, 2554-2558.

Hopfield, J. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences USA*, **81**, 3088-3092.

Hu, M. (1962). Visual pattern recognition by moment invariants. *IEEE Transactions on Informational Theory*, **IT-8**, 179-187.

Hubel, D. and Wiesel, T. (1959). Receptive fields of single neurons in the cat's striate cortex. *Journal of Physiology (London),* **148**, 574-591.

Hubel, D.H. and Wiesel, T.N. (1962). Receptive fields, binocular interactions and functional architecture in the cat's visual cortex. *Experimental Brain Research,* **22**, 415-419.

Hubel, D.H. and Wiesel, T.N. (1963). Receptive fields of cells in striate cortex of very young, visually inexperienced kittens. *Journal of Neurophysiology*, **26**, 994-1002.

Hubel, D.H. and Wiesel, T.N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, **195**, 215-243.

Hubel, D.H. and Wiesel, T.N. (1977). Ferrier Lecture. Functional structure of macaque visual cortex. *Proceedings of the Royal Society of London B 198,* 1-59.

Humphrey, N.K. and Weiskrantz, L. (1969). Size constancy in monkeys with inferotemporal lesions. *The Quarterly Journal of Experimental Psychology*, **21**, 225-238.

Iwai, E. and Mishkin, M. (1968). Extrastriate visual focus in monkeys: Two visual foci in the temporal lobe of monkeys. In: Yoshii, N. and Buchwald, N.A. (Eds.), **Neurophysiological Basis of Learning and Behaviour**. Osaka: Osaka University, pp. 23-33.

Iwai, E. and Mishkin, M. (1969). Further evidence of the locus of the visual area in the temporal lobe of the monkey. *Experimental Neurology*, **25**, 585-594.

James, W. (1890) **The Principles of Psychology** (Volumes I and II). New York: Holt.

Johnson, J.W. and Ascher, P. (1987). Glycine potentiates the NMDA response in cultured mouse brain neurons. *Nature*, **325**, 529.

Jolicoeur, P. (1990). Orientation Congruency Effects on the Identification of Disoriented Shapes. *Journal of Experimental Psychology: Human Perception and Performance*. **Vol. 16**, No. 2, 351-364.

Jolicoeur, P. and Cavanagh, P. (1992). Mental Rotation, Physical Rotation, and Surface Media. *Journal of Experimental Psychology: Human Perception and Performance*. **Vol. 18**, No. 2, 371-384.

Jolicoeur, P., Snow, D. and Murray, J. (1987). The time to identify disoriented letters: Effects of practice and font. *Canadian Journal of Psychology*, **41**, 303-316.

Jones, J.P. and Palmer, L.A. (1987a). An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, **58** (6), 1233-1258.

Jones, J.P. and Palmer, L.A. (1987b). The two dimensional spatial structure of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, **58** (6), 1187-1211.

Kandel, E.R. (1979). **Behavioral Biology of Aplysia**. Freeman, San Franciso.

Kandel, E.R. and Schwartz, J.H. (1982). Molecular biology of learning: Modulation of transmitter release. *Science*, **218**, 433-443.

Kikuchi, R. and Iwai, E. (1980). The locus of the posterior subdivision of the inferotemporal visual learning area in the monkey. *Brain Research*, **198**, 347-360.

Koch, C. and Crick, F. (1994). Some further ideas regarding the neuronal basis of awareness. In: Koch, C. and Davis, J.L. (Eds.), **Large Scale Neuronal Theories of the Brain**. MIT Press, 93-109.

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

Koenderink, J.J. and van Doorn, A.J. (1979). The internal representation of solid shape with respect to vision. *Biological Cybernetics*, **32**, 211-216.

Koriat, A. and Norman. J. (1988) Frames and Images: Sequential Effects in Mental Rotation. *Journal of Experimental Psychology: Learning, Memory and Cognition*. **Vol. 14**, No. 1, 93-111.

Koriat, A. and Norman, J. (1985). Mental rotation and visual familiarity. *Perception and Psychophysics*, **37**, 429-439.

Kosslyn, S.M. (1980). **Image and Mind**. Harvard University Press, Cambridge, MA.

Kosslyn, S.M., Alpert, N.M., Thompson, W.L., Maljkovic, V., Weise, S.B., Chabris, C.F., Hamilton, S.E., Rauch, S.L. and Buonanno, F.S. (1993). Visual mental imagery activates topographically organized visual cortex: Pet investigations. *Journal of Cognitive Neuroscience*, **5**, 263-287.

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

Le Bihan, D., Turner,R., Jezzard, P., Cuenod, C.A. and Zeffiro. T. (1992). Activation of human visual cortex by mental representation of visual patterns. In: *Book of Abstracts of Work in Progress*, 11th Annual Meeting of the Society for Magnetic Resonance in Medicine, p. 311.

Levitan, I. B. and Kaczmarek, L. K. (1991). **The Neuron: Cell and Molecular Biology.** Oxford University Press.

Lozo, P. (1996a). Dynamic tuning in the Feedforward Excitation-Feedback Presynaptic Facilitation (FFE-FBPF) neural circuit. *Proceedings of the Seventh Australian Conference on Neural Networks (ACNN'96)*, Canberra, ACT, April 10-12, pp. 79-84.

Lozo, P. (1996b). Neural circuit for match/mismatch and familiarity/novelty detection in SAART neural networks. (To appear in *Fourth International Symposium on Signal Processing and its Applications (ISSPA'96)*, Gold Coast, Australia, August 1996).

Lozo, P. (1996c). Neural circuit for self-regulated attentional learning in Selective Attention Adaptive Resonance Theory (SAART) neural networks. (To appear in *Fourth International Symposium on Signal Processing and its Applications (ISSPA'96)*, Gold Coast, Australia, August 1996).

Lozo, P. (1995). Selective Attention Adaptive Resonance Theory (SAART) neural network for neuro-engineering of robust ATR systems. *IEEE International Conference on Neural Networks,* Perth 27 Nov.- 1 Dec., pp. 2461-2466.

Lozo, P. (1993). Variable threshold ART-3 neural network. *Proceedings of the Fourth Australian Conference on Neural Networks (ACNN'93).* Melbourne, 1-3 Feb., pp. 5-9.

Lozo, P. and Lim, C-C. (1996). Neural circuit for object recognition in complex and cluttered visual images. (*Submitted May 1996*).

Lozo, P. and Nandagopal, D.N. (1996). Selective transfer of spatial patterns by presynaptic facilitation in a shunting competitive neural layer. (*Submitted May 1996*).

Lozo, P., Johnson, R.P., Nandagopal, D., Nussey, G. and Zyweck, T. (1991). An Adaptive Resonance Theory (ART) based neural network approach to classifying targets in infrared images. *Proceedings of the Second Australian Conference on Neural Networks,* Sydney, pp. 22-25.

Lozo, P., Lim, C.C., and Nandagopal, D.N. (1995). Translation invariant pattern recognition: Neural network architecture based on biological visual spatial attention. *Australian Journal of Intelligent Information Processing Systems*, **2 (1)**, 1-10.

Lozo, P., Lim, C.C. and Nandagopal, D. (1993). Towards Dynamic Neuro-computational Vision Part 1: Covert and exogenous visual spatial attention neural network with translation invariance. *IEEE Visual Signal Processing and Communications* - Workshop Proceedings, Melbourne 20-22 Sept., pp. 145-149.

Lozo, P., Nandagopal, D. and Lim, C.C. (1993). A self-organising neural network architecture for real-time and stable learning and recognition of translation variant patterns. *Proceedings of the First Australian and New Zealand Conference on Intelligent Information Systems.* Perth, 1-3 December, 10-14.

Lozo, P., Lim, C. and Nandagopal, D. (1994). A neural network model of 2-D visual spatial attention and translation invariant pattern recognition. *Proceedings of the Fifth Australian Conference on Neural Networks*, Brisbane, pp. 84-89.

Lynch, J.C., Mountcastle, V.B., Talbot, W.H. and Yin, T.C.T. (1977). Parietal lobe mechanisms for directed visual attention. *Journal of Neurophysiology,* **40**, 362-389.

Mach, E. (1902). **Science of Mechanics**.

Mach, E. (1914). **The analysis of sensation and the relation of the physical to the psychical**, C.M. Williams (Trans.), revised by S. Waterlow. London: Open Court Publishing.

Marr, D. (1982). **Vision**. W.H. Freeman, San Francisco, CA.

Marr, D. and Hildreth, E.C. (1980). Theory of edge detection. *Proc. R. Soc. Lond.* **B 207**, 187-217.

Marr, D. and Nishihara, K. (1978). Representation and recognition of the spatial organisation of three dimensional shapes. *Philosophical Transactions of the Royal Society (London),* **B200**, 269-294.

Maunsell, J.H.R. and Van Essen, D.C. (1983). The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *Journal of Neuroscience*, **3**, 2563-2586.

McCulloch, W.S. and Pitts, W.H. (1943). A logical calculus of ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, **5**, 115-133.

Menon, M.M. and Heinemann, K.G. (1988). Classification of patterns using a self-organizing neural network. *Neural Networks*, **1**, 201-215.

Mesulam, A.M. (1981). A cortical network for directed attention and unilateral neglect. *Annals of Neurology,* **10**, 309-325.

Miller, E.K., Li, L. and Desimone, R. (1991). A neural mechanism for working and recognition memory in inferior temporal cortex. *Science*, **254**, 1377-1379.

Miller, E.K., Li, L. and Desimone, R. (1993). Activity of neurons in anterior inferior temporal cortex during a short-term memory task. *Journal of Neuroscience,* **13,** 1460-1478.

Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, **14**, 247-279.

Millman, J. (1979). **Microelectronics: Digital and Analog Circuits and Systems**. International Student Edition. McGraw-Hill.

Mishkin, M. (1966). Visual mechanisms beyond the striate cortex. In: R. Russel, R. (ed.), **Frontiers of Physiological Psychology**. Academic Press, New York, 93-119.

Mishkin, M. (1982). A memory system in the monkey. *Philos. Trans. R. Soc. Lond. (Biol.)*, **298**, 83-95.

Miyake, S. and Fukushima, K. (1984). A neural network model for the mechanism of feature extraction. *Biological Cybernetics*, **50**, 377-384.

Miyashita, Y. (1993). Inferior temporal cortex: Where visual perception meets memory. *Annual Review of Neuroscience,* **16**, 245-63.

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

Motter, B.C. (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *Journal of Neurophysiology*, **70**, 909-919.

Mountcastle, V.B. (1978). Brain mechanisms for directed attention. *Journal of the Royal Society of Medicine,* **71**, 14-28.

Mountcastle, V.B., Motter, B.C., Steinmetz, M.A. and Sestokas, A.K. (1987). Common and differential effects of attentive fixation on the excitability of parietal and prestriate (V4) cortical visual neurons in the macaque monkey. *Journal of Neuroscience,* **7**, 2239-2255.

Mumford, D. (1991). On the computational architecture of the neocortex. I. The role of the thalamo-cortical loop. *Biological Cybernetics*, **65**, 135-145.

Mumford, D. (1992). On the computational architecture of the neocortex. II. The role of cortico-cortico loops. *Biological Cybernetics*, **66**, 241-251.

Mumford, D. (1994). Neuronal architectures for pattern-theoretic problems. In: Koch, C. and Davis, J.L. (Eds.), **Large Scale Neuronal Theories of the Brain**. MIT Press, 125-152.

Niebur, E., Koch, C. and Rosin, C. (1993). An oscillation-based model for the neuronal basis of attention. *Vision Research*, **33**, 2789-2802.

O'Dell, T.J., Hawkins, R.D., Kandel, E.R. and Arancio, O. (1991). Tests on the roles of two diffusable substances in LTP: Evidence for nitric oxide as a possible early retrograde messenger. *Proc. Natl. Acad. Sci. USA*, **88**, 11285-11289.

Ögmen, H. and Gagné, S. (1990). Neural network architectures for motion perception and elementary motion detection in the fly visual system. *Neural Networks*, **3**, 487-505.

Olshausen, B.A., Anderson, C.H. and Van Essen, D.C. (1992). A neural model of visual attention and invariant pattern recognition. *CNS Memo 18, Cognitive and Neural System Program*, Division of Biology, California Institute of Technology, Pasadena, CA.

Olshausen, B.A., Anderson, C.H. and Van Essen, D.C. (1993). A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information flow. *Journal of Neuroscience*, **13**, 4700-4719.

Oram, M.W. and Perrett, D.I. (1994). Modelling visual recognition from neurobiological constraints. *Neural Networks 1994 Special Issue Models of Neurodynamics and Behaviour*, Vol. 7., Nos. 6/7, 945-972.

Pallis, C.A. (1955). Impaired identification of faces and places with agnosia for colours. *Journal of Neurology, Neurosurgery and Psychiatry*, **18**, 218-224.

Palmer, S.E. (1983). The psychology of perceptual organization: A transformational approach. In: Beck, J., Hope, B. and Rosenfeld, A. (Eds.), **Human and Machine Vision**. Orlando: Academic, 269-339.

Peli, T. and Malah, D. (1982). A study of edge detection algorithms. *Computer Graphics and Image Processing*, **20**, 1-21.

Perrett, D.I. and Oram, M.W. (1993). Neurophysiology of Shape Processing. *Image and Vision Computing*, **11**,317-333.

Perrett, D.I., Rolls, E.T. and Caan, W. (1982). Visual neurons responsive to faces in the monkey temporal cortex. *Experimental Brain Research*, **47**, 329-342.

Perrett, D.I., Smith, P.A.J., Potter, D.D., Mistlin, A.J., Head, A.S., Milner, A.D. and Jeeves, M.A. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings of the Royal Society of London Series B*, **223**, 293-317.

Petersen, S.E., Robinson, D.L. and Morris, J.D. (1987). Contributions of the pulvinar to visual spatial attention. *Neuropsychologia*, **25**, 97-105.

Porat, M. and Zeevi, Y.Y. (1988). The generalized Gabor scheme for image representation in biological and machine vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **10**, 452-468.

Posner, M.I. (1980) Orienting of attention. *Quarterly Journal of Experimental Psychology*, **32**, 3-25.

Posner, M.I., Nissen, M.J. and Ogden, W.C. (1978). Attended and unattended processing modes: The role of set for spatial location. In: H.L. Pick and I.J. Saltzman (Eds.), **Modes of Perceiving and Processing Information**. Hillsdale, N.J.: Erlbaum.

Posner, M. I., Walker J.A., Friedrich, F.J. and Rafal R.D. (1984). Effects of parietal injury on covert orienting of visual attention. *Journal of Neuroscience*, **4**, 1863-1874.

Posner, M.I., Walker, J.A., Friedrich, F.A. and Rafal, R.D. (1987). How do the parietal lobes direct covert attention? *Neuropsychologia,* **25**, 135-145.

Posner, M.I. and Rothbart, M.K. (1992). Attention and conscious experience. In: Milner, A.D. and Rugg, M.D. (Eds.), **The Neuropsychology of Consciousness**. Academic Press, London, 91-112.

Posner, M.I. and Rothbart, M.K. (1994). Constructing neuronal theories of mind. In: Koch, C. and Davis, J.L. (Eds.), **Large Scale Neuronal Theories of the Brain**. MIT Press, 183-199.

Pratt, W. K. (1978). **Digital Image Processing**. Wiley and Sons.

Rafal, R.D. and Posner, M.I. (1987). Deficits in human visual spatial attention following thalamic lesions. *Proceedings of the National Academy of Science USA,* **84**, 7349-7353.

Riches, I.P., Wilson, F.A.W. and Brown, M.W. (1991). The effects of visual stimulation and memory on neurons of the hippocampal formation and the neighboring parahippocampal gyrus and inferior temporal cortex of the primate. *Journal of Neuroscience,* **11**,1763-1779.

Richmond, B.J., Wurtz, R.H. and Sato, T. (1983). Visual responses of inferior temporal neurons in awake rhesus monkey. *Journal of Neurophysiology,* **50**, 1415-1432.

Richmond, B.J. and Sato, T. (1987). Enhancement of inferior temporal neurons during visual discrimination. *Journal of Neurophysiology,* **58**, 1292-1306.

Riddoch, G. (1917). Dissociation of visual perception due to occipital injuries, with especial reference to appreciation of movement. *Brain*, **40**, 15-57.

Riddoch, M.J. and Humphreys, G.W. (1987). A case of integrative visual agnosia. *Brain*, **110**,1431-1462.

Rizzolati, G., Gentilucci, M., Matelli, M. (1985). Selective spatial attention: One center, one circuit or many circuits. In M. Posner, O.S.M. Marin (Eds.), *Attention and Performance XI*, pp. 251-265. Hillsdale, NJ: Erlbaum.

Robertson, L.C., Lamb, M.R. and Knight, R.T. (1988). Effects of lesions of temporal-parietal junction on perceptual and attentional processing in humans. *Journal of Neuroscience*, **8**, 3757-3769.

Robertson, L.C., Palmer, S.E. and Gomez, L.M. (1987). Reference Frames in Mental Rotation. *Journal of Experimental Psychology: Learning, Memory and Cognition*. **Vol. 13**, No. 3, 368-379.

Robinson, D.L. (1993). Functional contribution of the primate pulvinar. In: **Progress in Brain Research**, **95**, 371-380.

Rock, I. and Gutman, D. (1981). Effects of inattention on form perception. *Journal of Experimental Psychology: Human Perception and Performance,* **7**, 275-285.

Rockland, K.S. and Pandya, D.N. (1979). Laminar origins and terminations of cortical connections of the occipital lobe in the rhesus monkey. *Brain Research*, **179**, 3-20.

Rolls, E.T., Perret, D.I., Caan, A.W. and Wilson, F.A.W. (1982). Neuronal responses related to visual recognition. *Brain*, **105**, 611-646.

Rosenfeld, A. and Kak, A.C. (1982). **Digital Picture Processing**. Academic Press.

Sary, G., Vogels, R. and Orban, G.A. (1993). Cue-invariant shape selectivity of macaque inferior temporal neurons. *Science*, **260**, 995-997.

Sato, T. (1988). Effects of attention and stimulus interaction on visual responses of inferior temporal neurons in macaque. *Journal of Neurophysiology,* **60**, 344-364.

Sato, T., Kawamura, T. and Iwai, E. (1980). Responsiveness of inferotemporal single units to visual pattern stimuli in monkeys performing discrimination. *Experimental Brain Research,* **38**, 313-319.

Schuman, E.M. and Madison, D.V. (1991). A requirement for the intercellular messenger nitric oxide in long-term potentiation. *Science*, **254**, 1503-1506.

Schwartz, E.L. (1980). Computational anatomy and functional architecture of striate cortex: A spatial mapping approach to perceptual coding. Vision Research, Vol. 20, 645-669.

Schwartz, E.L., Desimone, R., Albright, T.D. and Gross, C.G. (1983). Shape recognition and inferior temporal neurons. *Proceedings of the National Academy of Sciences*, **80**, 5776-5778.

Shulman, G.L., Remington, R.R. and McLean, J.P. (1979). Moving attention through visual space. *Journal of Experimental Psychology: Human Perception and Performance*, **5**, 522-526.

Seibert, M. and Waxman, A.M. (1989). Spreading Activation Layers, Visual Saccades, and Invariant Representations for Neural Pattern Recognition Systems. *Neural Networks*, **2**, 9-27.

Seibert, M. and Waxman, A.M. (1990). Learning Graphs Representations from View Sequences; In D. Touretzky (editor). *Advances in Neural Information P(rocessing Systems 2*, pp. 258-265, Morgan Kaufmann.

Seibert, M. and Waxman, A.M. (1992). Adaptive 3D-object recognition from multiple views. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **14**, 107-124.

Seibert, M., Waxman, A.M. and Gove, A.N. (1995). Learning to distinguish similar objects. *Applications and Science of Artificial Neural Networks, SPIE-2492*, pp. 600-611.

Sillito, A.M., Jones, H.E., Gerstein, G.L. and West, D. (1994). Feature-linked synchronization of thalamic relay cell firing induced by feedback from the visual cortex. *Nature*, **369**, 479-482.

Singer, W. (1987). Activity-dependent self-organization of synaptic connections as a substrate of learning. In: Changeux, J.P. and Konishi, M. (Eds.), **The Neural and Molecular Bases of Learning**. John Wiley & Sons Limited, 301-336.

Singer, W. and Bedworth, N. (1973). Inhibitory interactions between X- and Y-units in the cat lateral geniculate nucleus. *Brain Research*, **49**, 291-307.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, **74** (11, Whole no. 498).

Spitzer, H., Desimone, R. and Moran, J. (1988). Increased attention enhances both behavioural and neuronal performance. *Science,* **240**, 338-340.

Stricker, E.M. and Zigmond, M.J. (1976). Brain catecholamines and the lateral hypothalamic syndrome. In D. Novin, W. Wyrwicka, and G. Bray (Eds.), **Hunger: Basic mechanisms and clinical implications.** New York: Raven Press.

Tanaka, K. (1993). Neuronal mechanisms of Object recogntion. *Science*, **262**, 685-688.

Tanaka, K., Fijita, I., Kobotake, E., Cheng, K. and Ito, M. (1993). Serial Processing of Visual Object-Features in the Posterior and Anterior Parts of the Inferotemporal Cortex: In Y. Ono, J.R. Squire, M.E. Raichle, D.I. Perrett and M. Fukuda (Eds.), **Brain Mechanisms of Perception and Memory**, pp 34-46. Oxford University Press.

Ting, C. and Chuang K-E. (1993). An adaptive algorithm for Neocognitron to recognize analog images. ***Neural Networks*, 6,** 285-299.

Tigges, J., Spatz, W.B. and Tigges, M. (1973). Reciprocal point-to-point connections between parastriate and striate cortex in the squirrel monkey (*Saimiri*). *J. Comp. Neurol.* **148**, 481

Treisman, A. and Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, **12**, 97-136.

Tsal, Y. (1983). Movements of attention across the visual field. *Journal of Experimental Psychology: Human Perception and Performance*, **9**, 523-530.

Tsotsos, J.K. (1991). Localizing stimuli in a sensory field using an inhibitory attention beam. *Technical Report, RBCV-TR-91-37*, Dept. of Computer Science, University of Toronto.

Turner, M.R. (1986). Texture discrimination by Gabor functions. *Biological Cybernetics*, **55**, 71-82.

Ullman, S. (1994). Sequence seeking and counterstreams: A model for bidirectional information flow in the cortex. In: Koch, C. and Davis, J.L. (Eds.), **Large Scale Neuronal Theories of the Brain**. The MIT Press, Cambridge, MA., 257-270.

Ungerleider, L.G. and Mishkin, M. (1982). Two cortical visual systems. In: Ingle, D.J., Goodale, M.A. and Mansfield, R.J.W. (Eds.), **Analysis of Visual Behavior**. Cambridge, Mass. MIT Press.

Van der Heijden, A.H.C. (1992). **Selective Attention in Vision.** Routledge.

Van Essen, D.C. and Anderson, C.H. (1990). Information processing strategies and pathways in the primate retina and visual cortex. In: Zornertzer S.F., Davis, J.L. and Lau, C. (Eds.), **An Introduction to Neural and Electronic Networks**. Academic Press, New York, pp. 43-72.

Van Essen, D.C., Anderson, C.H. and Olshausen , B.A. (1994). Dynamic routing strategies in sensory, motor, and cognitive processing. In: Koch, C. and Davis, J.L. (Eds.), **Large Scale Neuronal Theories of the Brain**. MIT Press, 271-299.

Van Essen, D.C. and Maunsell, J.H.R. (1980). Two-dimensional maps of the cerebral cortex. *Journal of Comparative Neurology*, **191**,255-281.

Van Essen, D.C., Olshausen, B., Anderson, C.H., and Gallant, J.L. (1991). Pattern recognition, attention, and information bottlenecks in the primate visual system. *Proceedings of SPIE Conference on Visual Information Processing: From Neurons to Chips*, **1473**, 17-28.

Van Essen, D.C., Felleman, D.J., DeYoe, E.A., Olavarria, J. and Knierim, J. (1990). Modular and hierarchical organization of extrastriate visual cortex in the macaque monkey. *Cold Spring Harbour Symposium in Quantitative Biology*. **55**, 679-696.

Von Helmholtz, H. (1866/1925). **Physiological optics**. J.P.C. Sothall (Ed.), New York: Dover.

Waxman, A.M., Seibert, M. and Bachelder, I.A. (1995). Visual processing of object form and environment layout. In: M. Arbib (Ed.), **The Handbook of Brain Theory and Neural Networks**, pp. 1021-1024. MIT Press.

Waxman, A.M., Seibert, M., Bernardon, A.M. and Fay, D.A. (1993). Neural Systems for Automatic Target Learning and Recognition. *Lincoln Laboratory Journal (Special Issue on Automatic Target Recognition)*, **6**, 77-116.

Weichselgartner, E. and Sperling, G. (1987). Dynamics of automatic and controlled visual attention. *Science*, **238**, 778-780.

Weiskrantz, L. and Saunders, R.C. (1984). Impairments of visual object transforms in monkeys. *Brain*, **107**, 1033-1072.

Weiskrantz, L., Warrington, E.K., Sanders, M.D. and Warshall, J. (1974). Visual capacity in the hemianopic field following a restricted occipital ablation. *Brain*, **97**, 709-728.

Whitfield, I.C. (1984). **Neurocommunication: An Introduction**. John Wiley & Sons.

Zeki, S. (1993). **A Vision of the Brain.** Blackwell Scientific Publications. Oxford.