

# SCAN: a neural model of covert attention

Citation for published version (APA):

Postma, E. O. (1994). *SCAN: a neural model of covert attention*. [Doctoral Thesis, Maastricht University]. Rijksuniversiteit Limburg. <https://doi.org/10.26481/dis.19940922ep>

## Document status and date:

Published: 01/01/1994

## DOI:

[10.26481/dis.19940922ep](https://doi.org/10.26481/dis.19940922ep)

## Document Version:

Publisher's PDF, also known as Version of record

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

**SCAN:  
A Neural Model of Covert Attention**

SCAN:

**SCAN:  
A Neural Model of Covert Attention**

**PROEFSCHRIFT**

ter verkrijging van de graad van doctor aan de Rijksuniversiteit  
Limburg te Maastricht, op gezag van de Rector Magnificus,  
Prof. dr. H. Philipsen, volgens het besluit van het College  
van Dekanen, in het openbaar te verdedigen op  
donderdag 22 september 1994 om 16.00 uur

door

**Eric Onno Postma**

geboren te Nijmegen in 1961

Promotores: Prof. dr. H. J. van den Herik  
Prof. dr. P.T.W. Hudson

Leden van de beoordelingscommissie:

Prof. dr. ir. A. Hasman (voorziter)  
Dr. ir. W.R.M. Dassen  
Prof. dr. A.H.C. van der Heijden (Rijksuniversiteit Leiden)  
Prof. dr. J.J. Koenderink (Universiteit Utrecht)  
Prof. dr. ir. drs. O.J. Vrieze

CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Postma, Eric O.

SCAN : A Neural Model of Covert Attention /

Eric O. Postma ; [ill. by the author]. - [S.l. : s.n.]

(Wageningen : Ponsen & Looijen). - Ill.

Thesis Maastricht. - With references. - With summary in Dutch

ISBN 90-9007320-5

NUGI 855

Subject headings: neural networks / covert attention/ computer vision.

Cover design: Paul Postma

# Contents

<b>Preface</b>	<b>ix</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Covert Attention . . . . .	1
1.2 Multidisciplinary Approach . . . . .	4
1.3 Objectives and Assumptions . . . . .	6
1.4 Thesis Overview . . . . .	8
<b>2 Covert Attention in the Brain</b>	<b>9</b>
2.1 Levels of Organization . . . . .	9
2.2 Neural Processing . . . . .	10
2.2.1 Noisy Processing . . . . .	10
2.2.2 Economical Wiring . . . . .	12
2.2.3 Many Processing Elements . . . . .	12
2.3 Cortical Processing . . . . .	13
2.3.1 The Retina . . . . .	13
2.3.2 The Striate Cortex . . . . .	15

2.3.3	Concurrent Processing . . . . .	17
2.3.4	The Temporal Pathway . . . . .	17
2.3.5	The Parietal Pathway . . . . .	18
2.3.6	Spatial Variance versus Spatial Invariance . . . . .	19
2.4	Neural Mechanisms of Attention . . . . .	20
2.4.1	Covert Attention in the Parietal Pathway . . . . .	20
2.4.2	Covert Attention in the Temporal Pathway . . . . .	21
2.4.3	Featural and Spatial Covert Attention . . . . .	22
2.5	The Psychology of Attention . . . . .	23
2.5.1	The Psychology of Spatial Attention . . . . .	23
2.5.2	The Psychology of Featural Attention . . . . .	25
2.6	Characteristics of a Vision Model . . . . .	26
2.6.1	Stochasticity . . . . .	26
2.6.2	Sparse Connectivity . . . . .	26
2.6.3	Profligacy . . . . .	27
2.6.4	Hierarchy . . . . .	27
2.6.5	Invariance . . . . .	27
2.6.6	Attention: Integration of the Characteristics . . . . .	28
<b>3</b>	<b>The Neural Lattice</b>	<b>31</b>
3.1	Threshold Elements . . . . .	32
3.1.1	Deterministic Threshold Elements . . . . .	32
3.1.2	Stochastic Threshold Elements . . . . .	33
3.1.3	Interacting Threshold Elements . . . . .	34

---

3.1.4	Energy Function . . . . .	34
3.2	Sparse Connectivity Patterns . . . . .	36
3.2.1	Neural Lattices with Nearest-Neighbour Connections . .	36
3.3	Statistical Mechanics of Ferromagnets . . . . .	38
3.3.1	The Phase Transition . . . . .	39
3.3.2	Spontaneous and Field-Induced Magnetization . . . . .	40
3.4	Statistical Mechanics of Neural Lattices . . . . .	42
3.4.1	Neural Lattices as Ising Lattices . . . . .	42
3.4.2	Critical Values and Connection Density . . . . .	47
3.5	Simulation Studies . . . . .	49
3.5.1	Simulations with an Initial Disordered State . . . . .	49
3.5.2	Simulations with an Initial Ordered State . . . . .	56
3.6	Summary and Conclusion . . . . .	60
<b>4</b>	<b>The Gating Lattice</b>	<b>63</b>
4.1	A Parallel Distributed Switch . . . . .	65
4.1.1	Gate Structure . . . . .	66
4.1.2	Gate Coupling . . . . .	67
4.1.3	Sublattices . . . . .	68
4.1.4	Synchronicity and Exclusivity . . . . .	71
4.1.5	Distributed Competition . . . . .	71
4.2	Statistical Mechanics of Antiferromagnets . . . . .	72
4.2.1	The Phase Diagram . . . . .	72
4.2.2	Closed Form of the Critical Curve . . . . .	73

4.3	Statistical Mechanics of Gating Lattices . . . . .	74
4.3.1	Calculation of Gating-Lattice Partition Function . . . . .	76
4.3.2	Gate Dynamics . . . . .	80
4.3.3	Order-Parameter Conserving Gate Dynamics . . . . .	82
4.3.4	Optimal Intrinsic-Noise Level . . . . .	83
4.3.5	Responsiveness to External Inputs . . . . .	84
4.4	Finite Size and Limited Time . . . . .	85
4.5	Simulation Studies . . . . .	87
4.5.1	Metastability in Gating Lattices . . . . .	87
4.5.2	Extrinsic Noise in the Gating Lattice . . . . .	90
4.6	General Discussion . . . . .	95
4.7	Extensions of the Gating Lattice . . . . .	97
4.7.1	Potts Gating Lattice . . . . .	97
4.7.2	Hierarchical Gating Lattices . . . . .	99
4.8	Conclusion . . . . .	101
<b>5</b>	<b>The SCAN Model</b>	<b>103</b>
5.1	An Artificial Searchlight . . . . .	104
5.2	The Structure and Operation of SCAN . . . . .	105
5.2.1	The Gating Network . . . . .	106
5.2.2	The Classifier Network . . . . .	111
5.2.3	The Integrated Model . . . . .	114
5.3	Formal Description of SCAN . . . . .	115
5.3.1	Formal Description of the Gating Network . . . . .	115

5.3.2	Formal Description of the Classifier Network . . . . .	118
5.4	Simulation Studies . . . . .	121
5.4.1	Performance as a Function of Intrinsic Noise . . . . .	121
5.4.2	Scanning a Natural Image . . . . .	124
5.4.3	Performance with Time-Varying Intrinsic Noise . . . . .	129
5.4.4	Improving SCAN: Scanning Scale Space . . . . .	132
5.4.5	Related Work . . . . .	133
5.5	General Discussion . . . . .	134
<b>6</b>	<b>Evaluating SCAN</b>	<b>137</b>
6.1	Gating in the Brain . . . . .	137
6.1.1	Gating by the Interneuron . . . . .	138
6.1.2	Gating by the Interneuron Network . . . . .	138
6.1.3	Gating Network in the Temporal Pathway . . . . .	139
6.1.4	Classification in the Inferotemporal Cortex . . . . .	140
6.1.5	Parallel Matching at V1 and V2 . . . . .	140
6.2	Quantitative Considerations . . . . .	141
6.3	Stable Perception . . . . .	142
6.4	Incompatibilities . . . . .	143
6.4.1	Representation of Visual Data . . . . .	143
6.4.2	Randomness of Individual Cortical Connections . . . . .	145
6.5	Concluding Remarks . . . . .	146
<b>7</b>	<b>Evaluation and Conclusion</b>	<b>147</b>

---

7.1	Evaluation of the Interdisciplinary Approach . . . . .	147
7.2	Evaluation of SCAN . . . . .	151
7.3	General Evaluation and Conclusion . . . . .	151
	<b>References</b>	<b>153</b>
	<b>Summary</b>	<b>173</b>
	<b>Samenvatting</b>	<b>177</b>
	<b>Curriculum Vitae</b>	<b>183</b>

# Preface

Recent insights into the structure and workings of the brain provide a rich source of inspiration for the construction of machines that exhibit the desired characteristics of human performance. SCAN (Signal Channelling Attentional Network), the neural-network model described in this thesis, is inspired by biological and psychological views on visual attention and possesses the flexibility and robustness of the processes underlying attention. Statistical-physics techniques are embodied in computer simulations in order to demonstrate the operation of the SCAN model. A computer-science approach allows both quantitative and qualitative assessments of the neural-network model. SCAN differs from previous similar approaches in that it addresses the *scaling problem*, the fact that real systems are very large and most neural-network models will only work when they are very small. By combining parsimony in wiring with a large number of processing elements, the SCAN model is capable of dealing with images of a realistic size.

My first interest in the subject of visual attention probably goes back to my childhood years, when I wondered why I saw everything out there, but at the same time could only see what I looked at. With hindsight, I was pondering the question of how the subjective experience of a complete overview of the visual world could be reconciled with the limited capacity to attend to its parts. During my study this problem evolved into one about the nature and degree of parallelism in vision. The neural-network paradigm provided the ideal touchstone for my ideas. Thus I have devoted my PhD research to the formulation of a neural-network model of visual attention.

I have been lucky to be surrounded by people who have assisted me in various ways. My special thanks go to Jaap van den Herik, my supervisor, who provided the opportunity to develop my ideas and guided me through the process

of writing a thesis. I also would like to thank Patrick Hudson, my second supervisor, who has on several occasions managed to raise my self-confidence from “too low” to “too high” within one day. He is to be credited for the emphasis on scalability and wiring economy in this study. The continuing collaboration in publishing with both may one day make me master the art of writing. Special thanks go to Bob Herschberg for many valuable suggestions and comments.

I had the great fortune to share a room and work with Hans Henseler. I thank him for his many suggestions on programming issues and fruitful discussions on neural networks. Jacques Lenting’s help on mathematical issues is greatly appreciated. I am grateful to Bruno Olshausen and Charlie Anderson for stimulating discussions on the biology of vision and the computational strategies underlying visual attention. I would like to express my thanks to Patrick Schoo, Johan Thole, and Maarten van der Meulen for managing the computers and their prompt help in case of hard- or software emergencies. I also thank Harm Bakker whose expertise on LateX and UNIX helped me out on many occasions. I gladly acknowledge the assistance furnished by IBM Nederland under whose aegis the current work was performed as part of the Joint Study Agreement DAEDALOS under contract number #289651.

Finally, I thank Jacqueline, my wife, for providing professional insights into the neuropsychology of vision and her careful proofreading of many drafts of the thesis.

While grateful for the many forms of assistance I have received, I insist that I remain responsible for any errors.

Eric Postma

Maastricht, July, 1994

# Chapter 1

## Introduction

### 1.1 Covert Attention

Vision is an active process. Observers sample the visual environment dynamically and this activity facilitates interpretation. The way in which this apparently occurs involves *overt attention*, i.e., eye movements. As illustrated by Yarbus's (1967) studies, the oculomotor system scans visual scenes by a series of eye fixations. However, a less apparent selective process also exists; it is not overtly visible and therefore known as *covert attention* (Posner, 1980). This process is believed to operate in close concert with the overt-attention system of eye movements by selecting future targets for foveation, i.e., eye fixation (Mackeben and Nakayama, 1993). A widely used metaphor for the covert-attention process is that of a *searchlight* (Crick, 1984). Covertly attending to part of an image can, according to this metaphor, be likened to a searchlight illuminating that part. The searchlight's movements may proceed independently of the eye movements, as can be easily verified by inspection of Figure 1.1 (after Anstis, 1974). Fixating the eyes on the centre of the figure effectively prevents eye movements. Nevertheless, by modulating the attentional searchlight, individual characters can be selected at will, rendering them perfectly perceivable (see Figure 1.2).

We believe that the covert-attention process plays an important role in any effective visual system. Therefore, we aim at reaching a formal specification of the process so that it can be applied in parallel-distributed vision machines.

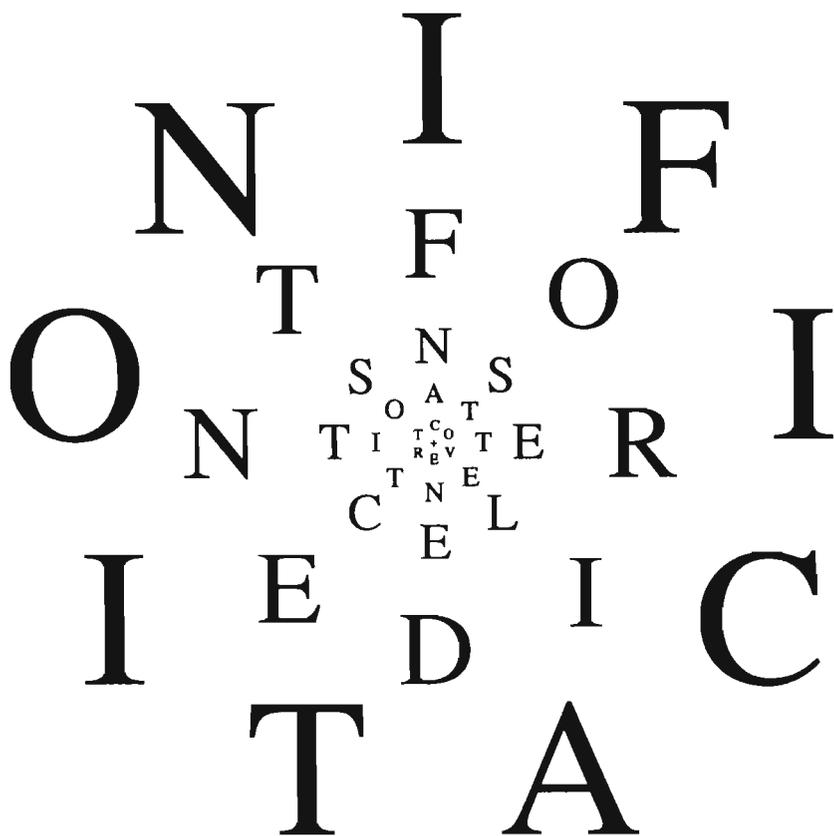


Figure 1.1: *Fixating the eyes on the centre of the figure still permits the attentional selection of any of the characters. The increasing size of the characters compensates for the loss in retinal resolution power with increasing eccentricity (after Anstis (1974)).*

Incorporating an attentional mechanism in vision machines leads to a radically different approach to image analysis. Images are actively explored instead of being passively sampled to gather information. Ballard (1991) argued for the

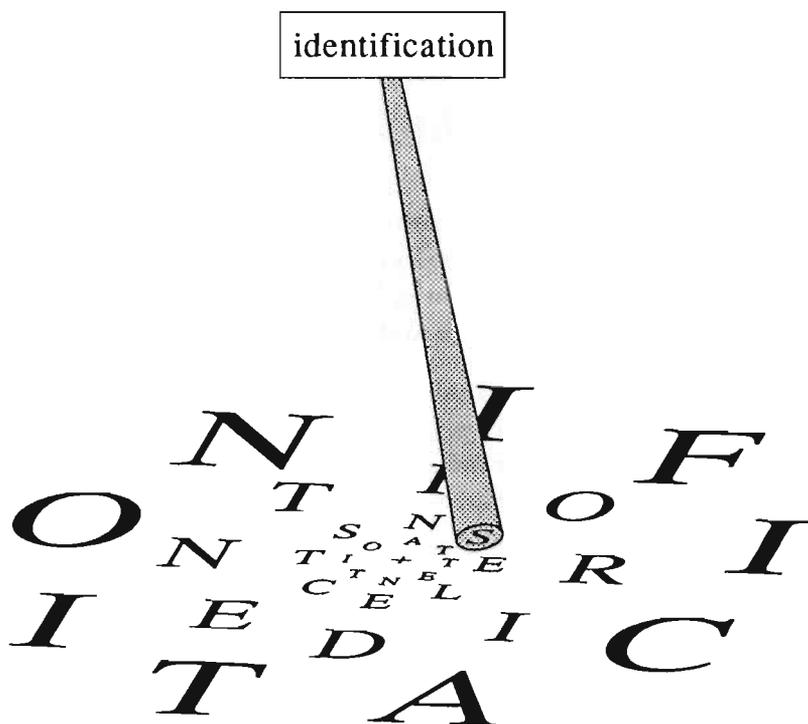


Figure 1.2: *Schematic of how an attentional searchlight might select objects for identification.*

use of an overt attentional mechanism in computer vision: a servo-driven camera that can be directed to parts of the visual scene. The use of a flexible sensor represents a first step towards biologically more realistic vision machines. The incorporation of a process analogous to covert attention represents a natural extension of this approach.

The problem statement of the present study reads: is it possible to construct a working model of covert attention as described by psychology which satisfies the constraints as set by biology?

In what follows, the multidisciplinary approach pursued is discussed. Then,

the detailed objectives of the study and the assumptions underlying the model are outlined. Finally, an overview of the thesis is provided.

## 1.2 Multidisciplinary Approach

Our approach combines ideas from four different disciplines: biology (neurophysiology), psychology (psychophysics), physics (statistical mechanics), and computer science (parallel computation). The following considerations have led to the choice of these disciplines and their merging in one approach.

(i) Biological findings on the visual system may reveal important principles underlying the attentional process. At the single-neuron level these findings provide hints to the microscopic interactions that are responsible for the emergence of attentional selection. Since our understanding of brain processes and their relation to the brain's structure is, to say the least, incomplete, care has to be taken in relating available data to functional principles. Therefore we preferably focus on relatively well-established biological characteristics.

(ii) Psychological findings represent the complement of biological findings in the sense that they reveal the functional characteristics of attention at the macroscopic (behavioural) level. Linking microscopic properties to these macroscopic properties enables the formulation of a model of attention as an *emergent* phenomenon (see Chapters 4 and 5).

(iii) The physics of large systems provides an appropriate formal tool to formalise the emergence of attention. Statistical mechanics deals with the relation between the microscopic dynamics and the macroscopic emergent behaviour of physical models (e.g., Chandler, 1987). The microscopic dynamics of such models are probably much simpler than those of biological systems. This need not require slavish modelling of the biology, since an important result from statistical mechanics is that macroscopic behaviour is in some well-described cases independent of the details of microscopic dynamics and interactions. This *universality* concept (e.g., Yeomans, 1992, see also Chapter 4) suggests that, even with relatively simple models, a realistic analogue of biological processing may be obtained.

(iv) Taking the brain's structure as a source of inspiration for our investigations implies that computational problems need to be considered as well. The

brain has to deal with the problems that confront computer scientists when they attempt to build parallel distributed computers (Heemskerk *et al.*, 1994). Apparently, the brain has found an efficient way to allocate its resources and to reduce its communication overhead (Nelson and Bower, 1990). The problem of building parallel computers capable of dealing with realistic problem sizes (i.e., problems of a scale comparable to those solved by the brain) leads to an appreciation of the *scaling problem* (Minsky and Papert, 1988) and its solution in the brain.

Minsky and Papert (1988), reviewing the reception of the first edition of their book *Perceptrons* pointed out that their “pessimistic evaluation of the perceptron” was the assertion that

...although certain problems can easily be solved by perceptrons on small scales, the computational costs become prohibitive when the problem is scaled up (p. 254).

In short, the size of a neural network *can* become a serious problem which is all too often overlooked. They went on, nevertheless, to say that:

Experiments with toy-scale problems have proved as fruitful in artificial intelligence as in other areas of science and engineering. Many techniques and principles that ultimately found real applications were discovered and honed in microworlds small enough to comprehend yet rich enough to challenge our thinking. But not every phenomenon encountered in dealing with small models can be usefully scaled up. Looking at the relative thickness of the legs of an ant and an elephant reminds us that physical structures do not always scale linearly: an ant magnified a thousand times would collapse under its own weight. Much of the theory of computational complexity is concerned with questions of scale (pp. 261-262).

Ideas from biology, psychology, and physics selected with the notion of scaling in mind ultimately lead to the specification of the SCAN (*Signal Channelling Attentional Network*) model of covert attention in Chapter 5.

### 1.3 Objectives and Assumptions

In this section we explicitly state the four main objectives that are pursued in this thesis. Additionally, we specify three assumptions and discuss each of them in detail. Within the context of the general problem statement above, the objectives of the current study are:

- specifying a scalable model of covert attention inspired by a biological example,
- formalising the model in terms of statistical mechanics,
- achieving translation invariance, and
- evaluating the model from a biological and psychological perspective.

We have based our approach on a set of assumptions about the nature and function of the covert-attention process. Three assumptions indicate the lines along which we approached our objectives:

- Covert attention is a parallel distributed process.
- Covert attention emerges from microscopic competitive dynamics.
- Covert attention enables invariant object perception.

Below, each of these assumptions is discussed in its appropriate context.

*Covert attention is a parallel distributed process.* The visual system is an intricate parallel distributed machine. Sequential behaviour, such as the selection of characters by the attentional searchlight (cf. Figure 1.1), has been interpreted as representing a “serial mode” of the underlying parallel machinery (e.g., Hurlbert and Poggio, 1985; Treisman, 1989). This interpretation leads to a “dual-mode” view of the visual system. Focused attention is associated with a serial mode of processing, whereas absence of focusing is associated with a parallel mode of processing. Treisman (1989), for instance, distinguishes between a *pre-attentive* (parallel) and an *attentive* (serial) mode. Several authors have argued against this view (e.g. Zohary and Hochstein, 1989;

Shulman, 1990; Van der Heijden, 1992), and indeed, the distinction between pre-attentive and attentive processing may not be as clear as suggested. Viewing both modes as the extremes of a continuum may be more appropriate, as is corroborated by a number of studies (e.g., Bundesen and Pedersen, 1983; Duncan and Humphries, 1989). The observed serial behaviour reflects the limitations of the underlying parallel process. Given the inherent parallelism of the brain (DeYoe and Van Essen, 1988), we adopt the more parsimonious assumption that attention is a parallel distributed process. This assumption has an important implication for the long-standing debate on the locus of attentional selection. According to the *early-selection* theory (Broadbent, 1958; Treisman, 1989) attentional selection precedes the object identification stage. In contrast, the *late-selection* theory (Deutsch and Deutsch, 1963) asserts that object identification precedes attentional selection. Although the debate has generated a large body of experimental results, no convincing evidence favouring either theory has been obtained (Allport, 1989). Our assumption that attention is distributed strongly suggests that there is no fixed locus of attention. Depending on the task the visual system is confronted with, the locus of selection may be *early* or *late*. Following our assumption one may argue for a "flexible" locus of attention (Yantis and Johnston, 1990).

*Covert attention emerges from microscopic competitive dynamics.* From the initial processing in the retina, up to the highest levels of the visual system, competitive interactions among (groups of) neurons prevail. The sharpening of contours due to competitive interaction among neurons at the peripheral stages of the visual system (see, e.g., Ratliff, 1965) may be taken to constitute a rudimentary form of attentional selection. Similarly, the decision that (part of) incoming visual data corresponds to an object representation is probably based on competitive interactions at the higher levels of the visual system and may also be considered to represent a form of attention (Grossberg, 1982). It seems therefore quite natural to assume that competition at the level of single (or small groups of) neurons represents the substrate of covert attention (cf. Walley and Weiden, 1973; Desimone, 1992). The competitive interactions among neurons on a microscopic scale give rise to the emergence of selective behaviour on the macroscopic scale of large groups of neurons.

*Covert attention enables invariant object perception.* Invariant object perception represents one of the most difficult computational tasks for artificial-vision systems. The major problem is how to extract or compute object representations that are invariant to changes in position (i.e., translation invariance).

One approach to this problem is based on the decomposition of the visual image into local features (Fukushima, 1980; Von der Malsburg, 1988; Mozer, 1991). Decomposition into features allows for their integration over large parts of the image so that positional dependencies are removed. The next step constitutes the binding of features belonging to a single object. The problem of how to do this is known as the *binding problem* (Barlow, 1981; Hinton, 1981; Sejnowski, 1986).

An alternative approach is based on the idea that covert attention is responsible for invariant perception. According to this idea, the attentional searchlight selects parts of the visual scene and remaps their contents into a pattern-recognition stage without affecting the spatial ordering of the selected pattern (Anderson and Van Essen, 1987; Baron, 1987; Nakayama, 1990). Because the spatial order is preserved from the onset, no binding problem arises. The assumption is made, in this study, that the existence of an attentional mechanism allows translation-invariant object representation.

## 1.4 Thesis Overview

The thesis is organized as follows. Chapter 2 contains a review of the biology and psychology of the visual system and, in particular, of covert attention. Five key characteristics are extracted and form the foundation of the SCAN model. Chapter 3 studies the neural lattice, an elementary selection network based on three of the five characteristics. Chapter 4 introduces an extension of the neural lattice, the gating lattice, forming the building-block for the model of attentional selection. Combination of the two remaining key characteristics with the gating lattice leads to the detailed specification of the SCAN model in Chapter 5. Chapter 6 evaluates the biological and psychological significance of the model and Chapter 7 concludes upon the model, its performance, and the general approach.

## Chapter 2

# Covert Attention in the Brain

In this chapter we review biological and psychological knowledge about the process of covert attention. In recent years knowledge about the visual system and its processes has expanded enormously (Stryker, 1990), providing a rich source of inspiration for building models of vision. Our review culminates in the formulation of five characteristics that define the foundations of our model of covert attention.

### 2.1 Levels of Organization

Any biological system, in particular the brain, has many levels of organization, each of which may be of importance for the functioning of the whole. For our purposes four levels of organization are enough (see, e.g., Shepherd and Koch, 1990):

1. *neurons,*
2. *neural circuits,*
3. *brain areas,* and
4. *behavioural systems.*

The biological aspects of covert attention mainly involve the first three levels, whereas the psychology of attention is primarily concerned with the level of

behavioural systems. In the following, the four levels are examined more closely.

## 2.2 Neural Processing

The neuron level is the lowest level of organization that we consider (although there are lower levels, see, e.g., Shepherd, 1990). The neuron, whose functioning is sketched here crudely, is the building block of the brain (Barlow, 1972). It acts as an input-output device that takes a number of signals as its inputs, compares it to an internal threshold, and generates spikes (or action potentials) as output whenever the threshold is exceeded. Figure 2.1 depicts the main components of the neuron. The sampling of input signals occurs at *dendrites*, the thresholding operation takes place at the *cell body*, and the spikes are generated at the initial part of the *axon* (i.e., the axon hillock). A spike, when generated, travels over the axon towards other neurons. Although the generation of a single spike is an all-or-none phenomenon, the rate at which spikes follow each other is related to the strength of the input signals (Adrian, 1946). The spikes generated in one neuron affect the generation of spikes in other neurons through junctions called *synapses*. It is at the synapse where the axon of a transmitting neuron makes contact with a receiving neuron. Generally this contact is at the dendrites of the receiving neuron, although exceptions exist (e.g., Crick and Anasuma, 1986). Incoming spikes effect the release of a chemical called a *neurotransmitter* at the *pre-synaptic* membrane of the synapse. The neurotransmitter is picked up at the *post-synaptic* membrane where it leads to an enhanced post-synaptic potential on the dendrite. Post-synaptic potentials are integrated at the cell body and, again, a spike is generated if the integrated potentials exceed some threshold.

### 2.2.1 Noisy Processing

The processes giving rise to an action potential are stochastic (Katz, 1966; Zucker, 1989). The amount of neurotransmitter released on arrival of a spike at the pre-synaptic membrane is not constant. It may vary from moment to moment. The release of transmitter substance is known to occur in quanta, i.e., fixed-sized portions of neurotransmitter. The quanta are released probabilistically according to a Poisson process with a rate that increases transiently

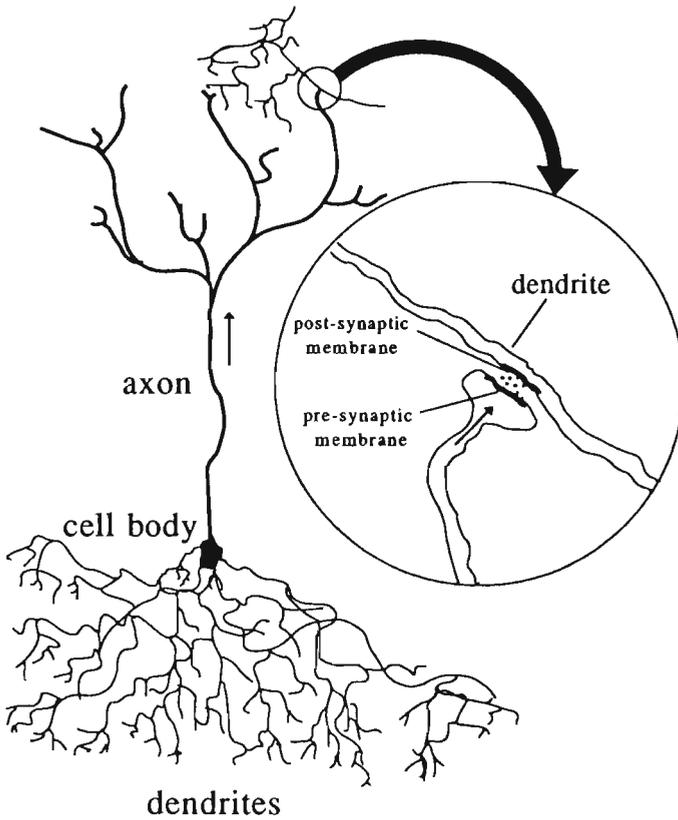


Figure 2.1: *Global outline of the neuron (after Katz (1966)).*

upon the arrival of a spike (Katz, 1966; Kuffler, Nicholls, and Martin, 1984). Consequently, neurotransmitter may be released when no spike arrives at the synapse while there may be no release when a spike does arrive. The building block of the brain appears therefore to be a computing element subject to a greater variance of transfer function than the building block of the digital computer.

### 2.2.2 Economical Wiring

Distant areas in the brain and body communicate through axons with lengths possibly exceeding a meter (Stevens, 1979). The majority of neurons in the brain, however, communicate through short-range axons with neurons in their close vicinity. There are two main advantages to short-range connectivity. The first is that it allows for *fast* communication. The longer the axon, the longer the delays. (It should be remarked, though, that long axons may be equipped with a myelin sheath that allows for a significant speed-up of spike propagation.) Spatial proximity allows thus for an efficient exchange of signals. The second advantage is that short-range connectivity is *cheap* in terms of spatial requirements (Jacobs and Jordan, 1992). As any printed-circuit designer knows, the number of communication lines that extend over long distances of the printed-circuit board should be kept to a minimum so as to exploit the limited (two-dimensional) space maximally. The same is true for wiring in the brain. The limited (three-dimensional) space available requires a parsimonious application of long-range connections, suggesting that, wherever possible, short-range connections should be preferred. This applies also to the level of neural circuits, the second level of organization in the brain. Since large communication requirements are associated with neural circuits, i.e., with small groups of neurons subserving the same function, spatial proximity in the brain often implies functional proximity (Nelson and Bower, 1990; Murre, 1992; Zeki, 1993). As demonstrated by the simulation studies of Durbin and Mitchison (1990), the structure of the peripheral part of the visual system can be understood as having evolved with minimizing wiring length as one of its limiting constraints.

### 2.2.3 Many Processing Elements

The extent of the communication problem discussed in the previous section becomes manifest when the number of neurons in the brain is taken into account. According to conservative estimates, there are about  $10^{11}$  neurons in the brain (Stevens, 1979). In contrast to the parsimony in wiring, there do not seem to be severe limitations on the number of neurons, i.e., processing elements. This overabundance allows for a profligate assignment of neurons to tasks in the brain. Viewed in combination with the stochastic nature of single-neuron processing, this leads to the notion of *distributed* or *population coding*

(Sparks, Lee, and Rohrer, 1990; Georgopoulos, Taira, and Lukashin, 1993). Since the output of single neurons is not a reliable indicator of the presence of an over-threshold input signal, the brain averages over a large number of them to achieve the required reliability.

## 2.3 Cortical Processing

Having reviewed some aspects of the lowest two levels of organization, this section focuses on the level of brain areas in the visual system. Recent developments in the neurosciences provide a wealth of data on structural and functional characteristics of the visual system. Felleman and Van Essen (1991) made an inventory of visual areas and their interconnections. They found over 30 cortical areas to be primarily visual in function. Felleman and Van Essen analyzed the approximately 300 connections among these areas and concluded that the visual system is hierarchically organized. An objective analysis of the connections corroborates their conclusion (Young, 1992). Figure 2.2 shows a diagram based on Felleman and Van Essen (1991).

The review of the biology of the visual system follows the structure of Figure 2.2. We start by discussing the sampling of visual signals by the retina (bottom of the figure). Subsequently we follow the course of visual inputs by discussing the structure and function of the striate cortex (primary visual cortex or V1, cf. Figure 2.2). Then we proceed by describing the gross division of labour in the visual system (roughly the left and right part of the figure). Finally we arrive at the level where object recognition is assumed to take place (upper part of the figure) and discuss the possible neural correlates of covert attention.

### 2.3.1 The Retina

Incoming visual information is picked up by the eye and focused onto the retina - a sheet of light-sensitive receptors at the back of the eye. Two main types of receptors exist in the retina, *rods* and *cones*. Cones are almost exclusively present in a central region of the retina called the *fovea*. Cones are colour-sensitive and their function is to enable high-resolution inspection of (parts of) visual objects. Rods are achromatic (not colour-sensitive) and are mainly

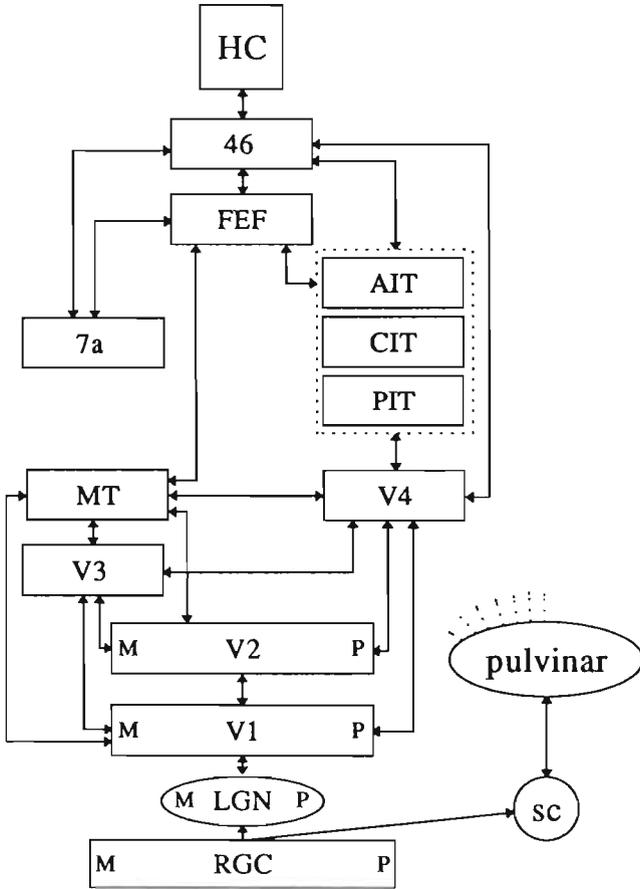


Figure 2.2: *Hierarchical structure of the visual system (after Felleman and Van Essen, 1991).*

found in the peripheral (non-foveal) region of the retina. There are about 6 million cones and 100 million rods in a human eye (e.g., Dowling, 1987). Through several intermediate processing stages, visual information picked up by the rods and cones converges on about 1.5 million "output cells": the Retinal Ganglion Cells (RGCs; see Figure 2.2). An RGC generates an output

signal whenever a stimulus is presented within its receptive field. There is a large variety in the receptive-field sizes of RGCs, resulting from the variety in receptor-RGC connectivity.

Koenderink and Van Doorn (1978) formulated a model for the spatial sampling characteristics of human vision. Their *stack model*, shown in Figure 2.3, helps us to appreciate the significance of retinal structure in spatial vision. The stack model is based on neuroanatomical, neurophysiological and psychophysical data. In the model, visual space is sampled by a number of sheets containing a fixed number of “sampling units” (i.e., RGCs). The largest sheet covers the entire visual field with coarse sampling units (i.e., the bottom layer of the stack). The smallest sheet covers the fovea (i.e., the top layer of the stack) with fine sampling units. For expositional purpose, Figure 2.3 shows an exploded view of the stack model. The individual layers should, however, be interpreted to form a single layer. Consequently, any point in the visual space is generally sampled by more than one receptor unit. A realistic stack model, in agreement with RGC counts, consists of about 40 layers each containing approximately 40,000 units.

The functional significance of the stack model becomes apparent by considering how a movement of the eye or of the observer affects the sampling pattern. Given such a movement the sampled pattern is shifted within a layer whereby its internal structure remains unaffected. Therefore, the activity pattern within a single layer provides a basis for coping with the problem of *translation invariance*. When approaching an object, the sampled pattern remains intact but is shifted by successive layers upwards in the stack (i.e., from crude sampling at the bottom to finer sampling at the top). Therefore, the activity pattern over the layers is suitable for dealing with the problem of *scale invariance* (cf. Koenderink and Van Doorn, 1978).

### 2.3.2 The Striate Cortex

Visual signals coming from the retinal network are relayed to a cortical area called striate cortex. This region (also known as V1) comprises about 20% of visual cortex (Felleman and Van Essen, 1991). The signals collected by the RGCs pass through a subcortical structure called the Lateral Geniculate Nucleus (LGN, see Figure 2.2) and map onto the striate cortex. Each of the approximately 1.5 million inputs to V1 connects to approximately 100 and

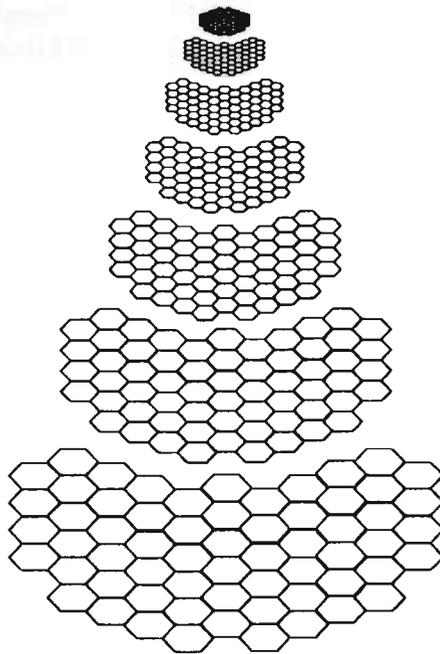


Figure 2.3: *Koenderink and Van Doorn's (1978) stack model.*

1000 V1 cells (e.g., Wilson *et al.*, 1990). The spatial structure of the retinal image is roughly preserved in the *retinotopic* map on the cortical surface. The map is unbalanced in that a disproportionately large area is devoted to the central (foveal) part of the retina. Considering the structure of the stack model such an imbalance makes sense. In the central region, the stacked layers contain more sampling units than in the periphery. To cope with the increased sampling density in the central region, the cortical area representing the central visual field is enlarged. In addition, the receptive-field size of cells in the striate cortex gradually increases with eccentricity (Hubel and Wiesel, 1974) as would be expected on the basis of the stack model.

### 2.3.3 Concurrent Processing

Throughout the assumed visual hierarchy, there is a large degree of concurrency on all levels. At the level of the striate cortex, visual signals are analyzed in parallel (*spatial parallelism*, Ullman, 1986). At the cortical-systems level, two main concurrent processing pathways can be discerned: the temporal (ventral) pathway and the parietal (dorsal) pathway (see, e.g., Desimone and Ungerleider, 1989).

The two pathways originate from relatively well separated classes of cells in the retina, striate cortex and further stages of the visual system: the parvocellular cells and the magnocellular cells (indicated by the P and M symbols at the bottom of Figure 2.2). We will discuss each of these main pathways in turn. The distinction between temporal and parietal pathways is used here rather than parvo- and magnocellular systems, because it is now clear that the separation of the parvo- and magnocellular systems is less clear than had previously been thought (Ferrera, Nealey, and Maunsell, 1992; Merigan and Maunsell, 1993).

### 2.3.4 The Temporal Pathway

Referring to Figure 2.2, the temporal pathway includes the visual areas V1, V2, V4, and the posterior inferotemporal (PIT), central inferotemporal (CIT), and anterior inferotemporal (AIT) areas, i.e., mainly the right part of the scheme. As evident from the discussion of area V1, a retinotopic representations is available at an initial stage. Moving up the temporal pathway, receptive field sizes increase gradually from V1 to V2 (approximately 1 to 2 times as large) and to V4 (2-8 times as large). The largest increase is associated with the transition from V4 to IT where receptive fields are more than 50 times as large as those in V1 (Desimone, Moran, and Spitzer, 1989). Accordingly, retinotopy decreases gradually and is totally absent at the level of the inferotemporal (IT) areas. This suggests a gradual loss of spatial dependence as one moves up the temporal pathway.

Evidently, the temporal pathway has a hierarchical organization (Van Essen and Maunsell, 1983) and has been argued to subservise the task of invariant object recognition (Gross and Mishkin, 1977; Kosslyn *et al.*, 1990; Plaut and Farah, 1991). Single cell recordings in IT and adjacent areas reveal response

characteristics of high specificity, such as faces (Perrett, Mistlin, and Chitty, 1987), hands (Gross, Rocha-Miranda, and Bender, 1972), and other complex objects (Tanaka *et al.*, 1991; Fujita, *et al.*, 1992, Tanaka, 1993). Patients suffering from damage to areas of the temporal pathway reveal impairments of object identification (*visual agnosia*, e.g., Brown, 1989).

Our interpretation of the function of the temporal pathway is one in which object-centred representations are formed and used as a prototype to be stored and retrieved from semantic structures. This interpretation is consistent with the known connections between the IT areas and the subcortical memory structures involved in the storage and retrieval of information (see, e.g., Mishkin and Appenzeller, 1987; Squire and Zola-Morgan, 1991; Li, Miller, Desimone, 1993; Miyashita, 1993).

### 2.3.5 The Parietal Pathway

The parietal pathway includes the areas V1, V2, and the posterior parietal (PP) areas V3, MT, and 7a (cf. Figure 2.2). Single-cell recordings in the PP areas reveal a range of characteristics that are related to the processing of spatial aspects of the visual input (Goldberg and Colby, 1989). Detailed neurophysiological studies of the PP area 7a suggest its involvement in the spatial transformation from gaze-centred to head-centred coordinates by combining retinal and eye-position information (Andersen, Essick, and Siegel, 1985; Andersen, 1987; Zipser and Andersen, 1988). Neuropsychological findings indicate the PP areas to be involved in the computation of a viewer-centred frame of reference, and an environment-centred frame of reference (Farah *et al.*, 1990).

Patients suffering from damage to their PP regions often show disturbances related to the analysis of spatial relations (e.g., Brown, 1989). A particularly illustrative disturbance is the (spatial) hemineglect. Vallar (1993) defined hemineglect as:

...a behavioural disorder whereby a patient fails to explore the half-space contralateral to the cerebral lesion and this deficit cannot be attributed to the impairment of elementary sensori-motor processes (p. 27).

Hemineglect patients seem to lack a spatial representation of one half of the visual world. Careful neuropsychological investigations have revealed, however, that hemineglect is an attentional impairment involving a disturbed distribution of attention in either a viewer-centred or an environment-centred reference frame (Farah *et al.*, 1990), but hemineglect has also been observed in an object-centred reference frame (Caramazza and Hillis, 1990; Driver and Halligan, 1991; Hillis and Caramazza, 1991) and, presumably may be found in combinations.

Our interpretation of the function of the parietal pathway is that it is responsible for the computation of spatial aspects of the visual environment.

### 2.3.6 Spatial Variance versus Spatial Invariance

The above discussion indicates that the two processing pathways pursue distinct computational goals. In what follows we focus on the spatial aspects of visual processing, neglecting for the moment other important parameters (e.g., colour, depth, and motion). Whereas the temporal pathway “filters out” information regarding the spatial relations among objects, the parietal pathway seems to exploit it. The *spatial invariance* of the object representation computed in the temporal pathway complements the *spatial variance* of visual information in the parietal pathway. If this interpretation is correct, the parietal and temporal pathways have to exchange information continuously in order to meet their computational objectives. Consider, for example, a cat that has spotted a mouse in his visual field. According to the functional division of labour, its temporal pathway computes a stable representation of the mouse leading to its identification. The position of the mouse will be registered by the parietal pathway leading to its localization. Each system in *isolation* yields an incomplete percept or even no percept at all. Identification without localization leads to perception of the object without knowledge of its position. (Identification might even fail due to the lack of information concerning the spatial relations of the object itself.) On the other hand, localization without identification yields perception of an unknown object at a known spatial position. (Analogously, localization might fail when it depends on the identity of the object.)

Communication between the temporal and parietal pathways is possible through direct connections, e.g., the reciprocal connections between V4 and

MT (Figure 2.2). Additionally, areas receiving input from both pathways, such as area 46 or higher centres (HC in Figure 2.2), may be used for exchange of information (cf. Young, 1992). Finally, given the reciprocity of almost all cortical connections, information from both pathways may be integrated at V1 and V2.

Obviously, several objects can be present in the visual field at several spatial scales. Therefore, a selection has to be made with respect to which object and what object features will have a reference frame attached. As will become clear, this task can be performed by the covert-attention process.

## 2.4 Neural Mechanisms of Attention

One of the most important processes in natural vision is *attention*. As discussed in Chapter 1, two forms of attention exist (Posner, 1980; Posner and Presti, 1987): overt attention, i.e., selection through eye movements (Yarbus, 1967; Wurtz, Goldberg, and Robinson, 1982), and covert attention, i.e., selection independent of eye movements. The present discussion is restricted to covert attention and its neural substrate.

### 2.4.1 Covert Attention in the Parietal Pathway

There is extensive neurophysiological evidence indicating a role of the PP areas of the parietal pathway in mediating spatial (covert) attention, i.e., the attentional selection of a spatially restricted part of the visual input (e.g., Wise and Desimone, 1988; Goldberg and Colby, 1989). The idea is in accordance with the functional role of the PP areas for computing spatial aspects of visual input. Parietal cells show an enhanced response when a stimulus in their receptive field is covertly attended to (see Heilman, Watson, and Goldberg (1987), for an overview). It is not clear whether spatial attention is being controlled by PP areas or results from modulatory signals from other areas (e.g., the pulvinar, superior colliculus (SC; see Figure 2.2) and other thalamic structures, see, e.g., Crick (1984) and LaBerge (1990)). Nevertheless, the fact that spatial attentional modulation is present in the areas subserving spatial analysis of patterns provides neurophysiological support for a spatial attention process.

As discussed in Section 2.3.5, the hemineglect caused by unilateral parietal damage is an attentional disturbance. Hemineglect patients are not capable of allocating attention to one half of internal representation of the visual world. Because internal representations in the parietal pathway are spatial representations, the attentional disturbance is in most cases spatial too.

### 2.4.2 Covert Attention in the Temporal Pathway

Although the posterior parietal cortex has often been suggested as the prime mediator of covert attention, the significant contribution of the inferotemporal areas to attentional selection is now generally appreciated. In the words of Goodale and Milner (1992):

...the focus of lesions causing the human attentional disorder of 'unilateral neglect' is parietotemporal [...], as is the focus for object constancy impairments. We conclude that spatial attention is physiologically non-unitary, and may be as much associated with the ventral system as with the dorsal (p. 24).

Desimone, Moran and Spitzer (1989) found that covert attention modulates the receptive-field structure of single cells in the temporal pathway. Without attentional modulation, the receptive field of a particular V4 cell covered a region S. Training a monkey to attend covertly to a red stimulus A positioned within S yielded strong responses of a cell whenever A was presented. When subsequently, the monkey was trained to attend a green stimulus B located at another position in S, presentation of A resulted in only a small response. Desimone *et al.*, (1989) interpreted these findings as follows:

The neuron responded as if the receptive field had contracted around the attended stimulus, so that the influence of stimuli at other locations in the field was reduced or eliminated (p. 173).

The attentional effect was only found to be present within V4 cells whose receptive fields covered both stimuli. Cells with only one of the stimuli in their receptive field showed no attentional modulation. Cells in IT, however, did show attentional modulation over a large part of the visual field.

Given our functional interpretation of the temporal pathway, it is to be expected that attentional effects in this pathway are primarily related to object-features or objects instead of locations. A recent study performed by Chelazzi *et al.*, (1993) suggests that this may indeed be the case. They presented monkeys with an object image (the cue) followed by a blank screen, after which a multi-object image was presented that might contain the cued object (the target). The monkeys were trained to make an eye movement to the target (if present) immediately after the onset of the multi-object image. Chelazzi *et al.*, recorded IT cells that responded to the cued object and found that if the cue matched one of the objects in a multiple-object image, these cells responded when the object was cued beforehand but not when it was not. Apparently, these IT cells are involved in the selection of objects.

### 2.4.3 Featural and Spatial Covert Attention

The presence of neural correlates of covert attention in both the parietal and temporal pathways indicates that attention is not a property of a single brain area. In fact, it is known that covert attention has also neural correlates in the superior colliculus and, in particular, the pulvinar that connects to almost all visual areas (represented by the dashed lines in Figure 2.2; Desimone, Wessinger, Thomas, and Schneider, 1990). Attention seems therefore to be a parallel distributed phenomenon emerging from an intricate attentional network (Corbetta *et al.*, 1990; Mesulam, 1990; Posner and Petersen, 1990; Desimone, 1992; Posner and Dehaene, 1994). Nevertheless, we limit our discussion to the temporal and parietal pathways in order to develop a first approximation to a working model of attentional processing in the brain.

Taking into account the functional specialization of the temporal and parietal pathways, the findings on attentional effects in both pathways are interpreted here as representing a combined attentional selection on the basis of object features (i.e., colour and shape) and object location (i.e., spatial coordinates). The selection on the basis of object features will be referred to as *featural (covert) attention*. In line with the alleged functional role of the temporal pathway, featural attention is thought to subservise the construction of an invariant representation along multiple parameters. Although the neural correlates of spatial attention are mainly found in the parietal pathway, they should somehow affect processing in the temporal pathway, because featural attention is often linked to a spatial position. The parietal pathway supports the

construction of a spatially invariant representation in the temporal pathway by communicating spatial information to it. Given the functional specialization of the temporal pathway, it might similarly support invariance for other stimulus parameters in the parietal pathway (e.g., object invariance).

## 2.5 The Psychology of Attention

The fourth level to be examined is the level of behavioural systems. This section discusses the functional aspects of attention as revealed by psychological investigations. A number of behavioural studies have been performed to unravel the characteristics of the covert-attention process. The results of these studies led to the formulation of the *searchlight* metaphor. According to this metaphor, covert attention selectively enhances restricted (contiguous) parts or features of visual input (Crick, 1984; Hurlbert and Poggio, 1985; Treisman, 1986; Posner and Presti, 1987). Processing of visual stimuli illuminated by the searchlight is facilitated at the cost of the processing of those lying outside the illuminated region. The searchlight metaphor emphasizes the spatial nature of attention.

### 2.5.1 The Psychology of Spatial Attention

In a standard covert-attention experiment (Posner, 1980), a subject is instructed to look at a *fixation point*. After some time, a *cue* is presented to generate a spatial expectation (e.g., an arrow). Following this, the *target* is presented. The time lapse between cue and target presentation is known as the stimulus-onset asynchrony (SOA). Usually, the subject is required to respond to the target by pressing a key. The reaction time is defined as the latency between target presentation and the subject's response.

The experiments reveal that cuing the correct location of a target results in a decreased response latency (Posner, 1980). Alternatively, cuing the incorrect location results in an increased latency. These findings can easily be interpreted in terms of the searchlight metaphor. A valid cue positions the searchlight on the expected location, yielding a facilitation in processing. Positioning the searchlight at an incorrect position hampers processing at the correct position lying outside the searchlight's beam.

Instead of cuing at the fixation point a cue can be presented at the location where the target will appear. Then it is not necessary to establish cue identity to direct the searchlight. Rather, detection (not identification) of the cue suffices to position the searchlight. This implies that the positioning information is directly available. The distinction between *controlled* and *automatic* shifts of attention is closely associated with the two forms of cuing (i.e., cuing at the fixation point or at the target location). The controlled shifts of attention are under the control of will. The effect of the arrow cue can be considered to induce a controlled shift of attention towards the expected location. Automatic attention is associated with the shift in attention not under voluntary control.

The processes of attentional selection and object identification are strongly related. The identification of objects has been argued to proceed in a single feedforward pass through the temporal hierarchy. Thorpe and Imbert (1988) report on a study in which subjects are presented with pairs of real everyday images, one immediately after the other. Subjects had to name the content of both images ("Charlie Chaplin", "Eiffel Tower", etc.). On those trials where the second image was correctly identified, the proportion correctly identified first images was 0.80 when the image duration was 100 ms. Increasing presentation time to 200 ms did not have a significant effect on this proportion. Thorpe and Imbert (1988) conclude that 100 ms presentation time is sufficient for accurate identification of natural images. They argue that given the biological facts, there is no time left for complex feedback processes.

The role of attentional selection in the feedforward identification process is nicely illustrated by a study of Bonda and Thorpe (1993). Subjects were presented with four natural images simultaneously, one on each quadrant of a rectangular region whose centre served as the fixation point. While keeping the eyes fixated on the fixation point, subjects were required to identify the image whose quadrant was cued preceding image presentation. Results showed that the proportion correctly identified images was almost the same for the four simultaneously-presented images as for a single image presented in isolation. These results suggest that the advance cuing positions the attentional searchlight on the appropriate quadrant, thereby effectively preventing interference from the images in the other quadrants. One of the major functions of attention may be to prevent interference from extraneous information (Chaudhuri, 1990; Green, 1991).

From the above results, we conclude that once the attentional searchlight is

positioned, the identification of a target object proceeds in a feedforward way without interference from other objects.

### 2.5.2 The Psychology of Featural Attention

Salient stimuli (e.g., a red tomato on a homogeneous contrastive background) invoke automatic shifts of attention. Such stimuli “attract” the attentional searchlight automatically. We may, therefore, associate automatic shifts with featural attention and controlled shifts with spatial attention. Featural attention has been studied extensively by Treisman and colleagues (Treisman, 1982; Treisman and Sato, 1990). According to the Feature Integration Theory (FIT), a visual scene can be segregated rapidly on the basis of boundaries defined by a change in a single feature. This rapid segregation is shown in Figure 2.4a. The boundary is defined by a change of a single feature, namely colour. Alternatively, as shown in Figure 2.4b, a boundary defined by a change of a feature-conjunction (colour and form) is not immediately perceived. According to FIT, spatial attention is invoked here in order to inspect each feature conjunction separately. Treisman (1982) associates the rapid segregation of Figure 2.4a with parallel processing, whereas the slow attentional segregation of Figure 2.4b is associated with serial processing.

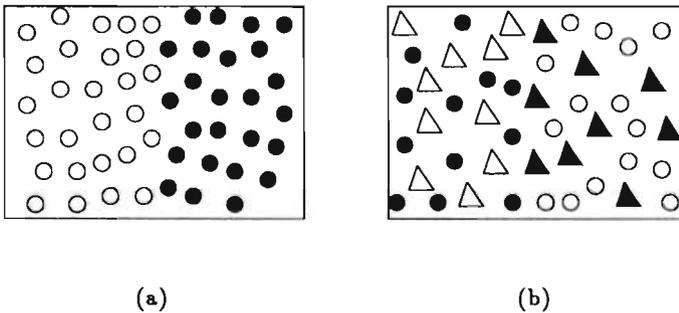


Figure 2.4: (a) Rapid segregation of the boundary defined by a single feature (colour); (b) Slow segregation of a boundary defined by a feature conjunction (colour and form).

## 2.6 Characteristics of a Vision Model

The biological and psychological aspects of covert attention discussed in the previous sections provide a source of inspiration for the specification of the covert-attention model to be formulated in Chapter 5. We list five characteristics that form the foundation of the model. Three of these characteristics pertain to the levels of neurons and neural circuits they are:

- stochasticity,
- sparse connectivity, and
- profligacy.

The remaining two characteristics

- hierarchy, and
- invariance

refer to the levels of brain areas and behavioural systems, respectively.

### 2.6.1 Stochasticity

The stochastic nature of neural processing (Katz, 1966) is adopted in our model by formalising the processing elements as probabilistic threshold neurons (e.g., Little, 1974; Amit, 1989).

### 2.6.2 Sparse Connectivity

The emphasis on sparse connectivity in the brain (Nelson and Bower, 1990) leads to the definition of a sparse connectivity pattern in our model. In particular, we define for each processing element a small set of nearest neighbours from which it receives connections.

### 2.6.3 Profligacy

The profligate availability of brain cells (Stevens, 1979) implies in our model that we place no heavy restrictions on the number of processing elements needed. Since the characteristics of stochasticity and sparse connectivity may lead to unreliable results, increasing the number of processing elements and averaging over their responses enhances reliability.

### 2.6.4 Hierarchy

Hierarchy is a characteristic that pertains to the structure of the visual system (cf. Figure 2.2) that is thought to reflect a *divide-and-conquer* strategy (Felleman and Van Essen, 1991). A complex task is best approached by decomposing it into smaller tasks, each of which is easier to solve than the original task. Repeating this task decomposition, one ultimately reaches a level at which the subtasks are within reach of the computational power of individual processing elements. The visual subsystems, specifically the object-identification pathway, follow this divide-and-conquer strategy. The decreasing magnitude of receptive-field sizes, going from the level of IT downwards to the level of V1, suggests that in this hierarchy the visual identification task is spatially decomposed. Incorporation of a spatial decomposition analogous to that in the identification pathway puts the first constraint on the overall structure of our covert-attention model.

### 2.6.5 Invariance

The functional interpretation of the temporal pathway is taken as a last characteristic. As the discussion in Section 2.3.4 made clear, the temporal pathway may be responsible for the computation of an invariant object representation. Clearly, such a representation requires selection of an object from the retinal image. Whether this selection is object-based (Duncan and Humphries, 1989, Bundesen, 1990) or location-based (Treisman, 1982; Van der Heijden, 1993) is still subject of debate. Probably, attentional selection can be based on both (Baylis and Driver, 1993). In reviewing neuropsychological evidence on this issue, Farah, Wallace, and Vecera (1993) proposed a model in which an input array exchanges information with two separate modules: an object-

representations module and, what we will call, a spatial-representations module. The presence of objects in the input array leads to the activation of their representations in the object-representation module and vice versa. In the words of Farah *et al.*, (1993):

Whereas stimulus input can activate any arbitrary configuration of locations, top-down activation from the object system will activate *regions of the array that correspond to objects* (p.136).

In our model of covert attention, top-down activation from the level of object representations is used to segregate objects from their background in the input array. The part of the input that receives the largest activation is selected and channelled towards the level of object representation. The top-down activation from the object-representations module can be easily combined with the top-down activation from the spatial-representations module by adding an analogue of the parietal pathway to our model (cf. Van der Heijden, 1992). In that way, attentional selection (i.e., the channelling of parts of the input) could be either object-based, location-based, or both. In our model, however, we have limited ourselves to constructing an analogue of attentional selection in the temporal pathway making the assumption that the necessary contribution of the parietal pathway, specification of location, can be implemented analogously. In accordance with our assumption that *covert attention enables invariant object perception*, the spatial structure of selected parts are preserved during channelling.

### 2.6.6 Attention: Integration of the Characteristics

The presumed role of the attentional searchlight in achieving translation invariance in the temporal pathway links the biology of this pathway with the psychology of covert attention. SCAN, our model of covert attention, is inspired by the part of the temporal pathway responsible for the computation of invariant object representations.

The build-up of attentional modulation with increasing levels in the hierarchically structured temporal pathway indicates that attention is *itself* a hierarchical process (cf. Koch and Ullman, 1985). The selective action of attention is based on a divide-and-conquer strategy. The complex task of selecting a

contiguous part out of the large retinal pattern can be decomposed into multiple local selection tasks each of which is concerned with a distinct subregion of the retinal pattern. Spatial decomposition leads to localized selection at the various levels. At the base level, selection extends over small distances. At higher levels, a reduced set of image parts is selected until, finally, at the top level a single part of the image, i.e., an object pattern, is selected.

SCAN responds to a conspicuous location of its input by focusing its searchlight on that location. Since our model is based on the temporal processing pathway, the dimensions along which conspicuity is defined are assumed to be under control of an object-based expectation mechanism. Consequently, when a particular object is expected, any part of the image that resembles an object is conspicuous. In SCAN, the attentional selection operates in a bottom-up fashion. Conspicuity is, however, under top-down control.

Before turning to a detailed specification of our model in Chapters 4 and 5, the next Chapter studies the significance of connection density in *neural lattices*, neural networks that are based on the first three characteristics.



## Chapter 3

# The Neural Lattice

The first step towards the formulation of a model of attention is the specification of a selective mechanism. In this chapter we discuss the *neural lattice*, a sparsely-connected neural-network realization of what may be regarded as the simplest selective mechanism. Although the selective capability of the neural lattice is too limited to be of use in our model of covert attention, its study enables a detailed specification of the effect of the connection density on network performance. Since a large connection density hampers the scalability of neural networks, we compare the performance of sparsely-connected neural lattices with the performance of one that is fully connected. The results obtained generalize to *gating lattices*, i.e., the building blocks for attentional selection in SCAN. The latter are introduced in Chapter 4.

Neural networks are based on a large number of interacting processing elements and are loosely inspired by the structure of real neural networks in the brain (see, e.g., Kohonen, 1977; Hinton and Anderson, 1981; Grossberg, 1982; Rumelhart, McClelland, and the PDP research group, 1986; Amit, 1989). The processing elements are sometimes referred to as (artificial) *neurons*, although the relative simplicity of the processing elements contrasts sharply with the complexity of their biological counterparts. As indicated in Chapter 1, our main objective is to capture a process analogous to covert attention in a scalable neural-network model without assuming too much complexity at the level of the underlying processing elements. For this reason the processing elements that we consider are reduced to simple binary threshold elements. The task of deciding whether an input signal exceeds a fixed threshold is most appro-

appropriate for such elements. Following the characteristic of profligacy we study the macroscopic properties of large numbers of processing elements. The microscopic characteristic of stochastic processing is reflected in the dynamics of the processing elements. The emphasis is, however, on the characteristic of sparse connectivity. In particular, we focus on the question: what is lost when full connectivity is abandoned and sparse connectivity is used instead?

After a discussion of the theory of threshold elements, we turn to the neural lattice, that is composed of a large number of stochastic threshold elements. We examine the emergent properties of a neural lattice by formulating it as an Ising lattice, a well-known model in statistical mechanics. Combining known properties of Ising lattices with quantitative results obtained from simulation studies we find that sparsely-connected networks are in some cases to be preferred to fully-connected networks when input signals have a Gaussian distribution and when selection should proceed independently of the network's initial state.

## 3.1 Threshold Elements

### 3.1.1 Deterministic Threshold Elements

Consider the task of determining whether an input signal  $h$  exceeds a certain threshold value  $\theta$ . In neural networks this task is readily solved by employing a binary threshold element (McCulloch and Pitts, 1943) whose state signals the presence of an over-threshold input and is represented by  $S$ . If the input signal is smaller than the threshold ( $h < \theta$ ) the threshold element is in its *OFF* state (which we denote by  $S = -1$ ), whereas when the input signal exceeds the threshold ( $h > \theta$ ) it is in its *ON* state ( $S = +1$ ):

$$S = \begin{cases} +1 & \text{if } h > \theta \\ -1 & \text{if } h < \theta . \end{cases} \quad (3.1)$$

The state associated with the case  $h = \theta$  is deliberately left undefined (see below), normally it is included in either one of the conditions. The discrete-time dynamics of the McCulloch-Pitts elements follow from this state equation. At each time instance  $t$ , the input at  $t - 1$  is compared to the threshold and

the state is set according to (3.1).

### 3.1.2 Stochastic Threshold Elements

The McCulloch-Pitts threshold element is deterministic and is therefore not in line with the characteristic of stochastic neural dynamics. As we have seen in Section 2.2.1, the firing of neurons do not follow from their inputs in a deterministic way. When the intrinsic noise of neural processing elements is taken into account, the state probability of the threshold element can be approximated within 1% by a sigmoid probabilistic function of the form (Little, 1974; Amit, 1989):

$$P(S = S') = \left[ 1 + \exp \left( -\frac{S'}{T}(h - \theta) \right) \right]^{-1}, \quad (3.2)$$

where  $P(S = S')$  denotes the probability of the threshold element assuming state  $S' \in \{-1, +1\}$ . The input  $h$  is scaled by the *intrinsic-noise* parameter  $T$  ( $T > 0$ , also known as the *temperature*). Interestingly, expression (3.2) has an analogon in the dynamics of Ising spins that is pursued in neural network models (Little and Shaw, 1975; Hopfield, 1982) and also in our model (see below). The average value of  $S$  in equilibrium,  $\langle S \rangle$ , can be computed by noting that (e.g., Hertz, Krogh, and Palmer, 1991):

$$\langle S \rangle = P(S = +1) \times (+1) + P(S = -1) \times (-1). \quad (3.3)$$

Substitution of (3.2) in (3.3) yields

$$\langle S \rangle = \tanh \left( \frac{h - \theta}{\frac{1}{2}T} \right). \quad (3.4)$$

In case  $h = \theta$  (and  $T > 0$ ), it follows that  $\langle S \rangle = 0$ .

### 3.1.3 Interacting Threshold Elements

From a biological perspective, a single element deciding whether an input exceeds a threshold is highly implausible. Instead, in biological systems such decisions are typically made by a large number of elements (McIlwain, 1976). Moreover, these elements do not operate independently, but interact to enhance the performance of the whole (e.g., Georgopoulos, Taira, and Lukashin, 1993).

Following the characteristic of profligacy we examine the effects of interacting threshold elements and define the *neural lattice*: a network of interconnected threshold elements. The neural lattice is required to respond to an over-threshold input signal by assuming a state where all its constituent elements are *ON*. Similarly, it responds to a under-threshold signal with a state in which all elements are *OFF*. As such, the global behaviour of the neural lattice may be viewed as that of a threshold element that detects an over-threshold input signal. A large number of threshold elements enables an instantaneous averaging over the entire population leading to a more reliable over-threshold detection within a fixed time-span than in case of a single stochastic element. In addition, the interactions among the elements may (as will become evident) enhance the robustness of detection. The detection performance of neural lattices with differing degrees of connectivity are evaluated on two points. First, the resistance to noise, superimposed on the input signals is taken as a measure of robustness. Second, the independency on the initial state is taken as a measure of flexibility. Although, the task faced by the neural lattice is a simple one, it enables us to obtain a detailed characterisation of the effects of the connection density on both measures.

### 3.1.4 Energy Function

Following Hopfield's (1982) approach, we characterise the neural lattice by an energy function  $E_{NL}$ . This function returns a scalar energy value  $E_{NL}$  for each lattice state  $\mathbf{S}$  (the vector of states  $S(i)$  of its elements).

We formulate the energy function for the neural lattice as

$$E_{NL}(\mathbf{S}) = - \sum_i \sum_j J(i, j) S(i) S(j) - \sum_i h(i) S(i), \quad (3.5)$$

where  $J(i, j) = J(j, i)$  is the weight of the connection between the elements  $i$  and  $j$ ,  $S(i)$  is the state of element  $i$ , and  $h(i)$  is the external input of element  $i$ . (Unless stated otherwise, we assume that  $h(i) = h$  for all  $i$ , with  $h$  a constant.)

For proper detection, the minima of  $E_{NL}$  should correspond to the network states with  $S(i) = -1$  for all  $i$  when  $h < 0$  and to states with  $S(i) = +1$  for all  $i$  when  $h > 0$ . This is the case when  $J_{ij} = J > 0$  for all distinct pairs  $i, j$  ( $J(i, i) = 0$  for all  $i$ ,  $J$  is a constant). Then the first righthand side term of (3.5) is minimal when either  $S(i) = -1$  for all  $i$  or  $S(i) = +1$  for all  $i$ . The last term of (3.5) is minimal when each element's state has the same sign as its input. In combination with the constraint expressed in the first righthand-side term, the last term minimizes the energy when  $S(i) = -1$  for all  $i$  and the majority of inputs is smaller than zero, or when  $S(i) = +1$  for all  $i$  and the majority of inputs is larger than zero.

The structure of the neural lattice follows from its energy function. The quadratic terms (i.e., the first righthand-side term) prescribe the presence of a connection between each pair  $i, j$  of threshold elements for which  $J(i, j) \neq 0$ . The linear terms (i.e., the last term of (3.5)) define the one-to-one mapping of external inputs to elements.

In terms of (3.2) the state of the  $i$ -th element in a neural lattice (where we assume  $\theta = 0$ ) is given by:

$$P(S(i) = S') = \left[ 1 + \exp \left( -\frac{2S'}{T} [h(i) + \sum_{j \neq i} J(i, j) S(j)] \right) \right]^{-1}, \quad (3.6)$$

The input to each element can be partitioned into an *external input*  $h_{ext}(i) = h(i)$  and *internal input*  $h_{int}(i) = \sum_j J(i, j) S(j)$ .

The detection of an over-threshold (i.e., positive) signal  $h$  is expressed in the average network state  $m$ :

$$m = \frac{1}{2} \left( 1 + \frac{\sum_i S(i)}{N} \right), \quad (3.7)$$

where  $N$  is the number of threshold elements in the lattice. When all threshold elements are *ON*  $m = 1.0$ , whereas when all elements are *OFF*  $m = 0.0$ .

The equilibrium properties of fully-connected neural networks can be computed directly in terms of mean-field theory (Hopfield, 1984; Hopfield and Tank, 1985; Amit, 1989; Hertz, Krogh, and Palmer, 1991; see also below) leading to an expression similar to (3.4). For this reason, most theoretical work on neural networks deals with the case  $J(i, j) \neq 0$  for all distinct pairs  $i$  and  $j$ . Our emphasis on scalability leads us to study networks which are not fully connected.

In the neural lattice, sparse connectivity does not allow for *direct* global communication among the threshold elements. As a result, the global decision on the presence of an over-threshold input signals is expected to be more sensitive to noise than it would be in the fully-connected case. To assess this sensitivity, we assume that the external inputs of the threshold elements arise from a common source that generates control signals  $h(i)$  that are normally distributed around a mean  $\langle h \rangle$  and standard deviation  $\sigma$ . We suspect that, when  $\sigma$  is small, full connectivity may not be essential and that sparse connectivity is sufficient for proper operation.

## 3.2 Sparse Connectivity Patterns

In this section we examine the effects of sparse connectivity patterns on a network's behaviour. With such connectivity patterns an exact mean-field analysis of the neural lattice becomes unfeasible (Baxter, 1982). We therefore study the performance of neural lattices by combining theoretical data and simulation studies.

### 3.2.1 Neural Lattices with Nearest-Neighbour Connections

As stated in the introduction of this chapter, our neural lattice is based on the characteristics of profligacy and noisy processing in that it consists of a (potentially) large group of stochastic elements. In line with the characteristic of sparse connectivity, we examine the average behaviour of sparsely interconnected neural lattices. In particular, we restrict our investigation to lattice structures in which each element is only connected with its nearest neighbours. Three neural lattices, differing in lattice dimensionality, are the subject of our study. The behaviour of each of the lattices is examined and compared against

the behaviour of a fourth one that is fully connected.

The *one-dimensional neural lattice* consists of a chain-like network structure (see Figure 3.1a); each element is connected with its two (left and right) neighbouring elements. When assuming a square lattice, each element of a *two-dimensional neural lattice* is connected with its four nearest neighbours (Figure 3.1b). The *three-dimensional neural lattice* is one in which, assuming a cubic lattice, elements are connected with their six nearest neighbours (Figure 3.1c). Finally, the "*N-dimensional*" neural lattice leads to a hypercube in which elements are connected with their  $N - 1$  nearest neighbours (Figure 3.1d).

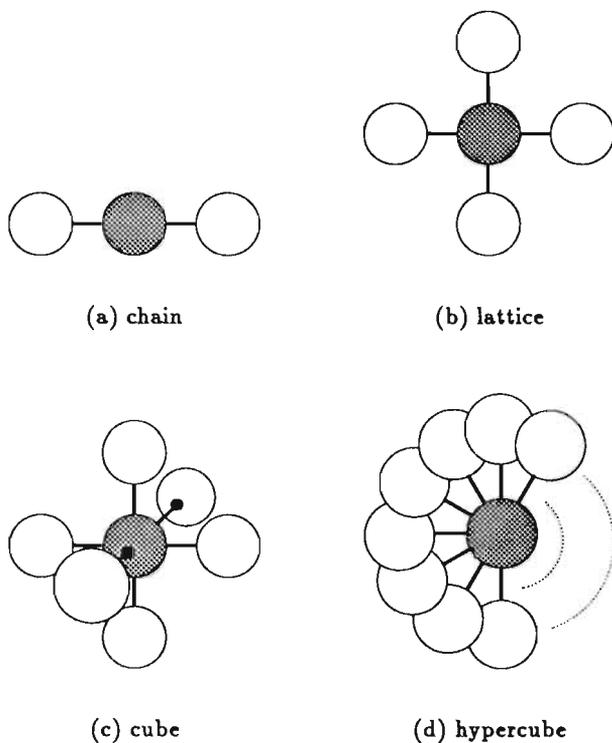


Figure 3.1: *Illustration of the connectivity patterns of the one-dimensional (a), two-dimensional (b), three-dimensional (c), and  $N$ -dimensional (d) neural lattices.*

Little (1974) and Hopfield (1982) showed that a neural network composed of threshold elements is equivalent to the Ising model of statistical mechanics (Ising, 1925). We formulate the neural lattice as a special case of the standard Hopfield neural network that is equivalent to the *Ising lattice*, i.e., a special case of the Ising model. The next section reviews the statistical mechanics of Ising lattices providing the background for the treatment of the statistical mechanics of neural lattices in the subsequent section.

### 3.3 Statistical Mechanics of Ferromagnets

The Ising lattice is a model of magnetism (Ising, 1925; Baxter, 1982). We consider the standard Ising lattice consisting of locally interconnected elements, called *spins*, and placed on an infinitely large grid. Spins can be thought of as small magnets that are in one of two orientations: *up* or *down*. Neighbouring spins are coupled by the magnetic force (“torque”) they exert onto each other. Consequently, when one spin is up its neighbour tends to align with it if the coupling is positive. An Ising lattice with all couplings positive forms a model of a ferromagnet (Baxter, 1982).

The Ising model energy function  $E_{Ising}$  is

$$E_{Ising}(\mathbf{S}) = - \sum_i \sum_j J(i, j) S(i) S(j) - \sum_i h(i) S(i), \quad (3.8)$$

where  $\mathbf{S}$  is the state vector whose coordinates are the spin states  $S(i) \in \{-1, +1\}$ ,  $J(i, j)$  ( $= J(j, i)$ ) is the strength of the coupling between spins  $i$  and  $j$ , and  $h(i)$  is the external field of spin  $i$ . In the standard ferromagnetic Ising lattice, we have, for all  $i$  and  $j$ ,  $J(i, j) = J > 0$  if  $i$  and  $j$  are nearest neighbours, and  $J(i, j) = 0$  otherwise; i.e., only nearest neighbouring spins are coupled.

For a ferromagnet, the first righthand-side term is minimized when all spins are aligned, i.e., are in the same state. As far as the energy is concerned, it does not matter whether all spins are up or all spins are down. This intrinsic symmetry is characteristic for the Ising lattice. The symmetry can be broken by the last term of (3.8). This term is minimized when all spins align with a homogeneous external field  $H$  (i.e.,  $h(i) = H$  for all  $i$ ). Consequently, a

non-zero field  $H$  breaks the symmetry by biasing the state where all spins are aligned with the external field.

The Ising energy function (3.8) is identical to the neural-lattice energy function (3.5): spins correspond to threshold elements, magnetic couplings to connection weights, and external fields to external inputs.

### 3.3.1 The Phase Transition

The energy of a system state  $\mathbf{S}$  is directly related to the probability of the system being in state  $\mathbf{S}$ . For the Ising lattice, the probability of the system being in state  $\mathbf{S}$  is given by the Boltzmann-Gibbs distribution (e.g., Plischke and Bergersen, 1989):

$$P(\mathbf{S}) = \frac{\exp(-E_{\text{Ising}}(\mathbf{S})/T)}{Z}. \quad (3.9)$$

$E_{\text{Ising}}(\mathbf{S})$  is the energy function defined in (3.8) and  $Z$  is a normalization constant called the partition function (“Zustandssumme”) that is defined as

$$Z = \sum_s \exp(-E_{\text{Ising}}(s)/T), \quad (3.10)$$

where the summation runs over all possible system states.  $T$  is the intrinsic-noise or temperature value of the system. The Boltzmann-Gibbs distribution indicates that the most probable states of the Ising lattice are states with low energy values. The intrinsic-noise value controls the randomness of the system. At very high noise ( $T \rightarrow \infty$ ) all states are equally likely and the system is completely random. At the other extreme ( $T \rightarrow 0$ ), only system states with minimal energy values are probable and the system is completely deterministic.

For neural lattices we are interested in the probability of states where the majority of threshold elements (spins) are either *ON* or *OFF* (up or down). In the context of Ising lattices, such states are called *ordered states*. These states correspond to the global minima of the energy function (when  $J > 0$ ) and their probability is scaled by  $T$ . Since  $T$  controls the randomness of the system, for some value of  $T$  the ordered state disappears and a disordered state appears.

The value of  $T$  for which this occurs is known as the critical temperature  $T_c$ . In the context of neural lattices,  $T_c$  will be referred to as the *critical intrinsic-noise value*. Physicists refer to the ordered and disordered states as *ordered* and *disordered phase*, respectively. At  $T_c$ , a *phase transition* takes place from a ordered to a disordered phase (or vice versa) (see e.g., Stanley, 1971).

The presence of order in the ferromagnetic Ising lattice can be monitored by its *order parameter*: the *magnetization* (per spin)  $m$ . In (3.7) we have defined  $m$  as an order parameter of the neural lattice. The equivalence of the neural lattice and the Ising lattice enables us to treat  $m$  analogously in both lattices. When  $m = 0.5$ , the Ising lattice is in its disordered state, whereas when  $m \neq 0.5$  it is in an ordered state.

### 3.3.2 Spontaneous and Field-Induced Magnetization

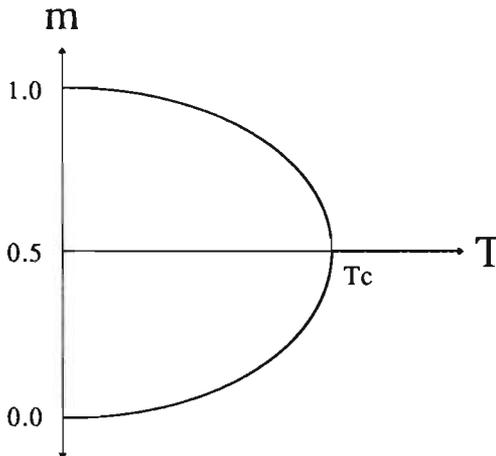


Figure 3.2: *Equilibrium magnetization of the ferromagnetic Ising lattice as a function of the intrinsic noise for  $H = 0$ .*

The macroscopic behaviour of the ferromagnetic Ising lattice is demonstrated in Figure 3.2. It shows the magnetization curve as a function of the intrinsic noise  $T$  when  $H = 0$ . Below the critical intrinsic-noise value  $T_c$ , the Ising lattice exhibits *spontaneous magnetization*. Starting from a disordered state ( $m = 0.5$ ), the Ising lattice will reach an equilibrium state  $m \neq 0.5$ . Below

$T_c$  the symmetry of the system, i.e., the balanced number of up and down spins, is broken and an asymmetric state is selected. In terms of the neural lattice, this implies that when  $T < T_c$ , and  $h(i) = H = 0$  for all  $i$ , the number of elements that is *OFF* differs from the number of elements that is *ON*. At  $T = 0$ , all elements are in the same state (either all *OFF* or all *ON*). The spontaneous emergence of order for  $T < T_c$  results from the nearest-neighbour interactions among the spins.

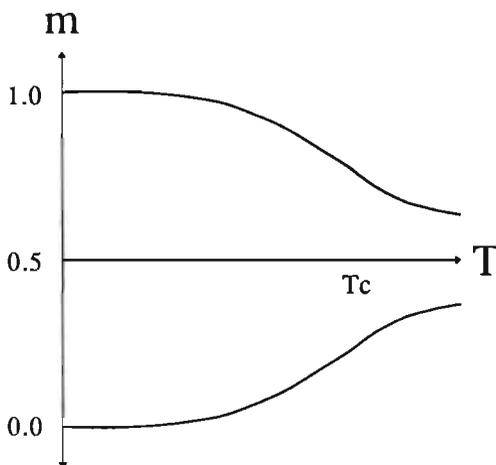


Figure 3.3: *Equilibrium magnetization of the ferromagnetic Ising lattice as a function of intrinsic noise for  $|H| > 0$  (after Yeomans (1992)).*

With non-zero  $H$ , the behaviour of the order parameter changes radically. The equilibrium magnetization curve for  $|H| > 0$  is shown in Figure 3.3. In contrast to the  $H = 0$  case, there is no discontinuity at the critical intrinsic-noise value for  $|H| > 0$ . Rather, the magnetization curve is continuous and the ordered state extends into the  $T > T_c$  region. Plotting the magnetization of an Ising lattice as a function of the external field yields a graph as shown in Figure 3.4. When  $T > T_c$ , the magnetization is a continuous sigmoidal function of the external field  $H$ . In the case  $T < T_c$ , there is a discontinuity at  $H = 0$ . The discontinuity disappears at the critical point  $T_c$ . At that point the magnetization curve has an infinite slope at  $H = 0$ .

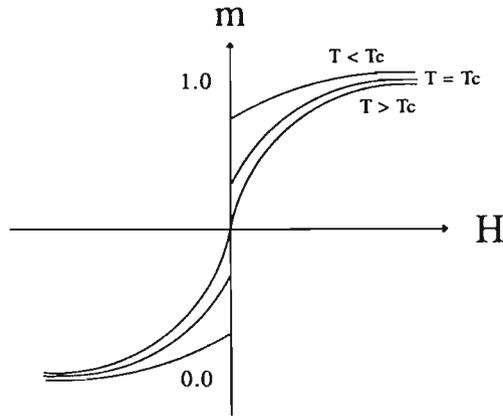


Figure 3.4: *Equilibrium magnetization of the ferromagnetic Ising lattice as a function of the external field  $H$  for various values of the intrinsic noise (after Yeomans (1992)).*

## 3.4 Statistical Mechanics of Neural Lattices

The statistical mechanics of Ising lattices can be applied directly to neural lattices. It has to be emphasized, however, that, in contrast to the physicists, we are primarily interested in the performance measures of robustness and flexibility. Moreover, whereas statistical mechanics deals with systems of infinite size and sufficiently long relaxation times, our studies focus on systems of finite size and their relaxation within a relatively small fixed time span. This leads to a change in several system characteristics as will be discussed in the next chapter.

### 3.4.1 Neural Lattices as Ising Lattices

The four neural lattices defined in Section 3.2.1 are equivalent to four distinct Ising lattices that have been well studied in statistical physics. These are: the one-dimensional, two-dimensional, three-dimensional, and mean-field Ising lattices. We rewrite the neural-lattice energy function (3.5) as

$$E_{NL}(\mathbf{S}) = -J \sum_i S(i) \sum_{j \in N(i)} S(j) - \sum_i h(i)S(i), \quad (3.11)$$

where the lattice geometry follows from the definition of the neighbourhood set  $N(i)$  and the connection weight is defined as  $J = q^{-1}$  where  $q$  is the number of nearest neighbours, i.e., the number of elements in  $N(i)$ . This lattice-dependent value of  $J$  ensures that the magnitude of the internal field is identical for all lattices. The normalization of the internal field allows a fair comparison of lattice performances.

### The One-dimensional Neural Lattice

The one-dimensional neural lattice is equivalent to the one-dimensional Ising lattice (Ising chain). The energy function of this neural lattice is obtained by defining the appropriate neighbourhood. Since neural lattices have a finite size, appropriate boundary conditions need to be defined. We assume periodic boundary conditions for all neural lattices. For the one-dimensional neural lattice the neighbourhood is defined as

$$N(i) = \{ (L + i - 1) \bmod L, (L + i + 1) \bmod L \}, \quad (3.12)$$

with  $L$  the linear dimension that, in this case, is equal to  $N$ , the number of elements in the chain. The main characteristic of the one-dimensional neural/Ising lattice is that it does not exhibit a phase transition for any  $T > 0$  (e.g., Chandler, 1987). Any (partially) ordered state is destroyed by the stochastic flipping of spins at non-zero intrinsic noise, so that eventually at equilibrium it ends up in a disordered state. The critical intrinsic-noise value of the one-dimensional neural lattice has a value of zero (i.e.,  $T_c^{1d} = 0$ ). For  $H = 0$ , an ordered state can therefore be maintained in the absence of intrinsic noise only.

The equilibrium properties of the one-dimensional Ising lattice are known exactly. Of particular interest for our purposes is the behaviour of the equilibrium magnetization as a function of an homogeneous external field ( $h(i) = H$ , for all  $i$ ). It is given by (e.g., Plischke and Bergersen, 1989):

$$m = \frac{1}{2} \left[ \frac{\sinh \frac{H}{T}}{\sqrt{\sinh^2 \frac{H}{T} + \exp\left(\frac{-4}{T}\right)}} + 1 \right]. \quad (3.13)$$

The shape of this function for  $T \in [0.1, 1.0]$  and  $H \in [0.0, +0.01]$  is illustrated in Figure 3.5. It is evident from this graph that at  $T$  close to zero, very small positive values of  $H$  manage to drive the one-dimensional Ising lattice into the ordered state with  $m \approx 1.0$ .

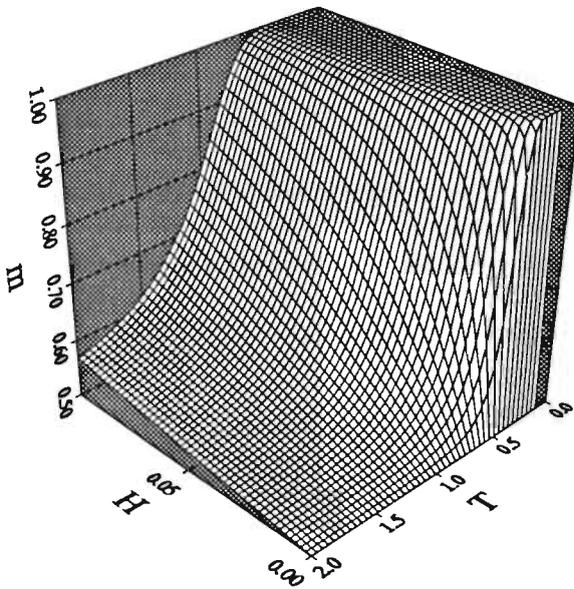


Figure 3.5: *Equilibrium magnetization of the one-dimensional Ising lattice as a function of the external field  $H$  and intrinsic noise  $T$ .*

For  $H = 0.0$  the ordered state with  $m \approx 1.0$  (or  $m \approx 0.0$ ) is stable only in the absence of noise; for non-zero values of  $H$ , the ordered states may survive an amount of intrinsic noise proportional to the magnitude of  $H$ , as can be seen from Figure 3.5. When  $H$  increases, the largest value of  $T$  for which  $m \approx 1.0$  is reached increases too.

### The Two-dimensional Neural Lattice

The two-dimensional neural lattice corresponds to the two-dimensional Ising lattice. Assuming double indices to denote the element's position on the lattice, the neighbourhood is defined as

$$N(i, j) = \{ ([L + i - 1] \bmod L, j), \quad (3.14) \\ ([L + i + 1] \bmod L, j), \\ (i, [L + j - 1] \bmod L), \\ (i, [L + j + 1] \bmod L) \},$$

with  $L = \sqrt{N}$ . In contrast to its one-dimensional version, this neural lattice *does* exhibit a phase transition at  $T_c > 0$ . For  $H = 0$ , the partition function of the equivalent Ising lattice has been derived by Onsager (1944). The value of the critical intrinsic noise is therefore known:  $T_c^{2d} \approx 2.27J$  (Plischke and Bergersen, 1989). Unfortunately, the equilibrium properties in the presence of an external field cannot be computed exactly for the neural lattice in two dimensions (Chandler, 1987). Therefore, approximation techniques have been employed to establish the equilibrium magnetization of the two-dimensional Ising lattice in the presence of an external field (see, e.g., Nienhuis and Nauenberg, 1975). Qualitatively, the dependency of the equilibrium magnetization on  $T$  and  $H$  is similar to that of the one-dimensional Ising lattice. A quantitative evaluation of the magnetization in the presence of an external field is deferred to the simulation studies in Section 3.5.

### The Three-dimensional Neural Lattice

The three-dimensional neural lattice is, again, equivalent to the three-dimensional Ising lattice. The triple-indexed neighbourhood is defined as

$$N(i, j, k) = \{ ([L + i - 1] \bmod L, j, k), \quad (3.15) \\ ([L + i + 1] \bmod L, j, k), \\ (i, [L + j - 1] \bmod L, k), \\ (i, [L + j + 1] \bmod L, k), \\ (i, j, [L + k - 1] \bmod L), \\ (i, j, [L + k + 1] \bmod L) \}.$$

with  $L = \sqrt[3]{N}$ . The partition function of the three-dimensional Ising lattice has not been solved exactly (Yeomans, 1992). Nevertheless, numerical studies have established the properties of the three-dimensional lattice fairly precisely. The value of the critical intrinsic noise has been approximated to be  $T_c^{3d} \approx 4.52J$  (Plischke and Bergersen, 1989). As for the two-dimensional lattice, the shape of the magnetization function for  $T, H > 0$  is subject of our simulation studies.

### The N-dimensional Neural Lattice

The N-dimensional neural lattice corresponds to an Ising lattice in which each spin couples to all other spins. As stated above, such lattices can be treated exactly with the mean-field approach (e.g., Chandler, 1987). Assuming a static external input  $H = h(i)$  for all  $i$ , an arbitrary element  $i$  on the neural lattice receives as average internal input  $\langle h_{int}(i) \rangle$ :

$$\begin{aligned} \langle h_{int}(i) \rangle &= H + q^{-1} \sum_j \langle S(j) \rangle \\ &= H + \langle S \rangle, \end{aligned} \quad (3.16)$$

where  $\langle S \rangle$  is the average over all states  $S(j)$ ,  $j \neq i$ . This leads to (cf. (3.3))

$$\langle S \rangle = \tanh \left( \frac{H + \langle S \rangle}{T} \right). \quad (3.17)$$

If  $H = 0$  a non-zero value of  $\langle S \rangle$  corresponds to an ordered state. Figure 3.6 shows the graph of (3.17) for three different values of  $T$ . An intersection of a graph with the diagonal (dotted line) corresponds to a solution of (3.17). As can be seen in the figure, non-zero solutions appear only for  $T < 1$  (the arrows in Figure 3.6; Chandler, 1987). Consequently,  $T_c = 1$  for the N-dimensional neural lattice.

For  $H > 0$ , the curves in Figure 3.6 shift to the left. Consequently, non-zero solutions appear for values of  $T > 1$ .

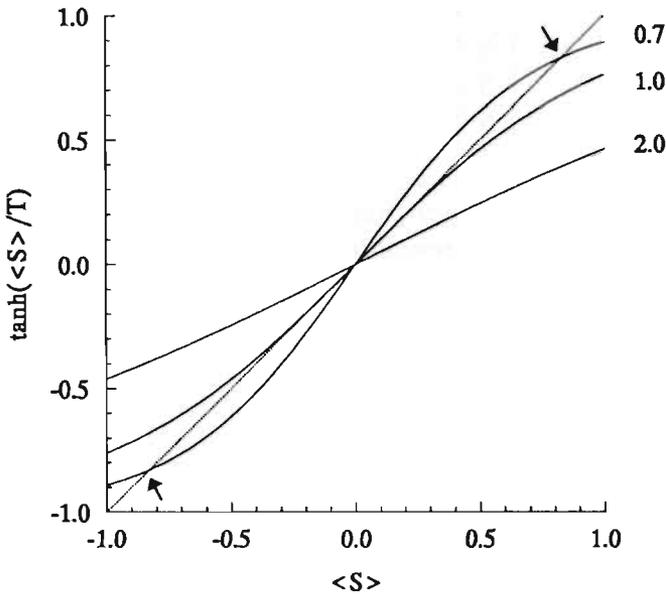


Figure 3.6: The mean-field equation for three values of  $T$ . The arrows point at the non-zero solutions for  $T = 0.7$ .

### 3.4.2 Critical Values and Connection Density

Table 3.1 lists the *normalized critical intrinsic-noise* values for the four neural lattices with normalized connection strengths  $J = q^{-1}$ .

Comparison of the four normalized values suggests that (at least for our lattice geometries) the amount of intrinsic noise needed to destroy an ordered state is proportional to lattice dimensionality. Alternatively, however, the proportionality may be due to the number of nearest neighbours  $q$ . It appears that both factors contribute to the proportionality. As a case in point Figure 3.7 displays the normalized critical values for four three-dimensional lattices (top curve) and three two-dimensional lattices (bottom curve).

The critical values for both the two- and three-dimensional lattices increase with the number of nearest neighbours  $q$ . Consequently, the number of nearest

$d$	$q$	$T_c(J = 1/q)$
1	2	0
2	4	$\approx 0.57$
3	6	$\approx 0.75$
N	N-1	1

Table 3.1: *Normalized critical intrinsic-noise values for the neural lattices of dimensionality  $d$ . ( $q$  is the number of nearest neighbours.)*

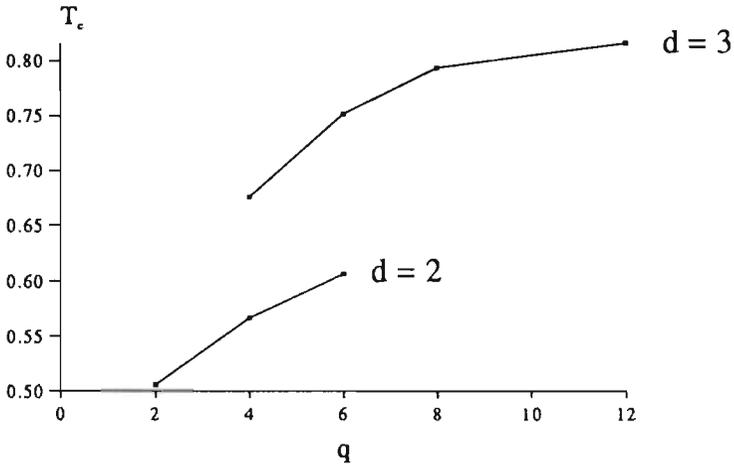


Figure 3.7: *Normalized critical intrinsic-noise values as a function of  $q$ . The top curve connect four values of  $T_c$  for three-dimensional lattice geometries. The bottom curve connects three values for two-dimensional lattice geometries. (Data points are taken from Baxter (1982).)*

neighbours is positively related to  $T_c$  for fixed dimensionality. Comparing the critical values of the two lattice dimensionalities at  $q = 4$  and  $q = 6$  signifies the positive contribution of lattice dimensionality when the number of nearest

neighbours is fixed.

The comparison of critical intrinsic-noise values reveals that increasing the connection density (either by enlarging the neighbourhood or by increasing lattice dimensionality) enhances the resistance to intrinsic noise. Having established that the normalized critical values associated with a range of connection densities map onto the unit interval, two striking features emerge. First, the transition from a one-dimensional to a two-dimensional neural lattice is associated with a moderate increase in connection density (four connections per element instead of two), whereas it yields a substantial increase in the resistance to intrinsic noise ( $\Delta T_c \approx 0.57$ ). Second, the transition from  $d = 3$  to  $d = N$ , associated with an enormous increase in connection density (number as well as length), leads to a relatively small increment in the resistance to intrinsic noise ( $\Delta T_c \approx 0.25$ ). These observations suggest that the robustness of the neural lattice is not linearly related to its connection density.

## 3.5 Simulation Studies

To allow for a quantitative comparison of the performance of the four neural lattices, simulation studies were performed. As stated in Section 3.1.3, the behaviour of the neural lattices is evaluated along two performance measures: *flexibility* (i.e., dependency on the initial state) and *robustness* (i.e., resistance to extrinsic noise). The flexibility of a neural lattice is assessed by initializing the lattice at  $t = 0$  in either an ordered or a disordered state and comparing the field-induced magnetization in either case. The robustness of a neural lattice is tested by presenting the lattice with input signals  $h(i)$ , taken from a Gaussian distribution (with mean  $\langle h \rangle$  and various values of the standard deviation  $\sigma$ ), and monitoring the field-induced magnetization. Our objective is to establish the relation between lattice dimensionality and the performance measures of flexibility and robustness.

### 3.5.1 Simulations with an Initial Disordered State

The first simulation deals with the behaviour of the neural lattices when starting from a disordered state, i.e., initially the number of *ON* and *OFF* elements are equal. The robustness of each of the four neural lattices is assessed by

presenting them with randomly distributed input signals.

### Procedure

The four lattices have the same size, i.e., the number of threshold elements,  $N$ . The one-dimensional lattice has a linear dimension of  $L = 16125 (= N)$ . The two- and three-dimensional lattices are defined as a square and cubic lattice with linear dimensions  $L_2 = 125$  ( $125^2 = N$ ) and  $L_3 = 25$  ( $25^3 = N$ ), respectively. The  $N$ -dimensional lattice contains  $N = 16125$  elements. To simulate the time-dependent behaviour of the neural lattices, we employ the following update rule (Glauber, 1963):

$$P(S(p) \rightarrow -S(p)) = \left[ 1 + \exp \left( -\frac{\Delta E_{NL}(S(p))}{T} \right) \right]^{-1} \quad (3.18)$$

where  $p$  is a vector containing the positional coordinates of the element.  $\Delta E_{NL}(S(p))$  is the change in energy induced by changing the state of the threshold element at position  $p$ :

$$\Delta E_{NL}(S(p)) = E_{NL}(\dots, S(p), \dots) - E_{NL}(\dots, -S(p), \dots), \quad (3.19)$$

The arguments of the two righthand-side terms differ only in the state  $S(p)$ . Substitution of equation (3.5) in (3.18) yields (cf. (3.6))

$$\Delta E_{NL}(S(p)) = -2S(p) \left( h(p) + J \sum_{p' \in N(p)} S(p') \right). \quad (3.20)$$

In the context of statistical mechanics, the Glauber rule is applied in Monte-Carlo simulations of the behaviour of, for instance, Ising systems (Binder, 1987). Glauber dynamics are also employed in neural networks to update stochastic threshold elements (see also, e.g., Ackley, Hinton, and Sejnowski, 1986; Amit, 1989; Hertz, Krogh, and Palmer, 1991).

At each *Monte-Carlo step per Spin* (MCS) or iteration, the Glauber updating rule (3.18) is applied to  $N$  randomly selected elements of the neural lattice. The random sequence is different for each iteration. A single simulation run

is defined as a consecutive run of 1000 MCSs. Each of the four lattices is presented with external input  $h(p)$  taken from a Gaussian distribution with mean  $\langle h(p) \rangle = 0.1$  and standard deviation  $\sigma \in \{0.0, 0.1, 0.2, 0.3, 0.4, 0.5\}$ . The value of  $\langle h(p) \rangle$  is chosen for two reasons. First, it is sufficiently smaller than the internal input  $J \sum S(p')$  (that may take values on the unit interval  $[0.0, 1.0]$ ) and does therefore not override the nearest-neighbour interactions responsible for ordering. Second, preliminary simulation studies indicated that for this value, a convergence is obtained within 1000 iterations according to our convergence criterion (see below). Simulations are performed for 10 different intrinsic-noise values:  $T \in \{0.05, 0.10, \dots, 0.50\}$ .

Lattice behaviour is monitored by computing the order parameter at the end of each MCS. Simulation runs start with  $m \approx 0.5$  (disordered state) by initializing each threshold element in the *ON* state with a probability of 0.5. All simulation runs are repeated 10 times.

## Results

The value of  $m$  at time  $t$  is denoted by  $m(t)$ . To evaluate the performance of the neural lattices we define the *convergence time*,  $t_{conv}$ , as the smallest value of  $t$  ( $T$  constant) for which the slope of a regression line through the values  $m(t)$  up to  $m(t + 100)$  has a value smaller than 0.001. Since our simulations run for 1000 iterations only, the largest convergence time is equal to 900. The *convergence magnetization*,  $m(t_{conv})$ , is the value of the magnetization at the convergence time. The four graphs in Figure 3.8 show the average convergence magnetizations for the four lattices as a function of the intrinsic-noise level.

Comparison of the four graphs in Figure 3.8 reveals that the overall average convergence magnetization increases with lattice dimensionality. Each graph contains five convergence-magnetization curves. From top to bottom these curves display the results for zero extrinsic noise  $\sigma = 0$  to maximal extrinsic noise  $\sigma = 0.5$ . A clear decrement in the average convergence magnetization as a function of  $\sigma$  is observed for all lattices. The decrement is strongest, however, for the one-dimensional lattice and becomes smaller with increasing dimensionality.

For the zero extrinsic noise ( $\sigma = 0$ ) curves, the largest convergence magnetization is obtained for the smallest value of the intrinsic noise  $T = 0.05$ . (In the zero intrinsic-noise limit all elements of the neural lattice operate

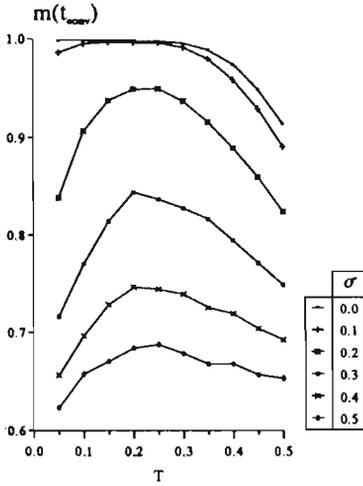
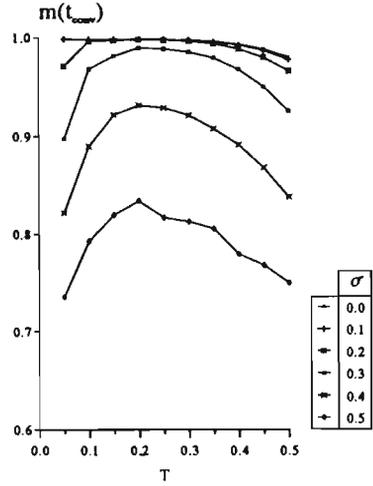
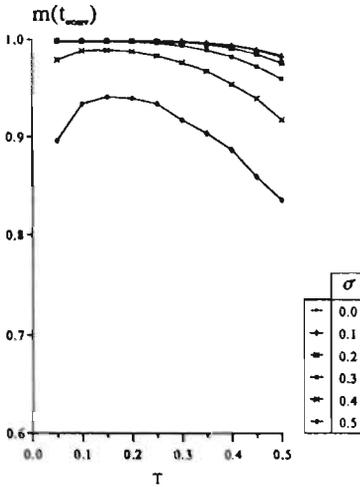
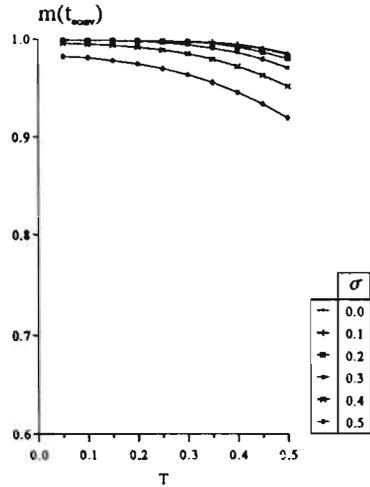
(a)  $d = 1$ (b)  $d = 2$ (c)  $d = 3$ (d)  $d = N$ 

Figure 3.8: Average convergence magnetization of the (a) one-dimensional, (b) two-dimensional, (c) three-dimensional, and (d)  $N$ -dimensional neural lattices with disordered initial states.

as deterministic threshold elements and turn therefore into their  $ON$  state:  $\lim_{T \rightarrow 0} m(t_{conv}) = 1.0$ .) With increasing extrinsic-noise values a maximum value of the convergence magnetization develops at  $T > 0.05$  for the first three graphs (Figures 3.8a-c). For the last graph (Figure 3.8d) the maxima for the five  $\sigma > 0$  curves are associated with the smallest intrinsic-noise level  $T = 0.05$ . The maximum values of the average convergence magnetization for each lattice and value of  $\sigma$  are listed in Table 3.2. Figure 3.9 displays these values as a function of lattice dimensionality. Each curve connects the data points associated with a single value of  $\sigma$ . Clearly, the resistance to extrinsic noise increases with lattice dimensionality.

$\sigma$	1d	2d	3d	Nd
0.0	0.9987 (0.0001)	0.9990 (0.0000)	0.9988 (0.0001)	0.9989 (0.0000)
0.1	0.9966 (0.0002)	0.9984 (0.0001)	0.9987 (0.0001)	0.9989 (0.0000)
0.2	0.9494 (0.0020)	0.9979 (0.0000)	0.9986 (0.0001)	0.9987 (0.0000)
0.3	0.8442 (0.0029)	0.9895 (0.0011)	0.9980 (0.0000)	0.9989 (0.0002)
0.4	0.7460 (0.0035)	0.9317 (0.0025)	0.9888 (0.0004)	0.9956 (0.0001)
0.5	0.6873 (0.0027)	0.8344 (0.0038)	0.9407 (0.0027)	0.9819 (0.0005)

Table 3.2: *Maximum values of the convergence magnetization for the four neural lattices. The entries are values averaged over ten equivalent simulation runs. The values between brackets are the standard errors of these average values.*

The effect of increasing the dimensionality of the neural lattice on the resistance to noise is expressed in the associated change in maximum convergence magnetization. By comparing the slopes of the four curves in can be seen that the gain in noise resistance, associated with increasing the lattice dimensionality, is largest for the transition from  $d = 1$  to  $d = 2$  and smallest for the transition from  $d = 3$  to  $d = N$ . In addition, the gain increases with larger values of  $\sigma$  for all transitions, except for  $\sigma = 0.5$  at the  $d = 1$  to  $d = 2$  transition.

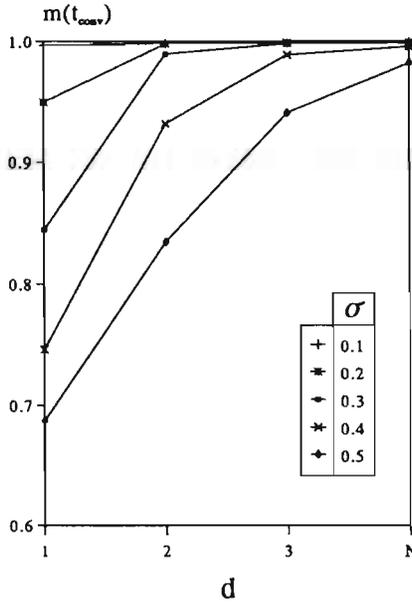


Figure 3.9: *Maximum convergence magnetizations as a function of lattice dimensionality.*

## Discussion

The main result of the simulations starting with a disordered state is that the robustness (i.e., the resistance to extrinsic noise) increases with lattice dimensionality. This is consistent with the positive relation between the critical intrinsic-noise value  $T_c$  and lattice dimensionality  $d$  (cf. Table 3.1). Larger values of  $T_c$  imply stronger bonds among neighbouring threshold elements that are harder to break by either intrinsic or extrinsic noise.

The emergence of maxima for  $\sigma > 0$  and  $1 \leq d \leq 3$  can be understood as follows. In the limit of zero intrinsic noise  $T \rightarrow 0$ , the threshold elements follow their input in a deterministic way. In the presence of extrinsic noise such deterministic behaviour increases the probability of groups of threshold elements

in their *OFF* state. Interestingly, intrinsic noise enhance the robustness to extrinsic noise. Starting from the deterministic limit, increasing the intrinsic-noise level enables individual threshold elements to “recover” from their wrong state through the stochastic updating. However, when the intrinsic-noise level grows too large, the increased randomness in updating distorts the formation of ordered states. This explains the appearance of maxima in Figure 3.8.

To assess the contribution of the connections we recall that the state of an arbitrary threshold element  $i$  on the neural lattice is determined by  $h_{ext}(i) + h_{int}(i) = h(i) + q^{-1} \sum_j S(j)$ . The distribution of  $h_{ext}(i)$  is Gaussian. With  $\sigma$  known, the probability of  $h_{ext}(i)$  being smaller than some specified value  $x$ ,  $P(h_{ext}(i) < x)$  can be computed. The distribution function  $F(x)$  of the Gaussian specifies the area under the curve from minus infinity to  $x$ :

$$F(x, \langle x \rangle, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x \exp\left(\frac{1}{2} \left[\frac{v - \langle x \rangle}{\sigma}\right]^2\right) dv. \quad (3.21)$$

The probability of an external input smaller than  $x$  is therefore

$$P(h_{ext}(i) < x) = F(x, \langle h \rangle, \sigma) = \Phi\left(\frac{x - \langle h_{ext}(i) \rangle}{\sigma}\right), \quad (3.22)$$

where  $\Phi$  is the distribution function of a “standard” Gaussian distribution with zero mean and unit variance, i.e.,  $\Phi(z) = F(x, 0, 1)$ , ( $z$  is the standardized value). The value of  $\Phi(z)$  can be obtained from tables in statistical texts (e.g., Kreyszig, 1970). The probability of a negative external input,  $P(h_{ext}(i) < 0)$ , is 0.16, 0.31, 0.37, 0.40, and 0.42, for  $\sigma = 0.1, 0.2, 0.3, 0.4$ , and 0.5, respectively. In a deterministic neural lattice without connections ( $J = 0$ ), the threshold elements follow the sign of their input. Consequently, the maximal convergence value is equal to  $1 - P(h_{ext}(i) < 0)$ . It follows that for  $J = 0$  the maximal convergence values are 0.84, 0.69, 0.63, 0.60, and 0.58, for  $\sigma = 0.1, 0.2, 0.3, 0.4$  and 0.5, respectively. Comparison of these values with those in Table 3.2 proves the positive contribution of the connections.

### 3.5.2 Simulations with an Initial Ordered State

The second simulation deals with the behaviour of the neural lattice when starting from an ordered state. The objective is to assess the flexibility of neural lattices. Instead of random initialization, in these simulations all elements are set to their *OFF* state at  $t = 0$ .

#### Procedure

The same procedure is followed as in the first simulations except for the initialization and the intrinsic-noise values used. For all simulation runs  $m(0) = 0.0$ . Simulations are performed for nineteen intrinsic-noise values:  $T \in \{0.10, 0.15, \dots, 1.00\}$ .

#### Results

The average convergence magnetizations obtained from the simulations with the lattices starting in the ordered state  $m(0) = 0.0$  are displayed in Figure 3.10.

The pattern of results differs significantly from the results of the simulations starting from a disordered state. A jump in the magnetization is observed for all lattices. We define the value of  $T$  where the jump takes place as the jump intrinsic-noise value  $T_{jump}$ . On the one hand, comparison of the  $\sigma = 0$  (top) curves in Figures 3.8a-d reveals that  $T_{jump}$  grows with lattice-dimensionality. On the other hand, within each graph,  $T_{jump}$  decreases with increasing  $\sigma$ .

The maximum value of the convergence magnetization for each lattice and value of  $\sigma$  is listed in Table 3.3 and displayed in Figure 3.11 as a function of lattice dimensionality.

When the extrinsic noise is small ( $\sigma = 0.0$  and  $0.1$ ), the largest convergence magnetization is obtained for the one-dimensional lattice. With increasing  $\sigma$ , the lattice dimensionality at which the largest magnetization is obtained grows too. For  $\sigma = 0.2$ , the two-dimensional lattice outperforms the others, whereas for  $\sigma = 0.3$  and  $0.4$  the three-dimensional lattice performs best. For the largest value of  $\sigma$  the  $N$ -dimensional lattice outperforms the others. It

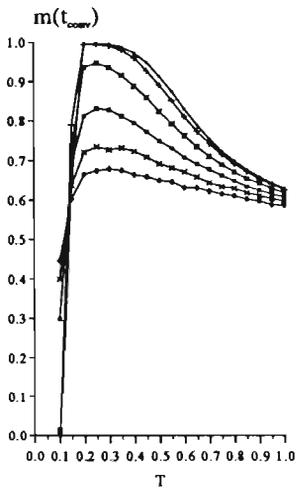
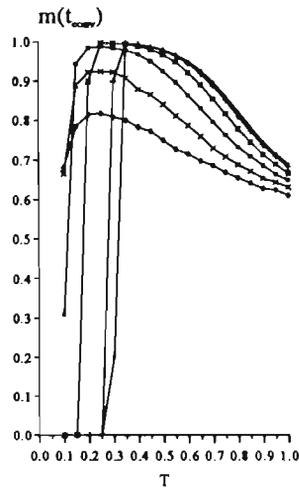
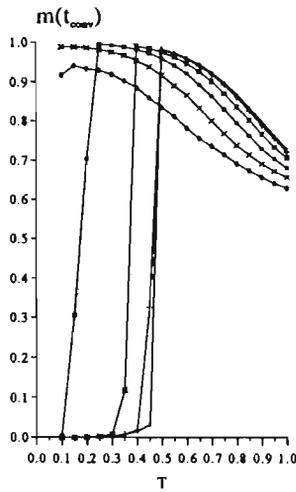
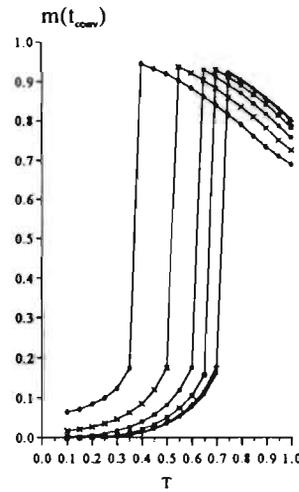
(a)  $d = 1$ (b)  $d = 2$ (c)  $d = 3$ (d)  $d = N$ 

Figure 3.10: Average convergence magnetization of the (a) one-dimensional, (b) two-dimensional, (c) three-dimensional, and (d)  $N$ -dimensional neural lattices with ordered initial states.

$\sigma$	1d	2d	3d	Nd
0.0	0.9971 (0.0001)	0.9960 (0.0000)	0.9830 (0.0001)	0.9260 (0.0002)
0.1	0.9945 (0.0004)	0.9955 (0.0000)	0.9814 (0.0001)	0.9228 (0.0001)
0.2	0.9461 (0.0017)	0.9973 (0.0001)	0.9903 (0.0001)	0.9302 (0.0003)
0.3	0.8305 (0.0021)	0.9880 (0.0010)	0.9955 (0.0001)	0.9309 (0.0003)
0.4	0.7342 (0.0047)	0.9242 (0.0031)	0.9890 (0.0014)	0.9385 (0.0003)
0.5	0.6782 (0.0025)	0.8176 (0.0046)	0.9416 (0.0019)	0.9455 (0.0005)

Table 3.3: *Maximum values of the convergence magnetization for the four neural lattices. The entries are values averaged over ten equivalent simulation runs. The values between brackets are the standard errors of these average values.*

should be noted that in the latter case the convergence magnetization falls below 0.95 and is expected to deteriorate further for  $\sigma > 0.5$ .

Comparison of the slopes in Figure 3.11 indicates that the largest gain is associated with the transition from  $d = 1$  to  $d = 2$  except for  $\sigma = 0.1$  which is associated with a small loss. The transition from  $d = 2$  to  $d = 3$  yields a gain for  $\sigma = 0.3$  to  $0.5$  and a loss for  $\sigma = 0.1$  and  $0.2$ . The transition to the  $N$ -dimensional lattice is associated with a loss for all values of  $\sigma$  except for a small gain for  $\sigma = 0.5$ . It should be noted that the dimension  $d$  at which the maximum convergence magnetization is obtained seems to increase with  $\sigma$ .

## Discussion

The main result of the simulations starting from an ordered state is that the flexibility, i.e., the ease with which a lattice turns from one ordered state to another, depends on both lattice dimensionality and amount of extrinsic noise. When the extrinsic noise is small ( $\sigma = 0.0$  or  $0.1$ ), the one-dimensional neural lattice outperforms the other lattices. For intermediate extrinsic noise ( $\sigma = 0.3, 0.4$ ), the two- and three-dimensional lattices achieve the largest convergence magnetizations. For the largest extrinsic noise value, the  $N$ -dimensional neural lattice performs best. Again, this is a consequence of the increase of the critical intrinsic-noise value  $T_c$  with the lattice dimensionality.

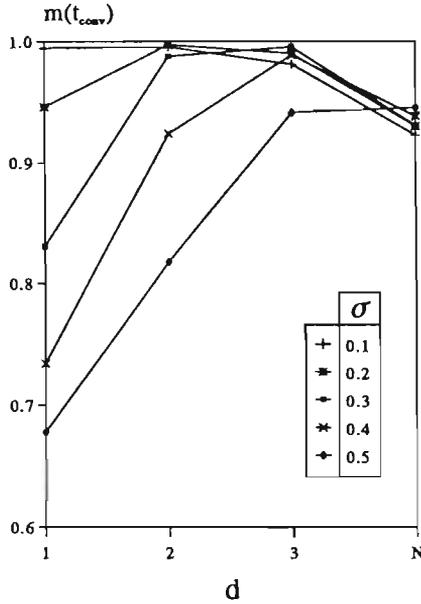


Figure 3.11: *Maximum convergence magnetizations as a function of lattice dimensionality.*

Since the value of  $T_c$  is an indicator of the strength with which the ordered state is maintained, it also indicates the difficulty with which an ordered state is destroyed. Changing from one ordered state into another, as required in the simulations, implies that the initial order has to be destroyed. In the zero intrinsic-noise limit  $T \rightarrow 0$  and  $\sigma = 0$  the input of each threshold element  $h_{ext}(i) + h_{int}(i)$  equals  $h(i) + q^{-1} \sum_j S(j) = 0.1 - 1.0 = -0.9$ . Due to this negative input, the threshold elements will not change their initial state. With increasing  $T$ , the probability of threshold elements flipping into their *ON* state (despite their negative input) increases. When  $h_{ext}(i) = 0$  and consequently  $h_{ext}(i) + h_{int}(i) = -1.0$ , an ordered state is destroyed at  $T = T_c$ . A positive value of  $h_{ext}(i)$  “helps” to destroy the ordered state by increasing the net input (to  $-0.9$  for  $h_{ext} = 0.1$ ). Consequently, the temperature at which the order is destroyed is expected to be smaller than  $T_c$ . As can be seen in Figure 3.10

this is indeed the case for  $d \leq 2$ . For the  $\sigma = 0$  curves,  $T_{jump}$ , the value of the intrinsic noise needed to destroy the initial ordered state, is  $\approx 0.35, 0.50$ , and  $0.75$ , for the two-, three-, and  $N$ -dimensional lattice, respectively.

In the presence of extrinsic noise ( $\sigma > 0$ ), the ordered state is more easily destroyed. Whenever the external input is larger than the internal input, the probability that the corresponding element turns into the *ON* state exceeds 0.5. Since the probability of a positive net input (i.e.,  $h_{ext} + h_{int} > 0$ ) increases with  $\sigma$ , the value of  $T_{jump}$  decreases. When the extrinsic noise grows too large, the convergence magnetization drops, as is clearly visible for all lattices.

### General Discussion

The simulations performed with disordered and ordered initial states yield the following two main results. First, the robustness of neural lattices to extrinsic noise increases with lattice dimensionality. Second, the flexibility of neural lattices starting from an ordered state is a function of the lattice dimensionality and the extrinsic noise value. If the extrinsic noise is zero, the best flexibility is obtained with the one-dimensional neural lattice. If the extrinsic noise is large, however, the  $N$ -dimensional lattice has the best flexibility. To get the best combination of the properties of robustness and flexibility, one should take into account the value of  $\sigma$  and the lattice dimensionality.

## 3.6 Summary and Conclusion

We have investigated the behaviour of three sparsely-connected neural lattices and a single fully-connected lattice. In particular, we have assessed their sensitivity to static extrinsic input noise and their dependency on the initial state. Simulation studies of the neural lattices enabled a characterisation of lattice performance as a function of its connection density. It turns out that while the fully-connected lattice exhibits the best robustness, the one-dimensional lattice shows the best flexibility in the absence of extrinsic noise. From our findings we conclude that full connectivity in biological or artificial systems, does not necessarily provides the best combination of flexibility and robustness. Given the problems associated with scaling up fully-connected neural networks (Minsky and Papert, 1988; Nelson and Bower, 1990) we conclude

that in the neural lattice only a relatively small price is to be paid when full connectivity is abandoned and a sparser connectivity pattern is used instead.

In the following chapter we introduce the *gating lattice*, an extension of the neural lattice, to enable selection of one out of three input channels. In line with the conclusion, the gating lattice is defined as a sparsely-connected two-dimensional lattice. As the gating lattice is the building block of the SCAN model, its sparse connectivity enhances the scalability of SCAN.



## Chapter 4

# The Gating Lattice

A selection mechanism giving the allocated resources access to a part of the visual data represents an inevitable component of any model of vision. We interpret resource allocation as the application of a central pattern recognizer to a selected location in a two-dimensional visual input field. The allocation of “recognition resources” to the appropriate location enables the identification of individual objects (i.e., characters) irrespective of input position. A parsimonious way of allocating recognition resources to a selected location is through *pattern routing* (Anderson and Van Essen, 1987). An example of pattern routing is shown in Figure 4.1. Given an input field containing an object pattern P, its content may be routed towards the input of a pattern recognizer. Translations of the object pattern over the input field, i.e., the movement of the pattern to a new location P', are compensated for by a routing mechanism. Consequently, routing enables the active selection of contiguous parts of the input field as well as for position-invariant pattern recognition. It therefore represents a convenient model for covert attention, since location-based selection and translation invariance are also features of the covert-attention process (see Chapter 2). In this chapter, we focus on the building block that provides routing in SCAN.

Pattern routing requires a mechanism capable of selecting and gating part of the input field to the input of the recognizer. A neural-network model called the *gating lattice* (Postma, Van den Herik, and Hudson, 1992a, 1993a) provides such a mechanism. The goal of the model is to allow for the gating of a spatially contiguous pattern without distorting its spatial order. The gating lattice is

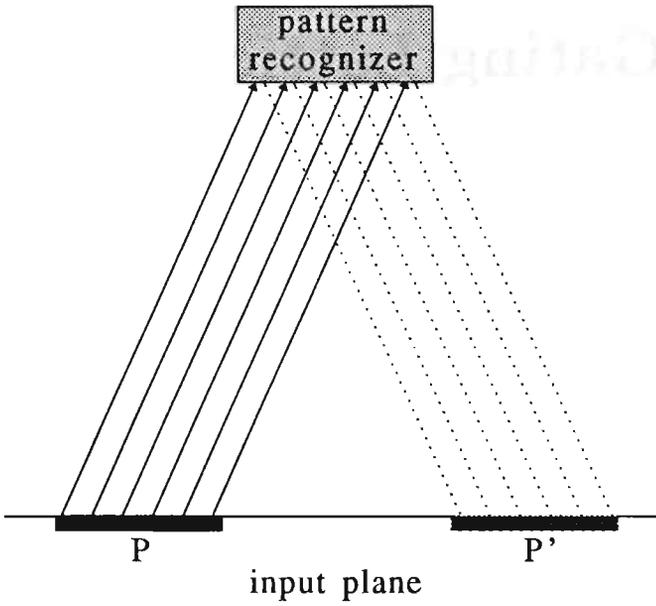


Figure 4.1: *An illustration of pattern routing.*

an extension of the neural lattice and is capable of selecting and gating one out of three parallel input streams. We show that the gating lattice, like the neural lattice, can be formalised as an Ising lattice. Formal analysis and simulation studies enables us to understand its macroscopic behaviour and determine the parameter values appropriate for flexible and robust gating.

We start with an informal characterisation of the gating lattice. Section 4.2 reviews the statistical mechanics of antiferromagnetic Ising lattices, leading to a formalisation of the gating lattice as an antiferromagnet in Section 4.3. Then, in Section 4.4, we discuss the effects of limited size and time on gating performance. Section 4.5 contains simulation studies of the gating lattice. Finally, Section 4.6 provides a discussion and concludes that the gating lattice represents an appropriate building block for pattern routing.

## 4.1 A Parallel Distributed Switch

Our study of the performance of neural lattices demonstrated that the optimal trade-off of robustness and flexibility is obtained with lattices of intermediate dimension when the extrinsic noise level is not too large (cf. Section 3.5.2). Our conclusion was that  $d = 2$  and  $d = 3$  are appropriate dimensions. With some emphasis on flexibility and scalability, which both decrease with increasing dimensionality, we define the gating lattice as a two-dimensional lattice with competitive (inhibitory) couplings. It acts as a parallel switch capable of selecting one out of  $n$  two-dimensional patterns (e.g., images); after selection the pattern is gated to a common output. For expositional purposes, we describe an example of a gating lattice involving the selection and gating of a one-dimensional input pattern. We consider a one-dimensional input pattern  $I$  with length  $p + n - 1$ . The elements of  $in$  are indexed as  $in(i)$  ( $i \in \{1, 2, \dots, p + n - 1\}$ ). The output of the gating lattice is a pattern  $out$  that has length  $p$  and that is indexed as  $out(j)$  ( $j \in \{1, 2, \dots, p\}$ ). In our example we assume two switching directions ( $n = 2$ ). Consequently, the gating lattice can select one of two contiguous subpatterns of  $in$  and gate their contents to the output  $out$ . It follows that the output pattern is defined as, either  $out(k) = in(k)$  for  $k \in \{1, 2, \dots, p\}$ , or  $out(k) = in(k + 1)$  for  $k \in \{1, 2, \dots, p\}$ . This is illustrated in Figure 4.2. The two arrays of circles represent the elements of  $in$  (bottom array) and  $out$  (top array). The gating lattice is represented by the tilted (solid and dashed) lines. The solid lines represent the

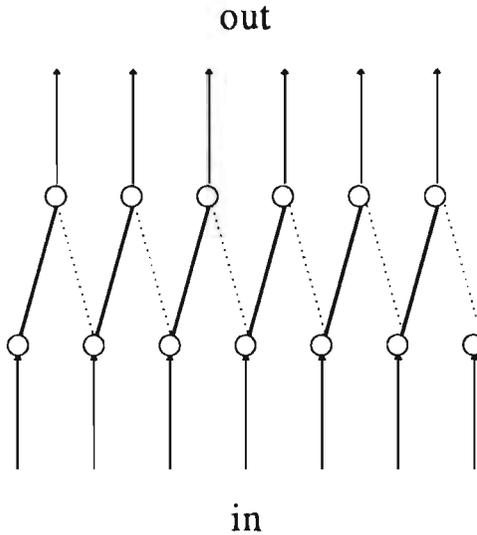


Figure 4.2: *An example of a parallel switch with  $n = 2$  switching directions.*

current switching direction, while the dashed lines represent the alternative switching direction. Pairs of tilted lines, i.e., inverted V's, may be interpreted as two-state switches. Because the switches are assumed to operate in parallel they form a parallel distributed switch. In the general case, the gating lattice has  $n$ -state switches, instead of two-state switches. Moreover, the input and output patterns have a two-dimensional structure and the switches in the gating lattice are arranged in a two-dimensional lattice.

### 4.1.1 Gate Structure

The elements of the gating lattice are called gates; they are binary elements that can be open or closed. Referring to the example in Figure 4.2, an open

gate corresponds to a solid tilted line while a closed gate corresponds to a dashed tilted line. Each gate is therefore responsible for the gating of a single pattern element (e.g., pixel). An example of a gate is depicted in Figure 4.3. Two neural elements make up the gate: a *signal element* (triangle) and a *control element* (circle). The signal element allows a signal to be gated in or out depending on its state. The control element is the threshold element also employed in the neural lattice; it controls the state of the signal element. Since gates are part of a two-dimensional structure, their position is represented by a double index. A gate at position  $(i, j)$  is denoted as  $g(i, j)$  and allows signal flow from its input  $in(i, j)$  towards its output  $out(i, j)$ . The state of a gate is denoted by  $G(i, j)$  (a spin variable with a value  $\in \{-1, +1\}$ ) and is controlled by an external control signal  $h(i, j)$ .  $G(i, j)$  is part of the control element. For  $G(i, j) = +1$ , the control element inhibits the signal element. Consequently, signal flow is prevented and  $out(i, j) = 0$ , i.e., the gate is closed. For  $G(i, j) = -1$ , the control element is inactive and the gate is open, then  $out(i, j) = in(i, j)$ . In general, the output of the gate at position  $(i, j)$  is given by

$$out(i, j) = \left[ 1 - \frac{1}{2}(G(i, j) + 1) \right] in(i, j). \quad (4.1)$$

We remark that for a complete description of gate behaviour, the only variables of relevance are the gate state  $G(i, j)$  and its control signal  $h(i, j)$ . Consequently, the gating lattice incorporates a special case of the neural lattice.

### 4.1.2 Gate Coupling

A gating lattice is formed by coupling gates with gates in their neighbourhood. Depending on the definition of the neighbourhood, this leads to a particular lattice geometry. We assume gating lattices to be *regular*. Lattices are called regular when the distance between all nearest neighbouring elements is equal. Figure 4.4 shows the three planar regular lattices: the *square*, the *hexagonal* (or honeycomb), and the *triangular* lattice. (In the previous chapter, only the square two-dimensional lattice geometry was considered.)

In the gating lattice, the neighbourhood  $N(i, j)$  of the gate at position  $(i, j)$  contains nearest neighbours of that gate only and is independent of gate position except for gates at the borders of the lattice (see below).

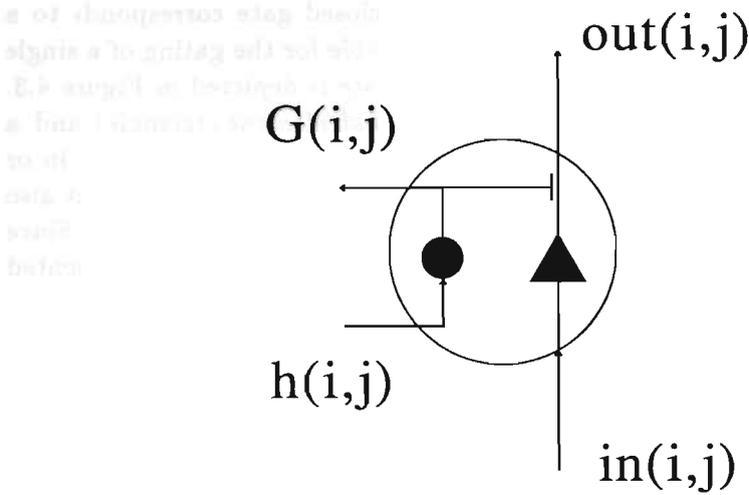


Figure 4.3: *The gate with its two neural elements: a control element (circle) and a signal element (triangle).*

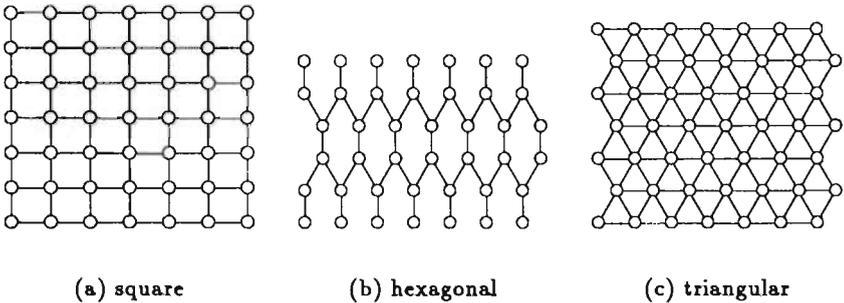


Figure 4.4: *The three regular planar lattices; (a) the square, (b) the hexagonal, and (c) the triangular lattice.*

### 4.1.3 Sublattices

An important concept is that of *sublattice*. A gating lattice contains  $n$  sublattices, each of which is responsible for the gating of one input pattern. We

define a sublattice as follows.

Let  $\Omega$  represent the set of all gates in an unbounded gating lattice, with  $\omega \in \Omega$  a subset. Also let the neighbourhood  $N(i, j)$  of each element  $g(i, j)$  of  $\Omega$  contain the nearest neighbours of  $g(i, j)$ .

**Definition 4.1** *A subset  $\omega$  of  $\Omega$  is called a sublattice if  $\omega$  is the largest subset of  $\Omega$  for which it is true that each member of  $\omega$  is not in the neighbourhood of any other member.*

With the above definition, a gating lattice can be partitioned into sublattices as follows. Consider as an example the partitioning of a square gating lattice into sublattices. Figure 4.5 shows a part of a square lattice with gates displayed as circles. We start by selecting an arbitrary gate  $g(i, j)$ , e.g., the gate in the centre of Figure 4.5. We assume that  $g(i, j)$  is member of  $\omega$  and mark this in the figure by a filled circle. Since  $g(i, j)$  is a member of  $\omega$ , its nearest neighbours are not. Consequently the neighbours of  $g(i, j)$  are left unmarked. Since we are after the largest subset of gates  $\omega$ , we proceed to the next-to-nearest neighbours of  $g(i, j)$ . These gates *can* be members of  $\omega$  and are marked accordingly. Continuation of this procedure leads to the partitioning of the square gating lattice into two regular sublattices. One sublattice is formed by all marked gates, the other is formed by all unmarked gates. In case there are three sublattices, as in the triangular lattice, after the first sublattice has been determined (marked), the second can be found by repeating the procedure for a gate in the neighbourhood of  $g(i, j)$ .

Figure 4.6 shows the three regular planar lattices where gates (circles) are shaded according to the sublattice of which they are part. Both the square and honeycomb lattice contain two sublattices (black and white), while the triangular lattice contains three sublattices (black, grey, and white). We remark that for each lattice, the number of gates on each sublattice is equal.

In case of bounded lattices (i.e., lattices containing a finite number of gates), two additional assumptions are required. The first assumption is that periodic boundary conditions are enforced on the lattice. With periodic boundary conditions the gating lattice becomes a toroidal structure. Second, the total number of gates in the lattice should be such that a continuation of sublattice structure is accommodated. This means that the total number of gates on the lattice  $N$  has to be a multiple  $n$  of the number of gates on a sublattice  $p$ , i.e.,

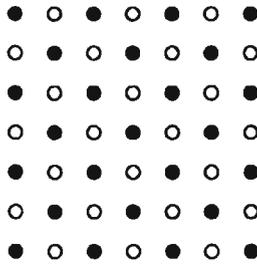
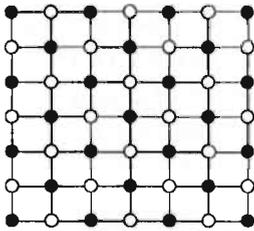
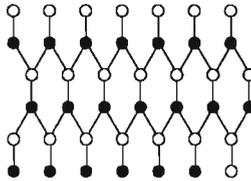


Figure 4.5: *Partitioning of the square lattice into two sublattices.*

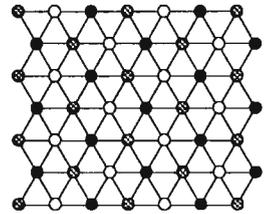
$$N = np.$$



(a) square



(b) hexagonal



(c) triangular

Figure 4.6: *Sublattices of (a) the square, (b) the hexagonal (b), and (c) the triangular lattice.*

In the gating lattice, each sublattice is responsible for the gating of a single input pattern. A gating lattice with  $n$  sublattices is capable of gating one out of  $n$  patterns. The inputs of the gates on one sublattice are the elements or pixels of the image. Images are mapped topographically onto the sublattices so that their spatial structure is maintained when gated.

In the following, we restrict our discussion to the *triangular gating lattice*, i.e., the gating lattice defined on the triangular lattice. A gating lattice based on this geometry has three sublattices ( $n = 3$ ), and can therefore select one out

of three input patterns.

#### 4.1.4 Synchronicity and Exclusivity

For the proper operation of a gating lattice, capable of selecting and gating one of  $n$  patterns, two requirements should be fulfilled: *synchronicity* and *exclusivity*. The synchronicity requirement states that all gates associated with one of the  $n$  patterns are synchronized, i.e., they are *all* open or *all* closed. Synchronicity assures that an input pattern is gated in its entirety.

The exclusivity requirement states that if the gates associated with one input pattern are open, all gates associated with the  $n - 1$  other patterns are closed. Exclusivity prevents the output of the gating lattice being a mixture of two or more input patterns. When both the synchronicity and exclusivity requirements are fulfilled, a gating lattice can be applied to the routing of visual patterns.

#### 4.1.5 Distributed Competition

In the (triangular) gating lattice, the synchronicity and exclusivity requirements are locally fulfilled by defining competitive couplings between a gate and its neighbouring gates. These couplings behave in such a way that neighbouring gates tend to have opposite states, i.e., open gates have neighbouring closed gates and vice versa. Competitive gate coupling facilitates the opening of all gates on one sublattice and the closing of all other gates. Locally, gates engage in a Winner-Take-All (WTA) competition (Grossberg, 1973; Feldman and Ballard, 1982), yielding a single open gate as the winner. The gating lattice employs *distributed competition* to achieve the opening of all gates on one sublattice, the *winning sublattice*. We define a *valid gating state* as the lattice state in which all gates on the winning sublattice are open and all other gates are closed.

In the following section, we discuss the antiferromagnetic triangular Ising lattice that turns out to be formally equivalent to the triangular gating lattice. Consequently, the formal techniques employed in analyzing antiferromagnetic Ising lattices can be applied to this gating lattice.

## 4.2 Statistical Mechanics of Antiferromagnets

When the coupling constant  $J$  of the Ising lattice is negative, one obtains a model of an antiferromagnet. Referring to the Ising-lattice energy function (3.8), the first righthand term is minimized when all spins are anti-aligned. Anti-alignment leads to a partitioning of the antiferromagnetic Ising lattice into groups of up and down spins, or to be more precise, into sublattices. Figure 4.6 can be interpreted as showing the sublattices of the square, hexagonal, and triangular antiferromagnetic Ising lattices. For the square and honeycomb Ising lattices, one sublattice is formed by up spins while the other is formed by down spins. For the triangular Ising lattice one (or two) are formed by up spins whereas, two (or one) are formed by down spins. In antiferromagnetic Ising lattices there is an intrinsic symmetry concerning the different sublattices.

### 4.2.1 The Phase Diagram

The macroscopic behaviour of antiferromagnetic Ising lattices is best presented in phase diagrams. Figure 4.7 shows the phase diagram for the triangular Ising lattice.

In the figure, the shaded region corresponds to values for  $T$  and  $H$  for which the antiferromagnetic Ising lattice is in an ordered phase, whereas the unshaded region corresponds to values for which it is in a disordered phase. The critical points  $T_c$  demarcate the border between two phases.

We have not yet fully addressed the role of the external field  $H$ . As the phase diagram illustrates, the ordered phase also becomes disordered for certain large values of  $H$ . This can be understood as follows. Let us take the square antiferromagnetic Ising lattice as an example. It has an analogous phase diagram as the triangular lattice (cf. Figure 4.7). In the ordered phase each up spin is surrounded by four down spins. Consequently, the contribution of this spin to the energy is  $-4J$ . A state where this spin is flipped into its down state contributes to the energy with  $+4J$  and is therefore unlikely at low temperatures. Taking the external field into account, the contribution of the up spin to the energy equals  $-4J - H$ , whereas the contribution of the down spin equals  $+4J + H$ . Let us now assume that  $H = -4J$ , both contributions to the energy are zero. Consequently, both states have an equal probability. For  $H < -4J$ , the state with all spins down is energetically favourable. The same

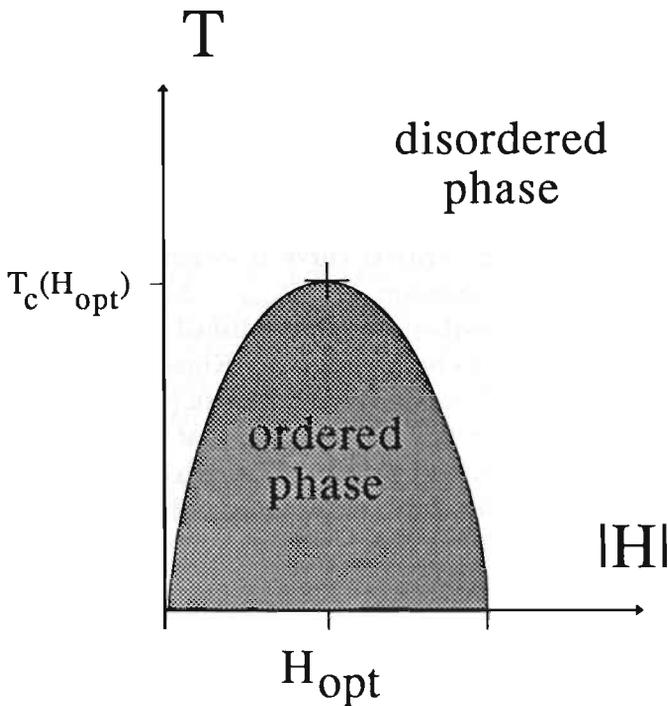


Figure 4.7: Phase diagram of the antiferromagnetic triangular Ising lattice

account applies to the case where a down spin is surrounded by up spins and  $H > +4J$ . Consequently, when  $H$  is too small or too large it overrides the nearest-neighbour couplings  $J$  and “draws” all spins downwards or upwards. For the square Ising lattice, the optimal value of  $H$ ,  $H_{bias}$ , is equal to zero and is associated with the highest critical temperature  $T_c$ .

#### 4.2.2 Closed Form of the Critical Curve

The border between the ordered and disordered phase in Figure 4.7 is known as the *critical curve*, i.e., the function  $T_c(H)$ . An approximate analytical expression of  $T_c(H)$  for the square antiferromagnetic Ising lattice has been derived by Müller-Hartmann and Zittartz (1977). Building upon this result,

Akutsu and Akutsu (1990) arrived at an approximation to the critical curve of the triangular Ising lattice:

$$2 \cosh \left( \frac{2J}{T} - \frac{2H}{3T} \right) = \frac{2J}{T}. \quad (4.2)$$

According to this result, the critical curve is symmetric around the  $H = 3$  line where  $T_c$  reaches its maximum, i.e.,  $H_{bia_s} = 3.0$  for the triangular Ising lattice. We note that an earlier study established the critical curve of the triangular antiferromagnet to be asymmetric (Kinzel and Schick, 1981). By improving their approximation, Akutsu and Akutsu (1990) arrived at a slightly more complex expression for the critical curve that is in agreement with the asymmetry found by Kinzel and Schick. The maximum value of  $T_c$  is now associated with  $H = 3.1$  (two-thirds of the spins up) or  $H = -3.1$  (two-thirds of the spins down).

Given this quantitative result on the antiferromagnetic Ising lattice and our qualitative understanding of its macroscopic behaviour, we are now able to formulate the gating lattice in a statistical-mechanics framework.

### 4.3 Statistical Mechanics of Gating Lattices

We formalise the triangular gating lattice according to the antiferromagnetic triangular Ising lattice. Instead of the spin variable we employ the gate variable  $G$ . Moreover, we index gates according to their position on the sublattice:  $G_x(i, j)$  where  $x$  refers to the sublattices labelled A, B, and C ( $x \in \{A, B, C\}$ ). The external field  $H$  is redefined in the gating lattice as an external *control signal* and is partitioned into the sublattice control signals  $H_x$ . The energy function for the gating lattice is

$$E_{GL}(\mathbf{G}) = \sum_{x \in \{A, B, C\}} \left( \sum_{i, j} G_x(i, j) \sum_{k, l \in N(i, j)} G_y(k, l) - (H_{bia_s} - H_x) \sum_{i, j} G_x(i, j) \right), \quad (4.3)$$

where  $\mathbf{G}$  is the vector containing all gate states and  $H_{bias}$  denotes a fixed bias. The neighbourhood  $N(i, j)$  contains six elements that are not part of sublattice  $x$  ( $y \neq x$ ) and will be defined in Section 4.5. The biases are set to the value of  $H$  for which  $T_c$  is maximal according to (4.2). The value of  $H_{bias}$  maximizes the value of the last term of (4.3) when one-third of the gates is open and two-thirds closed, i.e.,  $H_{bias} = 3.1$ . Analogous to the energy function of the antiferromagnetic Ising lattice, the first righthand side term of (4.3) is minimized if neighbouring gates are pairs of open and closed gates. Again there is an intrinsic symmetry in this term in that for  $H_x = 0$  (for all  $x$ ) the global minima, i.e., the three valid gating states, have equal energy values. Given non-zero control signals the last term breaks the symmetry by lowering the energy associated with the corresponding valid gating state.

The state of the gating lattice is expressed in the order parameter  $m$  that we (re)define for the gating lattice as

$$m = \frac{1}{2} [\#_A - (\#_B + \#_C) + 1], \quad (4.4)$$

where  $\#_x$  denotes the proportion of open gates on sublattice  $x$ . It should be remarked that through defining the gating-lattice order parameter in this way, we implicitly assume that the order parameter is maximal for the valid gating state with all gates on sublattice A open. In the simulation studies described below, we aim at the target state of  $m \approx 1.0$  where (the winning) sublattice A is open and both the B and C sublattice are closed. In the case that either sublattice B or sublattice C is open and the others are closed, the order parameter is zero ( $m = 0.0$ ). Because of the triple symmetry of the gating lattice our results generalize to the other cases where B or C are the winning sublattice.

For similar reasons, we (re)define the control signal  $H$  as

$$H = H_A = -H_B = -H_C. \quad (4.5)$$

A positive value of  $H$  implies that the control input to the gates on sublattice A is positive and the input to the gates on sublattices B and C is negative.

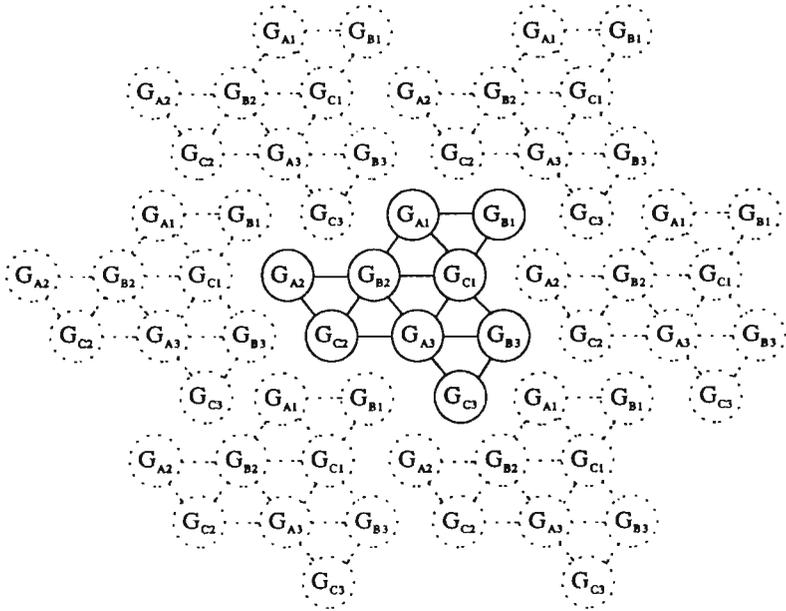


Figure 4.8: *Gating lattice used for the computation of the partition function. Solid circles are gates. The dashed circles illustrate the periodic boundary conditions (cf. Schick, Walker, and Wortis (1976, 1977)).*

### 4.3.1 Calculation of Gating-Lattice Partition Function

In Chapter 3, we saw that the probability of a state can be computed directly when the partition function  $Z$  is known. In this section we compute the partition function for a gating lattice containing  $3N = 9$  gates and analyze the probability of a valid gating state as a function of  $T$ ,  $H_{bias}$ , and  $H$ . For the small gating lattice, we employ periodic boundary conditions leading to the lattice structure shown in Figure 4.8. The same lattice was used by Schick, Walker, and Wortis (1976, 1977), in their renormalization-group analysis of the triangular Ising lattice.

Our objective is to compute  $P_{\nu,gs}$ , the probability of a valid gating state. For

our small gating lattice, states are represented by a nine-digit string  $s$  of 1s and 0s, for open and closed gates respectively. The nine digit positions in the string represent the gate states  $G_{A1}, G_{B1}, G_{C1}, G_{A2}, G_{B2}, G_{C2}, G_{A3}, G_{B3}, G_{C3}$ , respectively. (The subscripts correspond to the labels in Figure 4.8.) The valid gating states are represented by the strings 100100100, 010010010, and 001001001. We define  $P_{vgs}$  as:

$$P_{vgs} = \frac{e^{-E_{GL}(100100100)/T} + e^{-E_{GL}(010010010)/T} + e^{-E_{GL}(001001001)/T}}{Z_{GL}}, \quad (4.6)$$

with the partition function of the small gating lattice,  $Z_{GL}$ , defined as

$$Z_{GL} = \sum_s e^{-E_{GL}(s)/T}. \quad (4.7)$$

The summation runs over all  $2^9$  states.

To assess the effect of  $H_{bias}$  and  $T$  on the probability of a valid gating state we computed  $P_{vgs}$  for  $T \in (0.0, 3.0]$  and  $H_{bias} \in [0.0, 6.0]$ . The results are shown in Figure 4.9.

Near  $T = 0.0$ , the probability of a valid gating state equals 1.0 for  $H_{bias}$  not too close to 0.0 and 6.0. With increasing  $T$ , the probability drops to lower values. For  $H_{bias} \approx 3.0$  the probability remains close to 1.0 up to the critical intrinsic-noise value  $T_c \approx 1.4$ . The optimal value of  $H_{bias}$ , i.e., the value for which  $P_{vgs}$  is maximal, shifts towards more negative values with increasing  $T$ . This asymmetry of the critical curve (surface) is in agreement with the results of Kinzel and Schick (1981) and Akutsu and Akutsu (1990) discussed in Section 4.2.2. Figure 4.10 shows the intersection of the surface with the  $T = 0.6$  plane, i.e., well below the critical intrinsic-noise value. A clear optimum is visible at  $H_{bias} = 3.1$ . The optimum remains in the close vicinity of this value up to  $T_c$ . This result is consistent with the result of Akutsu and Akutsu (1990).

To exclude the effect of  $H_{bias}$ , we define the *reduced* gating-lattice partition function  $Z_{GL}^{red}$ . The reduced partition function differs from the standard partition function in that the summation is restricted to run over all states with exactly three open gates. There are 84 such states leading to 84 terms in the reduced partition function.

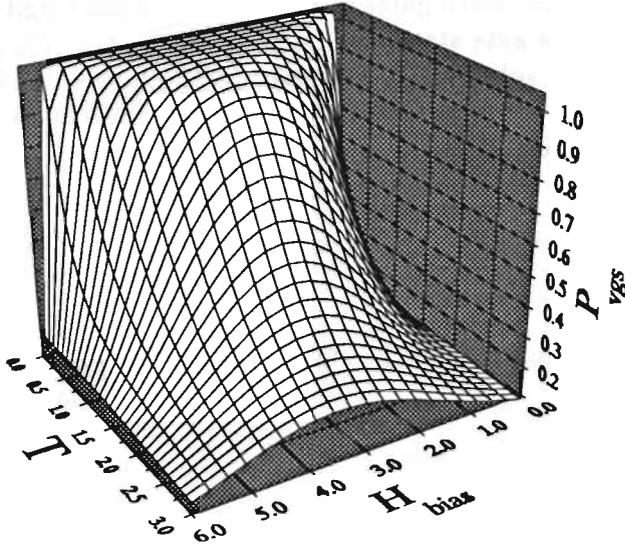


Figure 4.9: Probability of a valid gating state,  $P_{vgs}$ , as a function of  $T$  and  $H_{bias}$ .

Figure 4.11 shows the resulting surface plot. Clearly,  $P_{vgs}$  is independent of  $H_{bias}$  and varies with  $T$  only. Although,  $P_{vgs}$  starts decaying at  $T_c > 1.4$ , the slope is less steep than the slope at  $H_{bias} = 3.1$  in Figure 4.9. This difference is caused by the smaller number of terms (states) in the reduced partition function leading to a larger value of  $P_{vgs}$ , that is inversely proportional to  $Z$ .

The probability of the valid gating state with all gates on sublattice  $A$  open,  $P(A)$  is defined as

$$P(A) = \frac{e^{-E_{GL}(100100100)/T}}{Z_{GL}}. \quad (4.8)$$

To examine the probability  $P(A)$  as a function of the control signal  $H$  ( $= H_A = -H_B = -H_C$ ), we computed  $P(A)$  for  $H \in [0.0, 0.1]$  and  $T \in (0.0, 3.0]$ .

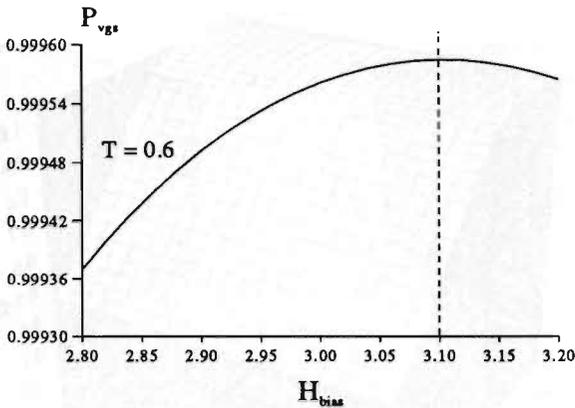


Figure 4.10: Probability of a valid gating state,  $P_{vg}$ , as a function of  $H_{bias}$  for fixed  $T = 0.6$ .

The results are shown in Figure 4.12.

As these Figures show, near  $T = 0.0$ ,  $P(A)$  increases rapidly with increasing  $H$ . However, the probability declines with increasing  $T$  at a rate inversely proportional to the strength of the control signal. Comparison of the results for the complete (Figure 4.12a) and reduced (Figure 4.12b) partition functions reveals a global pattern of results that is similar, except for the slightly larger values of  $P(A)$  for non-zero  $H$ . The results for larger values of  $H$  ( $H \in [0.0, 1.0]$ ) are shown in Figure 4.13.

For large control inputs, the probability  $P(A)$  remains equal to 1.0 up to a critical value  $T_c$  that increases with  $H$ . Again, the results obtained with the reduced partition function (Figure 4.13b) differ from those obtained with the complete partition function (Figure 4.13a) in that the plateau in the former extends to larger values of the intrinsic noise.

The exact results obtained with our nine-gate lattice provide insight into the behaviour of the gating lattice as a function of  $H_{bias}$ ,  $T$ , and  $H$ . In the following, we define two types of gate dynamics corresponding to the two types of partition functions analysed. With these gate dynamics we perform simu-

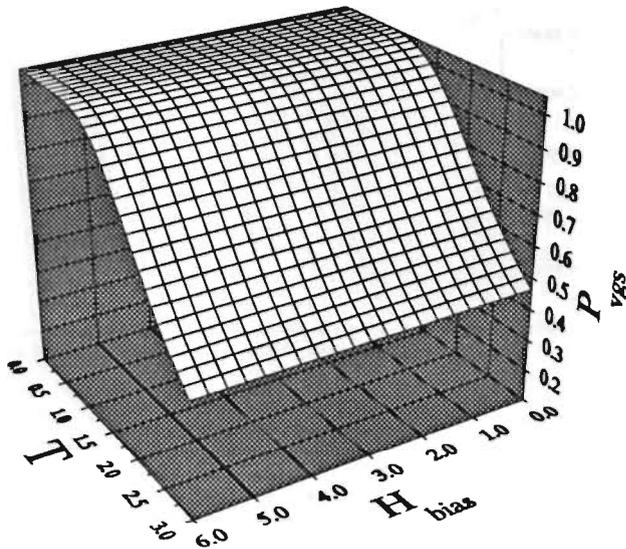


Figure 4.11: Probability of a valid gating state,  $P_{vgs}$ , computed with the reduced partition function as a function of  $T$  and  $H_{bias}$ .

lation studies of large gating lattices (i.e., larger than  $3N = 9$ ) and compare the results obtained with the exact results of the nine-gate lattice.

### 4.3.2 Gate Dynamics

In our first type of gate dynamics, the updating of individual gates follows Glauber's (1963) rule (cf. Section 3.5). Gate  $g_x(i, j)$  on sublattice  $x$  opens with probability

$$P(G_x(i, j) \rightarrow -G_x(i, j)) = \left[ 1 + \exp\left(-\frac{\Delta E_{GL}(G_x(i, j))}{T}\right) \right]^{-1}, \quad (4.9)$$

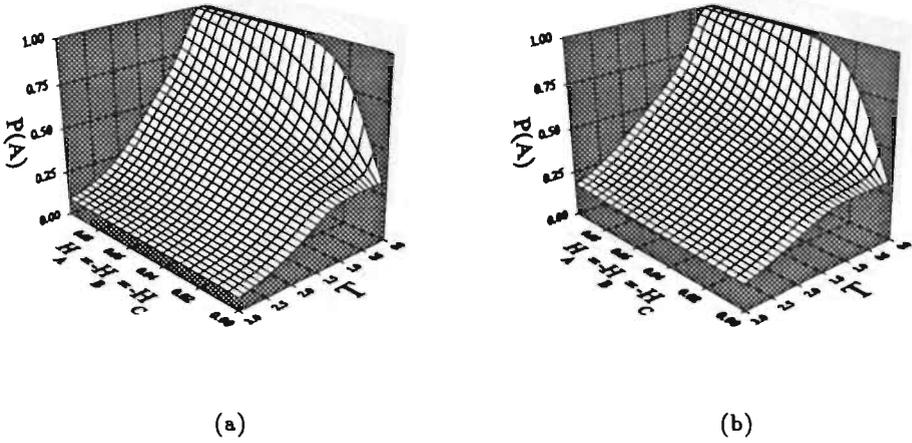


Figure 4.12: Probability of valid gating state  $A$ ,  $P(A)$  as a function of small control signals  $H$  and  $T$  for the complete partition function (a) and the reduced partition function (b).

with  $\Delta E_{GL}(G_x(i, j))$  the change in gating-lattice energy accompanying the transition from a closed to an open gate  $G_x(i, j)$ .  $\Delta E_{GL}(G_x(i, j))$  is defined analogously to (3.19), which leads to

$$\Delta E_{GL}(G_x(i, j)) = -2G_x(i, j) \left( H_{bias} - H_x - \sum_{k,l \in \mathcal{N}(i,j)} G_y(k, l) \right). \quad (4.10)$$

In our simulation studies, we randomly select lattice coordinates and apply (4.9) to the gate at these coordinates. Repeating this process simulates an asynchronous process of gate dynamics. In contrast to synchronous updating, in which all gates are updated simultaneously, asynchronous updating is a serial process that mimics unsynchronised parallel dynamics (e.g., Amit, 1989). Asynchronous updating avoids the problems associated with simultaneous updating (see, e.g., Lenting, 1992). The stochastic gate dynamics are in accordance with the characteristic of stochasticity as stated in Section 2.6.1.

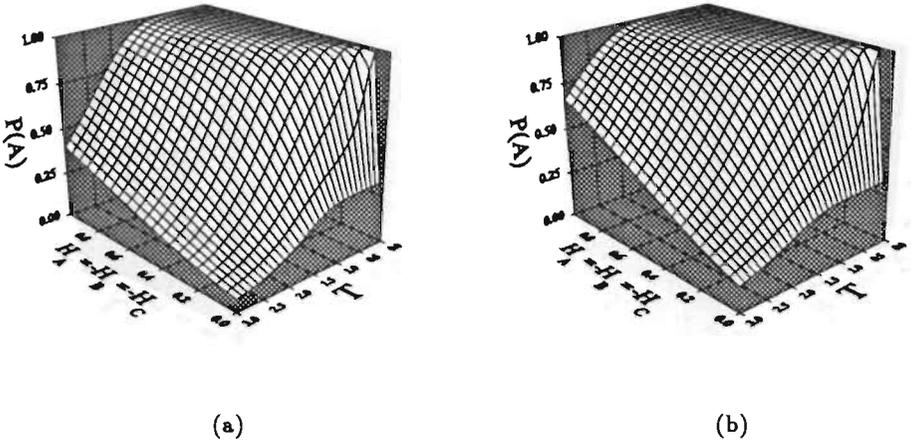


Figure 4.13: Probability of valid gating state  $A$ ,  $P(A)$  as a function of large control signals  $H$  and  $T$  for the complete partition function (a) and the reduced partition function (b).

### 4.3.3 Order-Parameter Conserving Gate Dynamics

A disadvantage of gate updating with (4.9) is that it depends on the absolute value of the control signals. When the control signals contain a positive bias that exceeds 6.0 or a negative bias that falls below 0.0, the competitive couplings become ineffective. Our second type of gate dynamics, based on Kawasaki (1972) and proposed as activity-conserving dynamics for neural networks in Postma, Van den Herik, and Hudson (1993b), does not suffer from this problem. Instead of changing the states of single gates, in Kawasaki updating, the states of neighbouring pairs of gates are exchanged. Whenever a nearest-neighbouring gate pair  $G_x(i, j)$ ,  $G_y(k, l)$  ( $x \neq y$  and  $g_y(k, l) \in N(i, j)$ ) is encountered, their states are exchanged with probability  $P(G_x(i, j) \leftrightarrow G_y(k, l))$  that is defined as:

$$P(G_x(i, j) \leftrightarrow G_y(k, l)) = \left[ 1 + \exp \left( - \frac{\Delta E_{GL}(G_x(i, j), G_y(k, l))}{T} \right) \right]^{-1}, \quad (4.11)$$

where  $\Delta E_{GL}(G_x(i, j), G_y(k, l))$  is the change in energy associated with the exchange of gate states:

$$\Delta E_{GL}(G_x(i, j), G_y(k, l)) = 2G_x(i, j) \times \left( H_x - H_y + \sum_{m, n \in N(i, j)} G_x(m, n) - \sum_{m, n \in N(k, l)} G_x(m, n) \right). \quad (4.12)$$

It should be noted that the  $H_{bias}$  term has vanished. Gate updating according to (4.11) is therefore independent of  $H_{bias}$  and ignores any bias present in the control signals  $H_x$ . When we employ these order-parameter conserving dynamics, we assume that the number of open gates on the gating lattice is equal to  $N$  (i.e., the number of gates on a sublattice). This means that lattice states, generated by (4.11), correspond to the states that are part of a reduced partition function.

The intrinsic-noise value is one of the main parameters of concern in the gate dynamics. In the following, we discuss how the choice of  $T$  affects the gating-lattice behaviour.

#### 4.3.4 Optimal Intrinsic-Noise Level

Our objective is to achieve a selective opening of all gates on one sublattice whenever the corresponding control signal is the largest of all control signals. The distributed competition in the lattice is to be exploited maximally by choosing the optimal intrinsic-noise level  $T_{opt}$ . If, on the one hand, we assume  $T_{opt} < T_c$ , the gating lattice ends up in an equilibrium state that is ordered, i.e., a valid gating state with  $m \approx 0.0$  or  $m \approx 1.0$ . If, on the other hand, we take,  $T_{opt} > T_c$ , the equilibrium state of the gating lattice is disordered with  $m \approx \frac{1}{3}$ . As we have seen in the previous chapter, two-dimensional neural lattices are flexible, i.e., relatively insensitive to (ordered) initial states. The intrinsic tendency to end up in an ordered state for  $T < T_c$  may therefore be exploited by selecting the  $T_{opt} < T_c$ .

### 4.3.5 Responsiveness to External Inputs

Although the gating lattice needs to be robust to external noise, at the same time it should remain responsive to (small) external inputs  $H_x$ . To gain insight into the responsiveness of the gating lattice, we discuss the behaviour of a gating lattice of infinite size. Then we treat the more realistic case of finite gating lattices.

In the statistical mechanics of Ising lattices, the tendency of spins to align with an externally applied field  $H$  is called the *susceptibility*  $\chi$  defined as:

$$\chi = \left( \frac{\partial m}{\partial H} \right)_T, \quad (4.13)$$

where  $m$  is the order parameter. Referring to Figure 3.4 in Chapter 3 (showing the equilibrium value of  $m$  as a function of  $H$ ), the susceptibility is associated with the slope of the  $m(H)$  curve. The susceptibility is also linked to the stochastic fluctuations of gate states. When the susceptibility is large, these fluctuations are highly correlated over large distances on the lattice. For a finite lattice, the finite-size susceptibility  $\chi_L$  follows from the fluctuation relation (see, e.g., Binder and Landau, 1984):

$$\chi_L = \frac{L^2}{T} (\langle G^2 \rangle - \langle G \rangle^2), \quad (4.14)$$

where  $L$  is the linear dimension of the gating lattice ( $L = \sqrt{3N}$ ) and  $G$  is the gate variable defined as  $G = G_A$  if part of sublattice A, and  $G = -G_B$  or  $G = -G_C$  if part of one of the other sublattices. With these expressions, the responsiveness of an infinitely large gating lattice to external inputs can be understood as follows (cf. Chandler, 1987).

- When  $T < T_c$ , an infinitesimal value of  $H$  breaks the symmetry and produces the opening of gates on sublattice A when positive and the opening of gates on sublattice B or C when negative. Given the discontinuity of the magnetization function  $m(H)$  at  $H = 0$  (cf. Figure 3.4), the susceptibility (derivative) at  $H = 0$  is divergent.
- At or near the critical point,  $T \approx T_c$ , the energetic barrier between the three symmetrical states disappears. Presenting the lattice with a small

external input does not suffice to break the symmetry. The fluctuations at the critical point are highly correlated, they extend over the entire lattice. In the infinite-size limit  $L \rightarrow \infty$ ,  $\chi_L$  becomes infinite too, since the correlation term in (4.14) ( $\langle G^2 \rangle - \langle G \rangle^2$ ) is non-zero.

- For  $T > T_c$ , stochastic fluctuations yield a zero correlation term, and consequently, the susceptibility vanishes.

The responsiveness of the gating lattice appears to be maximal for  $T < T_c$ . This observation is consistent with the optimal intrinsic-noise value mentioned earlier. The problem remains to establish the optimal value within this interval, since a too small value of  $T$  slows down the Glauber updating process. Before turning to simulation studies to determine the optimal intrinsic-noise value, we discuss the validity of statistical-mechanics results for gating lattices of finite size and limited relaxation times.

## 4.4 Finite Size and Limited Time

When applying statistical mechanics to the gating lattice, there are two things to be noted. First, the gating lattice is of finite size, whereas most theoretical results on Ising lattices become exact in the limit of infinite systems only. Second, the time allowed for the gating lattice to reach equilibrium is limited for practical reasons (see below), whereas the equilibrium behaviour of Ising lattices may only be obtained after much longer “observation” times.

This section discusses the effects of finite size and limited time on the behaviour of the gating lattice. The major implication of the limited size of the gating lattice, is that for all values of  $T > 0$  and  $H = 0$ , the equilibrium value of the order parameter  $m$  is equal to  $\frac{1}{3}$  (see, e.g., Binder and Heermann, 1988). The reason is that for finite systems at  $T < T_c$ , there is always a non-zero probability that the system will switch from one open sublattice to another. The average time elapsed for observing such a spontaneous switch is called the *ergodic time*. The ergodic time is an exponential function of  $L$ , the linear size of the gating lattice. In practice, spontaneous switches are only observed after sufficient time for very small lattices. Consequently, for finite (but not too small) gating lattices there *does* exist a spontaneous ordering for  $H = 0$ . As a case in point, real ferromagnets are finite but still exhibit spontaneous

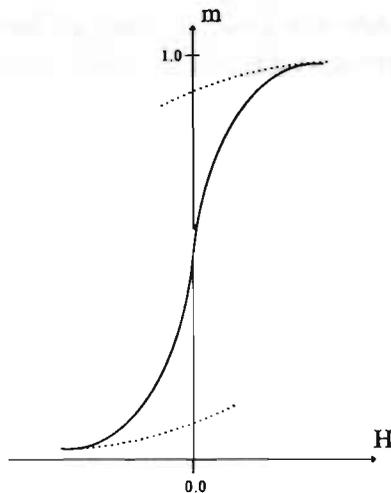


Figure 4.14: Order parameter  $m$  as a function of input  $H$  for finite-size gating lattices. Solid curve: true equilibrium. Dashed curve: value of  $m$  observed for limited time. (after Binder and Landau (1984)).

magnetization. Although the probability that a spontaneous reversal in the magnetization occurs is nonzero, one has to wait for a time that is greater than the present age of the universe (Lawrie, 1990). In finite systems, the tendency for spontaneous ordering can therefore still be exploited.

What happens for  $H \neq 0$  is of particular interest. Figure 4.14 (Binder and Landau, 1984) shows the true equilibrium value of  $m$  (solid curve), and the value of  $m$  obtained for limited time (dashed curve). Apparently, for limited observation time, the gating lattice may exhibit behaviour that is inconsistent with the input value  $H$ . For small positive values of  $H$ , the gating lattice may remain in an ordered state with  $m$  near 0.0. In the statistical mechanics of Ising lattices, such states are called *metastable states* (e.g., Binder and Stoll, 1973).

The existence of metastable states is of relevance for gating-lattice perfor-

mance, since it may prohibit a flexible response from changing inputs. Below, we perform simulation studies in which gating lattices are presented with small inputs  $H$  to study metastability and to assess its effect on the flexibility of gating. Moreover, we vary the amount of extrinsic noise to appraise the robustness of the gating lattice.

## 4.5 Simulation Studies

Our simulation studies aim at gaining insight into the behaviour of large, albeit finite, gating lattices. First, the appearance of metastable behaviour is examined and related to dynamic scaling and universal behaviour. Second, the robustness of gating lattices is studied and discussed.

### 4.5.1 Metastability in Gating Lattices

In this section, we discuss the simulations performed with the gating lattice to study metastability.

#### General Procedure

A triangular lattice geometry is obtained by defining the neighbourhood of a gate  $g(i, j)$  on the square lattice as

$$N_{g_l}(i, j) = \left\{ \begin{array}{l} ( [L + i - 1] \bmod L, [L + j + 1] \bmod L ), \\ ( i, [L + j + 1] \bmod L ), \\ ( [L + i - 1] \bmod L, j ), \\ ( [L + i + 1] \bmod L, j ), \\ ( i, [L + j - 1] \bmod L ), \\ ( [L + i + 1] \bmod L, [L + j - 1] \bmod L ) \end{array} \right\}. \quad (4.15)$$

The resulting lattice is a triangular lattice skewed by  $60^\circ$  (see Figure 4.15; cf. Watson and Ahumada, 1989) with periodic boundary conditions. The triangular lattices contain  $3N = L^2 = 333^2$  gates. The lattice is thus dimensioned to

allow for a continuation of sublattice structure (i.e.,  $3N$  is divisible by 3). An *iteration* is defined as  $3N$  updates, according to (4.9), applied to a randomly selected gate. A simulation run consists of 2500 iterations.

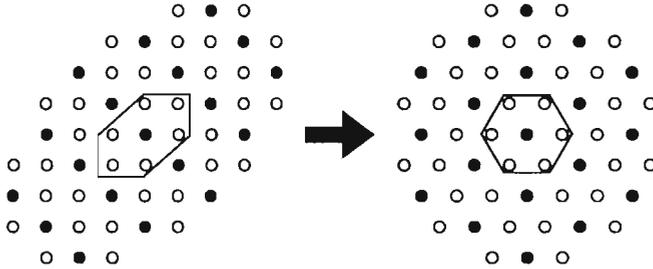


Figure 4.15: *Skewing a square lattice (left) by  $60^\circ$  yields a triangular lattice (right). The filled circles form a sublattice. The hexagon represents the neighbourhood of the filled circle in the center.*

Simulations proceeded as follows. The gating lattice is initialized in an ordered state with  $m(0) = 0.0$ . Control signals are set as follows  $H_A = -H_B = -H_C = 0.01$ . (This value leads to metastable behaviour as has been determined in a pilot simulation study.) Simulation runs are repeated for five intrinsic-noise values, i.e.,  $T = 1.28, 1.29, 1.30, 1.31$  and  $1.32$ . After ten iterations,  $m$  is recorded yielding a time-dependent curve  $m(t)$ .

## Results

The qualitative pattern of results is evident from the graph in Figure 4.16. It shows five  $m(t)$  curves, each one labelled by the value of  $T$ . The inset shows the graphs for the first 200 iterations. For the smallest value  $T = 1.28$ , a value of  $m(t) \approx 0.05$  is reached within a small time ( $\approx 10$  iterations) that remains stable over the remaining simulation time (the shaded interval). For  $T > 1.29$ , however, a small region of metastability is followed by a steady increase in  $m$ . At the largest intrinsic-noise value  $T = 1.32$ , metastability is (almost) absent, but the convergence magnetization is smaller than the convergence magnetizations for the smaller values of  $T$ .

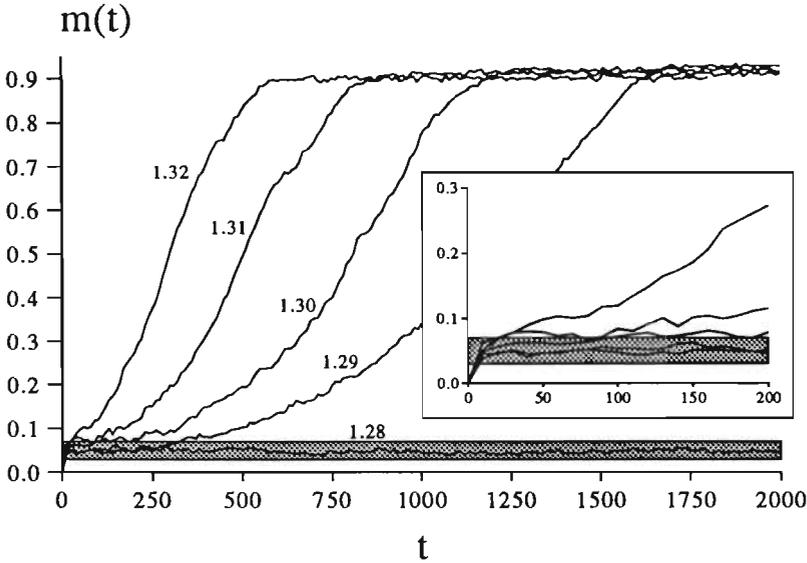


Figure 4.16: *Metastability in the gating lattice.*

## Discussion

The appearance of metastability, associated with small fields, may be interpreted as an intrinsic threshold of the gating lattice. The appearance of a sufficiently flat metastable region (the shaded interval in Figure 4.16) with small values of  $H$  indicates that the gating lattice is not able to respond within a fixed number of iterations to control signals that do not exceed  $H^*$ , the *threshold value* of  $H$ . It is clear from the simulation results that the region of metastability increases with decreasing values of  $T$

These observations can be formalised through *dynamic scaling* (e.g., Binder and Stoll, 1973). According to dynamic scaling, the non-linear relaxation functions obtained for different values of  $H$  near  $H^*$  and  $T$  near  $T_c$ , can be mapped onto each other by means of a rescaling of  $H$  and  $t$ . To allow for such mapping, the relaxation functions  $m(t)$  are first “normalized” through a simple rescaling of  $m(t)$ . The scaled order parameter  $m_*(t)$  is defined as:

$$m_s(t) = \frac{m(t)}{m_{conv}(t)}, \quad (4.16)$$

where  $m_{conv}(t)$  is the convergence value of  $m(t)$  (cf. Chapter 3).

According to dynamic scaling, the scaled relaxation curves  $m_s(t)$  obtained for different values of  $H$  and  $T$  can be mapped onto each other through a rescaling of  $H$  to  $H_s$ , as

$$H_s = H \left| \frac{T - T_c}{T_c} \right|^{\beta\delta}, \quad (4.17)$$

and a rescaling of  $t$  to  $t_s$ , as

$$t_s = t \left| \frac{T - T_c}{T_c} \right|^{\Delta_{\delta_m\delta_m}} \quad (4.18)$$

The exponents  $\beta$ ,  $\delta$ , and  $\Delta_{\delta_m\delta_m}$  are called *critical exponents* that describe the behaviour of variables, such as the order parameter and susceptibility near  $T_c$ . The values of these exponents for the (two-dimensional) gating lattice are as follows:  $\beta = \frac{1}{9}$ ,  $\delta = 14$  (Yeomans, 1992), and  $\Delta_{\delta_m\delta_m} = 1.85$  (Binder, 1976). An important property of these exponents is that they do not change with the microscopic details of the lattice. They are closely linked to the dimensionality of the system (Stanley, 1971). This feature is an expression of the principle of universality.

Because both  $\beta$  and  $\delta$  are positive, the dynamic scaling relation (4.17) indicates that the intrinsic threshold  $H^*$  is minimal at  $T = T_c$ . Consequently, the accuracy for small inputs of the gating lattice is maximal at  $T_c$ . However, according to (4.18) the time needed to establish equilibrium is maximal at  $T_c$ , a phenomenon known as *critical slowing down* (e.g., Heermann, 1990). The trade-off between speed and accuracy determines the value of  $T_{opt} < T_c$ .

### 4.5.2 Extrinsic Noise in the Gating Lattice

Our second simulation study focusses on the effect of extrinsic-noise level on gating quality and speed. We compare the pattern of results obtained with

the two updating rules described in Section 4.3.2 and Section 4.3.3. Where possible, we relate our findings to the exact results of the nine-gate lattice (Section 4.3.1).

## Procedure

Simulations are performed with triangular gating lattices containing  $99^2$  gates. The control signals are set as follows:  $H = H_A = -H_B = -H_C = 0.01q$ , with  $q$  the number of nearest neighbours of a gate (i.e.,  $q = 6$ ). Initially, gating lattices are in the ordered state with all gates on sublattice C open and all others closed (i.e.,  $m(0) = 0.0$ ). Gates are updated with either (4.9) or with (4.11). The same convergence criterion is used as in the neural-lattice simulations. Simulation runs of 1000 iterations are performed for fifteen intrinsic-noise values, i.e.,  $T \in \{0.1, 0.2, \dots, 1.5\}$  for the simulations with Glauber dynamics and  $T \in \{0.2, 0.4, \dots, 3.0\}$  for the simulations with Kawasaki dynamics. Four extrinsic-noise values are studied, i.e.,  $\sigma \in \{0.00, 0.05, 0.10, 0.20\}$ . Furthermore, each simulation run is repeated ten times and results (i.e.,  $m(t_{conv})$  and  $t_{conv}$ ) are averaged to enhance statistical accuracy.

## Results

The results of the simulations with Glauber updating are displayed in Figure 4.17. It shows the convergence values of the order parameter,  $m(t_{conv})$  as a function of the intrinsic-noise value  $T$  for the four extrinsic-noise values  $\sigma$ . The maximum convergence value (i.e., the peak of the curve) decreases with increasing  $\sigma$ . (For  $\sigma = 0.20$ , the maximum convergence value is not discernible.) Left of the peaks where the maximum convergence value is obtained, performance deteriorates rapidly with decreasing  $T$ , in particular for  $\sigma = 0.0$  and  $\sigma = 0.05$ . Right of the peaks, the decrease in  $m(t_{conv})$  proceeds much slower with growing  $T$ . Table 4.1 lists the quantitative results of the simulation. The column labelled  $T$  contains the intrinsic-noise values, while the columns labelled by  $\sigma$ , contain the corresponding average values of  $m(t_{conv})$  (averaged over ten runs). The numbers between brackets are standard errors of the mean.

The convergence times associated with the data points in Figure 4.17, are shown in Figure 4.18 and listed in Table 4.2. A peak in the convergence time is

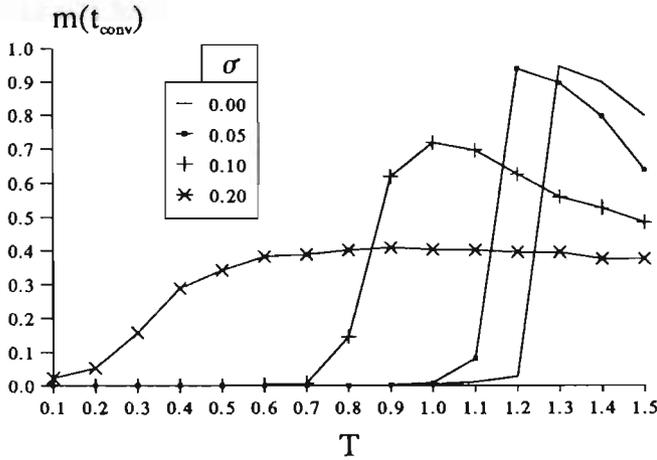


Figure 4.17: *Convergence values of the gating-lattice order parameter as a function of  $T$  for four values of  $\sigma$ . The results are obtained with Glauber updating.*

clearly visible for each value of  $\sigma$ . The values of  $T$  where the convergence time peaks are 1.3, 1.2, 0.9, and 0.4, for  $\sigma = 0.00, 0.05, 0.10$ , and  $0.20$ , respectively. It should be noted that the peak values correspond to the critical values of  $T$  (critical slowing down). For  $\sigma = 0.0$  and  $\sigma = 0.05$ , the intrinsic-noise values at which  $m(t_{conv})$  is maximal (see Figure 4.17) correspond to these critical values. For  $\sigma = 0.10$ , there is a slight discrepancy, i.e.,  $T = 0.9$  vs.  $1.0$ , while for  $\sigma = 0.20$ , the discrepancy may be large.

Figure 4.19 displays the convergence values of the order parameter obtained with Kawasaki updating. The quantitative results are listed in Table 4.3. The global pattern of results is similar to the pattern of results obtained with Glauber updating. The two main differences are that the convergence values with Kawasaki updating are smaller and that the four intrinsic-noise values associated with the maxima are larger.

The convergence times of the simulations with Kawasaki updating are shown in Figure 4.20 and listed in Table 4.4. The values of  $T$  where the convergence time

$T$	$\sigma = 0.00$	$\sigma = 0.05$	$\sigma = 0.10$	$\sigma = 0.20$
0.1	0.00 (0.0000)	0.00 (0.0000)	0.00 (0.0000)	0.02 (0.0010)
0.2	0.00 (0.0000)	0.00 (0.0000)	0.00 (0.0000)	0.05 (0.0028)
0.3	0.00 (0.0000)	0.00 (0.0000)	0.00 (0.0000)	0.16 (0.0063)
0.4	0.00 (0.0000)	0.00 (0.0000)	0.00 (0.0000)	0.29 (0.0142)
0.5	0.00 (0.0000)	0.00 (0.0000)	0.00 (0.0001)	0.34 (0.0050)
0.6	0.00 (0.0000)	0.00 (0.0000)	0.00 (0.0004)	0.38 (0.0089)
0.7	0.00 (0.0000)	0.00 (0.0000)	0.01 (0.0024)	0.39 (0.0107)
0.8	0.00 (0.0000)	0.00 (0.0000)	0.15 (0.0447)	0.40 (0.0072)
0.9	0.00 (0.0001)	0.00 (0.0002)	0.62 (0.0247)	0.41 (0.0055)
1.0	0.01 (0.0001)	0.01 (0.0008)	0.72 (0.0161)	0.40 (0.0087)
1.1	0.01 (0.0003)	0.08 (0.0514)	0.70 (0.0138)	0.40 (0.0117)
1.2	0.03 (0.0013)	0.94 (0.0048)	0.63 (0.0142)	0.40 (0.0058)
1.3	0.95 (0.0006)	0.90 (0.0053)	0.56 (0.0123)	0.40 (0.0038)
1.4	0.90 (0.0020)	0.80 (0.0078)	0.53 (0.0081)	0.38 (0.0038)
1.5	0.80 (0.0076)	0.64 (0.0073)	0.49 (0.0054)	0.38 (0.0028)

Table 4.1: *Convergence values  $m(t_{conv})$  obtained with Glauber updating.*

peaks are 2.2, 2.0, 1.8, and 1.0, for  $\sigma = 0.00, 0.05, 0.10,$  and  $0.20$  respectively. Except for  $\sigma = 0.20$  these values correspond with maximal values of  $m(t_{conv})$ .

## Discussion

In all simulations the gating quality is affected by the presence of extrinsic noise. Comparison of the results obtained with Glauber dynamics with those obtained with Kawasaki dynamics indicates that the maxima of the latter occur at higher (critical) intrinsic-noise values. The critical values of  $T$  obtained with Kawasaki dynamics (2.2, 2.0, 1.8, and 1.0) differ from those obtained with Glauber dynamics (1.3, 1.2, 0.9, and 0.4). The increased values of  $T$  for Kawasaki dynamics in comparison to those for Glauber dynamics cause a deterioration of gating quality  $m(t_{conv})$ . The simulation results do not agree with the enhanced equilibrium probability for the reduced partition function discussed in Section 4.3.1. We suspect that the increased values of  $T$  obtained with Kawasaki dynamics result from the initial state. To verify this,

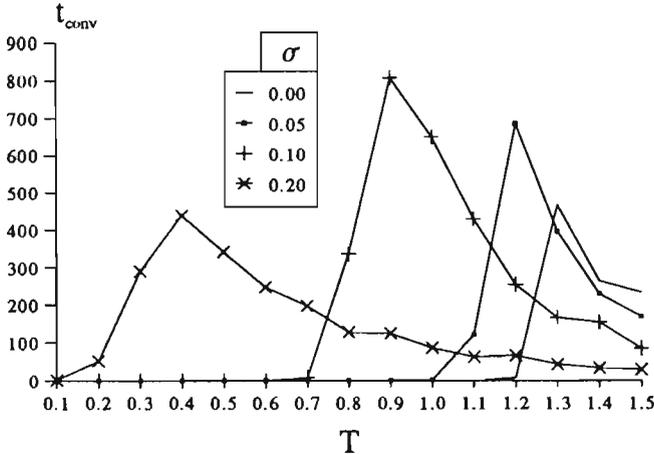


Figure 4.18: *Convergence times of the gating-lattice order parameter as a function of  $T$  for four values of  $\sigma$ . The results are obtained with Glauber updating.*

we ran simulations with Kawasaki dynamics starting from a disordered state  $m(0) = \frac{1}{3}$ . The results are shown in Figure 4.21. Evidently, for Kawasaki dynamics, ordered initial states hamper the gating quality considerably.

From these results we conclude that a gating lattice with Kawasaki dynamics is less flexible than one with Glauber dynamics. The choice for either of the two dynamics depends on the magnitude of control signals and the required gating quality (i.e.,  $m(t_{conv})$ ). When the bias in control signals is expected to be large, Kawasaki dynamics should be used despite the cost in gating quality. Otherwise, Glauber dynamics may be used.

Our simulations with Glauber dynamics show that the optimal performance is achieved at  $T$  just below  $T_c$ . When  $T$  is set much smaller than  $T_c$ , gating performance drops to zero because the lattice cannot recover from its initial ordered state. In this sense, the value of  $T$  ( $< T_c$ ) acts as a threshold. For  $T \leq 1.2$ , a noiseless control signal  $H = 0.01q$  cannot be detected within 1000 iterations. Interestingly, at  $T = 1.2$  gating performance improves from

$T$	$\sigma = 0.00$	$\sigma = 0.05$	$\sigma = 0.10$	$\sigma = 0.20$
0.1	0.0 ( 0.00)	0.0 ( 0.00)	0.0 ( 0.00)	1.6 ( 1.19)
0.2	0.0 ( 0.00)	0.0 ( 0.00)	0.0 ( 0.00)	51.5 ( 7.50)
0.3	0.0 ( 0.00)	0.0 ( 0.00)	0.0 ( 0.00)	289.9 (20.09)
0.4	0.0 ( 0.00)	0.0 ( 0.00)	0.0 ( 0.00)	439.7 (35.91)
0.5	0.0 ( 0.00)	0.0 ( 0.00)	0.0 ( 0.00)	342.0 (11.67)
0.6	0.0 ( 0.00)	0.0 ( 0.00)	0.0 ( 0.00)	248.2 (17.86)
0.7	0.0 ( 0.00)	0.0 ( 0.00)	8.1 ( 8.10)	198.2 (17.40)
0.8	0.0 ( 0.00)	0.0 ( 0.00)	336.2 (108.37)	127.9 ( 5.89)
0.9	0.0 ( 0.00)	0.0 ( 0.00)	807.5 ( 35.12)	124.7 (14.51)
1.0	0.0 ( 0.00)	2.6 ( 2.60)	649.8 ( 42.74)	85.9 (11.36)
1.1	0.0 ( 0.00)	121.2 (87.90)	429.7 ( 21.45)	62.7 ( 7.48)
1.2	7.8 ( 4.11)	684.9 (25.50)	255.5 ( 18.59)	66.3 ( 6.68)
1.3	467.5 (10.49)	396.7 (22.39)	166.9 ( 10.77)	43.2 ( 4.30)
1.4	264.8 (11.63)	230.2 (11.82)	155.0 ( 8.92)	33.3 ( 5.17)
1.5	233.8 (18.29)	169.6 (10.51)	84.5 ( 10.01)	29.0 ( 3.56)

Table 4.2: *Convergence times  $t_{conv}$  obtained with Glauber updating.*

$m(t_{conv}) \approx 0.0$  for  $\sigma = 0.0$  to  $m(t_{conv}) \approx 1.0$  for  $\sigma = 0.05$ . From these results we conclude that for optimal gating performance,  $T$  should be set to  $T_c - \epsilon$ , where  $\epsilon$  is a small positive value.

## 4.6 General Discussion

The performance of the gating lattice in the presence of extrinsic noise may be related to studies on random-field Ising models (RFIM, Imry and Ma, 1975). These studies concentrate on the behaviour of Ising lattices when confronted with Gaussian random fields that are centred around zero. It has been a matter of controversy whether the *lower critical dimension*  $d_c$ , i.e., the dimension below which spontaneous magnetization cannot occur, equals 2 or 3. Monte-Carlo simulation studies, performed by Andelman, Orland, and Wijewardhana (1984) yield zero spontaneous magnetization for a two-dimensional RFIM and non-zero spontaneous magnetization for the three-dimensional RFIM. They therefore concluded that  $d_c = 2$ . Their conclusion is in agreement with the

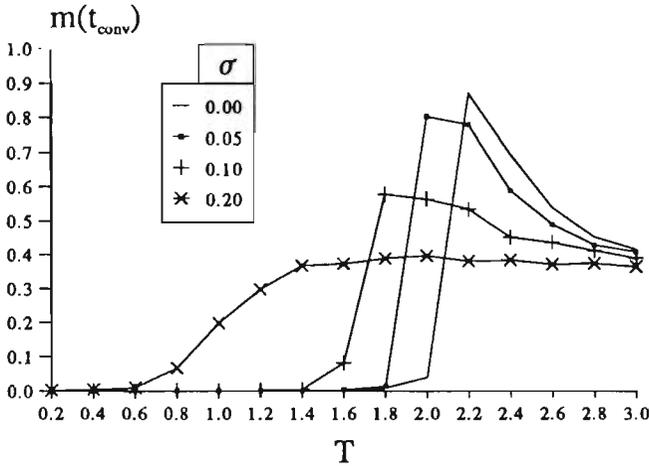


Figure 4.19: Convergence values of the gating-lattice order parameter as a function of  $T$  for four values of  $\sigma$ . The results are obtained with Kawasaki updating.

conjecture of Imry and Ma (1975) who introduced the RFIM. It is now generally accepted that  $d_c = 2$  (Plischke and Bergersen, 1989). With increasing  $\sigma$ , Andelman *et al.* find decreasing values of  $T_c$ . This is consistent with our simulations with neural and gating lattices in which we also find such a decrement of  $T_c$ .

The extrinsic-noise value above which there is no magnetization for finite  $T$ ,  $\sigma_c$ , may be taken as an indicator of robustness. The absence of magnetization for the two-dimensional RFIM implies that for the two-dimensional lattice,  $\sigma_c = 0.0$ . For the three-dimensional RFIM, Andelman *et al.* report  $\sigma_c \approx 2.75J$  while  $\sigma_c \approx 4.79J$  for the  $N$ -dimensional (mean-field) RFIM (Aharony, 1978).

$T$	$\sigma = 0.00$	$\sigma = 0.05$	$\sigma = 0.10$	$\sigma = 0.20$
0.2	0.00 (0.0000)	0.00 (0.0000)	0.00 (0.0000)	0.00 (0.0002)
0.4	0.00 (0.0000)	0.00 (0.0000)	0.00 (0.0000)	0.00 (0.0005)
0.6	0.00 (0.0000)	0.00 (0.0000)	0.00 (0.0000)	0.01 (0.0006)
0.8	0.00 (0.0000)	0.00 (0.0000)	0.00 (0.0000)	0.07 (0.0043)
1.0	0.00 (0.0000)	0.00 (0.0000)	0.00 (0.0000)	0.20 (0.0082)
1.2	0.00 (0.0000)	0.00 (0.0001)	0.00 (0.0001)	0.30 (0.0113)
1.4	0.00 (0.0000)	0.00 (0.0001)	0.01 (0.0005)	0.37 (0.0083)
1.6	0.00 (0.0001)	0.00 (0.0002)	0.08 (0.0324)	0.37 (0.0083)
1.8	0.01 (0.0004)	0.01 (0.0010)	0.58 (0.0276)	0.39 (0.0036)
2.0	0.04 (0.0043)	0.80 (0.0128)	0.56 (0.0134)	0.40 (0.0031)
2.2	0.87 (0.0044)	0.78 (0.0091)	0.53 (0.0065)	0.38 (0.0044)
2.4	0.69 (0.0068)	0.59 (0.0087)	0.45 (0.0049)	0.38 (0.0046)
2.6	0.54 (0.0031)	0.49 (0.0060)	0.44 (0.0039)	0.37 (0.0027)
2.8	0.45 (0.0041)	0.43 (0.0039)	0.41 (0.0039)	0.37 (0.0049)
3.0	0.42 (0.0030)	0.41 (0.0027)	0.39 (0.0018)	0.37 (0.0027)

Table 4.3: *Convergence values  $m(t_{conv})$  obtained with Kawasaki updating.*

## 4.7 Extensions of the Gating Lattice

We have defined the gating lattice as a two-dimensional triangular lattice and were therefore able to exploit its equivalence with the antiferromagnetic triangular Ising lattice. Instead, the gating lattice may be defined on a three-dimensional lattice to enhance its robustness. The equivalent Ising lattice, the antiferromagnetic cubic Ising lattice, has two degenerate states (i.e.,  $n = 2$ ). Consequently, selection is limited to one out of two patterns.

### 4.7.1 Potts Gating Lattice

Alternatively, the gating lattice may be defined as a lattice of *Potts gates*. Potts gates have  $n$  states instead of two. They are defined as Potts spins (Baxter, 1982). The energy function of a two-dimensional gating lattice with Potts gates is

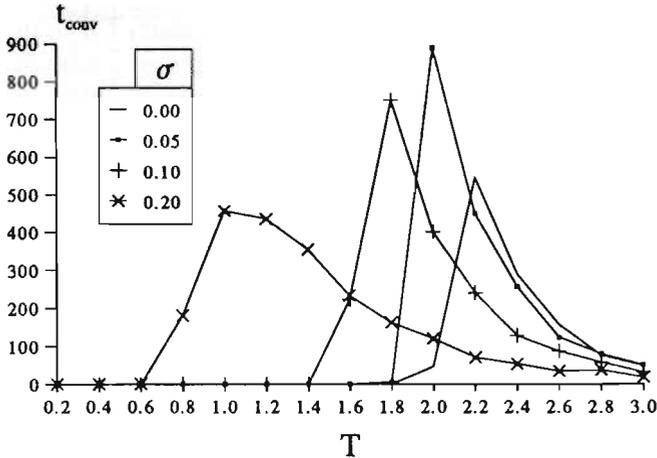


Figure 4.20: Convergence times of the gating-lattice order parameter as a function of  $T$  for four values of  $\sigma$ . The results are obtained with Kawasaki updating.

$$E_{GL}^{Potts}(\mathbf{G}) = - \sum_{i,j} \sum_{k,l \in N(i,j)} \delta(G_n(i,j), G_n(k,l)) - \sum_{i,j} H_n(i,j) G_n(i,j). \quad (4.19)$$

In this equation,  $G_n(i, j)$  is the  $n$ -ary gate state of the Potts gate at the lattice coordinates  $(i, j)$ , i.e.,  $G_n(i, j) \in \{1, 2, \dots, n\}$ . The nearest-neighbour coupling function  $\delta(\cdot)$  is defined as follows:  $\delta(G_n(i, j), G_n(k, l)) = 1$  if  $G_n(i, j) = G_n(k, l)$  and 0 otherwise. The Potts input  $H_n(i, j)$  is assumed to be a vector of  $n$  input values. The  $m$ -th input value couples to the  $m$ -th state of the Potts gate ( $m \in \{1, 2, \dots, n\}$ ). When all inputs are zero, the Potts gating lattice has an  $n$ -fold ground-state degeneracy. Each of these symmetrical ground states is defined as  $G_n(i, j) = m$ , for all coordinates  $(i, j)$ , with  $m \in \{1, 2, \dots, n\}$ . The number of ground states corresponds to the number of patterns selectable for gating (= the number of sublattices). The advantage of Potts gates is that the number of sublattices follows directly from  $n$  and is independent of the dimensionality of the lattice. Consequently, a three-dimensional Potts gating

$T$	$\sigma = 0.00$	$\sigma = 0.05$	$\sigma = 0.10$	$\sigma = 0.20$
0.2	0.0 ( 0.00)	0.0 ( 0.00)	0.0 ( 0.00)	0.0 ( 0.00)
0.4	0.0 ( 0.00)	0.0 ( 0.00)	0.0 ( 0.00)	0.0 ( 0.00)
0.6	0.0 ( 0.00)	0.0 ( 0.00)	0.0 ( 0.00)	1.1 ( 1.10)
0.8	0.0 ( 0.00)	0.0 ( 0.00)	0.0 ( 0.00)	182.3 (13.94)
1.0	0.0 ( 0.00)	0.0 ( 0.00)	0.0 ( 0.00)	457.7 (17.46)
1.2	0.0 ( 0.00)	0.0 ( 0.00)	0.0 ( 0.00)	435.3 (24.13)
1.4	0.0 ( 0.00)	0.0 ( 0.00)	0.0 ( 0.00)	354.1 (27.55)
1.6	0.0 ( 0.00)	0.0 ( 0.00)	224.0 (96.70)	234.3 (15.10)
1.8	0.0 ( 0.00)	5.6 ( 3.73)	750.5 (40.85)	163.4 ( 7.30)
2.0	47.2 (11.41)	889.3 (10.70)	400.0 (20.10)	121.2 ( 8.57)
2.2	545.2 (22.02)	448.8 (26.35)	240.5 (15.06)	71.1 ( 9.22)
2.4	289.3 (13.23)	256.0 (15.20)	128.4 ( 8.68)	54.1 ( 4.46)
2.6	156.8 (13.29)	123.6 (11.41)	87.0 ( 6.65)	34.6 ( 2.56)
2.8	76.5 ( 5.46)	79.9 (10.58)	58.0 ( 8.47)	37.8 ( 7.65)
3.0	48.9 ( 7.98)	51.4 ( 7.06)	32.0 ( 4.35)	18.5 ( 1.11)

Table 4.4: Convergence times  $t_{conv}$  obtained with Kawasaki updating.

lattice can be defined with an arbitrary number of sublattices to enhance the robustness. (As argued by Berker (1984) for the Potts gating lattice the lower critical dimension  $d_c = 2$  as for the regular lattice.) Our triangular gating lattice belongs to the *universality class* of three-state Potts models (Yeomans, 1992). This means that the critical exponents of the gating lattice correspond to those of the three-state Potts model.

#### 4.7.2 Hierarchical Gating Lattices

Another extension of gating lattices leads to *hierarchical gating lattices*. Hierarchical gating lattices are based on the real-space renormalization-group approach (Maris and Kadanoff, 1978; Wilson, 1979) and form a hierarchy of coupled gating lattices (see Postma, 1993). The idea is best illustrated by means of a two-level hierarchy based on the small gating lattice in Figure 4.8. Suppose that the nine-gate lattice forms the bottom level and that a three-gate lattice forms the top level of the hierarchy. Each of the three gates  $A$ ,  $B$ ,

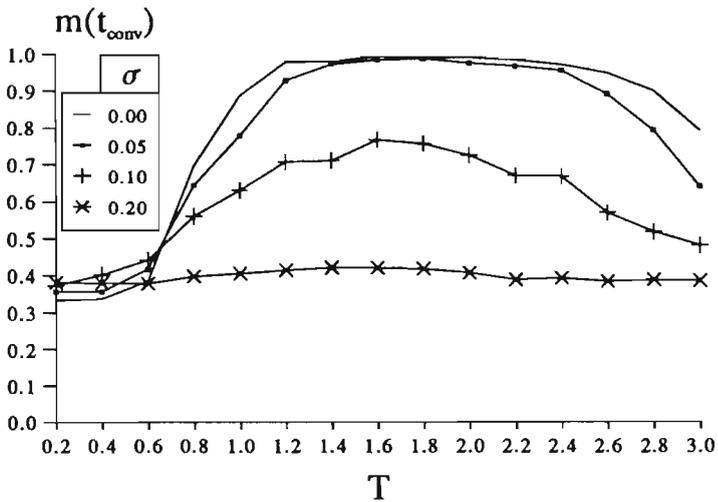


Figure 4.21: Convergence values  $m(t_{conv})$  as a function of  $T$  for Kawasaki dynamics with disordered initial states.

and  $C$  in the top lattice is connected with three gates in the bottom lattice. Gate  $x$  ( $\in \{A, B, C\}$ ) is connected with the gates labelled  $x_1$ ,  $x_2$ , and  $x_3$  (see Figure 4.8). The resulting structure is closely related to the renormalization procedure followed by Schick *et al.* (1977). Instead of connections, they employ a *projection operator* that maps the states of the three gates  $x_1$ ,  $x_2$ , and  $x_3$  onto gate  $x$ . (The projection operator is often defined by the majority rule: if the majority of the gates  $x_1$ ,  $x_2$  and  $x_3$  is open (closed), the *renormalized* gate  $x$  is open (closed) too.) In the renormalization-group procedure of Schick *et al.* (1977), the partition functions of the bottom and top lattice are identical except for (among others) the parameter  $K = J/T$ . This parameter of the nine-gate lattice renormalizes to  $K'$  in the three-gate lattice. Equating both partition functions yields a *recurrence relation*  $K \rightarrow K'$  with a fixed point that is defined as  $K^* = K = K'$ . Starting from  $K < K_c = J/T_c$  in the bottom lattice, the recurrence relation yields a renormalized value  $K' < K$ . Taking  $K'$  as the new value and inserting it into the recurrence relation leads to an even smaller value. Repeating this procedure leads to a  $K = 0$  fixed point which corresponds to complete disorder on the lattice (infinite intrinsic noise). When the initial value of  $K > K_c$ , repeated application of the recurrence re-

lation leads to the  $K = \infty$  fixed point which corresponds to the completely ordered state (zero intrinsic noise). Only when starting from an initial value  $K = K_c$ , do the recursion relations yield a non-trivial fixed point  $K^* = K_c$ . Returning to our hierarchical gating lattice, the procedure described above may be translated into our model by assuming the connections between both lattices to be reciprocal. The repeated application of the recurrence relation may then be replaced by the dynamics of the hierarchical gating lattice. When  $K$  is chosen properly, the hierarchical network converges to an ordered state. Hierarchical gating lattices may alleviate the need for a global control signal while retaining the property of robustness.

## 4.8 Conclusion

In this chapter we introduced the gating lattice as a building block for pattern routing. Two requirements for an effective system which will be capable of operating in the real world are robustness and flexibility (or responsiveness). To meet these requirements a number of variants have been studied.

First, instead of only using a central (extrinsic) gating controller, opening gates on a particular sublattice, we have added distributed (intrinsic) control in the form of a periodic continuation of locally interconnected WTA networks. Distributed competition facilitates the transition to an ordered state, enabling the selection of complete patterns despite noisy central control signals. This insensitivity to extrinsic noise has been defined here as robustness. Second, the responsiveness of the gating lattice, seen as sensitivity to small changes in the control signals, has been shown to be maximal just below the critical intrinsic-noise level. The requirement for flexibility set out in Chapter 3 is met by the gating lattice. This flexibility supports the continuous shifting of attention that reality requires.

The gating lattice provides an appropriate model for pattern routing, being capable of selecting one out of three input patterns. In the next chapter, the gating lattice is taken as the basic building block for pattern routing in the SCAN model of covert attention.



## Chapter 5

# The SCAN Model

In this Chapter we introduce the *Signal Channelling Attentional Network* (SCAN) model. SCAN achieves translation-invariant pattern processing through attentional selection. The largest part of SCAN consists of a gating network that takes a (potentially) large two-dimensional pattern, i.e., a retinal image, as its input. After selection of a contiguous subimage the gating network transports the attended pattern to the rest of SCAN which involves a classifier neural network to identify the pattern selected. Of the five characteristics introduced in Chapter 2, the SCAN model now combines the characteristics of hierarchy and invariance with the other three (stochasticity, sparse connectivity, and profligacy) that have been implemented in the gating lattices.

One approach to translation-invariant pattern processing involves having a constant transport medium with subimages transformed to a canonical representation (e.g., Burkhardt, 1989; Fang and Häusler, 1990; Glünder, 1990). Such approaches place heavy demands upon the transformation function, which must be capable of generating a canonical representation of an object regardless of its position. If there is more than one object in the visual field, some selection is still required to keep objects apart. Our approach, is to maintain a constant subimage with a variable transport medium. In SCAN, the transport medium *selects* objects, and an explicit selection is not required.

The chapter is organized as follows. After a qualitative description of SCAN in Sections 5.1 and 5.2, a formal specification of the structure and the dynamics

is given in Section 5.3. In Section 5.4 SCAN's performance is assessed through simulation studies. Section 5.5 concludes by discussing the significance of the characteristics of hierarchy and invariance for SCAN. A critical evaluation of the model in the light of its biological and psychological relevance is the subject of Chapter 6.

## 5.1 An Artificial Searchlight

The SCAN model incorporates a hierarchically structured network of gating lattices called the gating network (Postma, Van den Herik, and Hudson, 1993a, 1994a). At the base level of the hierarchy *spatial parallelism* (Ullman, 1986) is employed: large numbers of small gating lattices operate simultaneously to cover the whole input image. Going up the hierarchy, the degree of parallelism diminishes and the size of the input area treated by each gating lattice becomes larger. The decreasing degree in parallelism finally results in a single pattern arriving at the output of the network. The selections made by the individual gating lattices lead to several connected pathways running from the base level upwards to subsequent levels. The competitive interactions in the gating lattices ensure that only the pathway associated with the largest input value reaches the output of the gating network.

Figure 5.1 shows a sketch of a gating network. The shaded lines represent pathways of open gates. Their thickness reflects the strength of the associated (control) input. Whenever two lines meet, going from the base level of the pyramid to the top, the thicker of the two disrupts continuation of the thinner one. Consequently, the thickest line associated with the strongest input forms a connected pathway from the base level up to the top level. The parallel with the metaphor of covert attention discussed in Chapter 2, leads us to call this pathway the *attentional searchlight* (Postma, Van den Herik, and Hudson, 1992b). In the gating network the position of the attentional searchlight is controlled by the distribution of input values, e.g., the intensity of individual pixels. In SCAN, each value of the input distribution is taken to be the result of the cross-correlation of a part of the input with a template to achieve an optimal integration with ART networks (Carpenter and Grossberg, 1987a, 1987b). For expository purposes, the slightly adapted version of ART and its integration with the gating network are specified in Section 5.3.

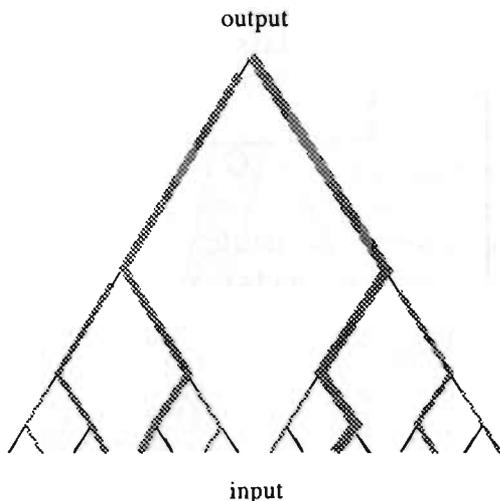


Figure 5.1: *Diminishing degree of parallelism in SCAN.*

## 5.2 The Structure and Operation of SCAN

A global outline of SCAN's structure is shown in Figure 5.2. The two components of SCAN are the gating network and the classifier network. The input of the gating network is the pattern  $I$ , of which a contiguous subimage, the *attended pattern*  $A$ , is selected and gated towards its output pattern  $O$ . The shaded region represents the attentional searchlight.

The classifier network takes pattern  $O$  as its input and maps it onto a binary vector that represents class membership. The classifier network is based on the ART2 neural network proposed by Carpenter and Grossberg (1987a, 1987b). The network generates an expectation pattern  $E$  that is compared in parallel with all contiguous subimages of  $I$ .

For convenience, we assume here that the patterns  $E$ ,  $A$ , and  $O$  possess a two-dimensional image-like structure. It should be emphasized that such a structure is by no means required. In general, the patterns only need to be organized consistently to allow for matching and identification (see Chapter 6, for a discussion of this issue).

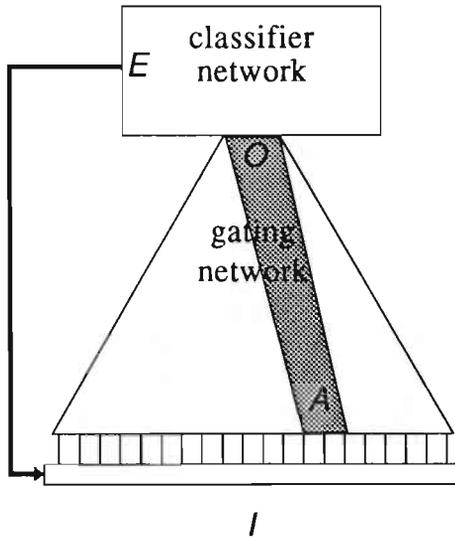


Figure 5.2: *The SCAN model.*

Our description of the SCAN model starts with a qualitative treatment of the gating and classifier networks. Following this, the model is specified formally.

### 5.2.1 The Gating Network

Describing the gating network is simplified by distinguishing between a data part and a control part. The data part of the gating network deals with the channelling of subimages; the control part is responsible for directing the searchlight by ensuring that each gating lattice receives the appropriate control signals. Below we first discuss the data part of the gating network. Then, we proceed by showing how the control part allows for autonomous selection of data.

### Data Part

Given a retinal image  $I$ , the task of SCAN is to select a contiguous subimage  $A$  and to channel it towards the output  $O$  of the gating network. In order to maintain the spatial structure of  $A$  during channelling, all elements (i.e., pixels) of  $A$  are processed by gating lattices. The two-level gating network shown in Figure 5.3 illustrates this. For clarity, all patterns and gating lattices are drawn as one-dimensional entities. The twelve hexagons at the bottom represent the elements of the input image  $I$ . The output pattern  $O$  (the hexagons at the top) is assumed to contain 4 elements. There are nine ways of selecting a contiguous subimage  $A(i)$  from  $I$ . Each of these selections corresponds to attending one of nine locations  $i$  ( $i \in \{1, 2, \dots, 9\}$ ). Whenever location  $i$  is attended, pattern  $A(i)$  ( $= \{i, i + 1, i + 2, i + 3\}$ ) is channelled in its entirety to the output  $O$  of the gating network. This is done by the gating lattices, represented in the figure by the horizontally aligned arrays of 12 circles (i.e., gates). Each of the three sublattices is defined as the set of every third gate, starting from the first, second, or third gate on the left. As illustrated in the magnified inset (the large dashed circle in Figure 5.3), only the input (*in*) and output (*out*) of a gate are involved in the data part of the gating network, not its control input (*h*) and control output (the gate state  $G$ ).

At the top level of the gating network, a single gating lattice can select one out of three patterns. These patterns are the outputs of the three gating lattices at the base level of the network. In combination, the four gating lattices can select and gate one out of nine patterns. Figure 5.4 displays two network states in which the attended pattern (shaded box) is channelled by an interconnected sequence of gating lattices. Figure 5.4a shows the selection of  $A(1)$  and Figure 5.4b shows the selection of  $A(5)$ . The paths along which the elements of the attended pattern are channelled are drawn as solid lines. Open gates are drawn as filled circles.

### Control Part

Proper operation of the gating network requires that the gating lattices are supplied with the appropriate control signals. In SCAN, the control signals are derived from the input image. The values of these control signals are a function of the content of the subimages to be attended to. Since, in our example, there are nine such subimages, the control part of SCAN receives

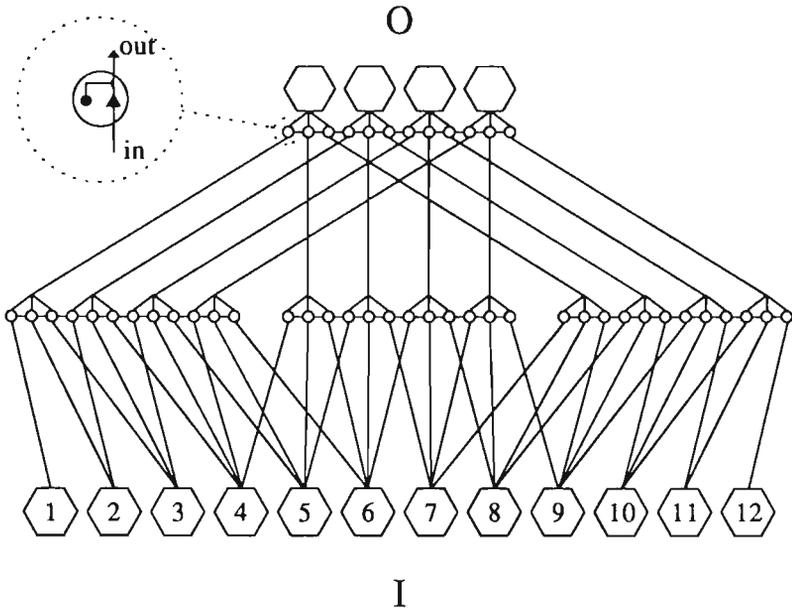
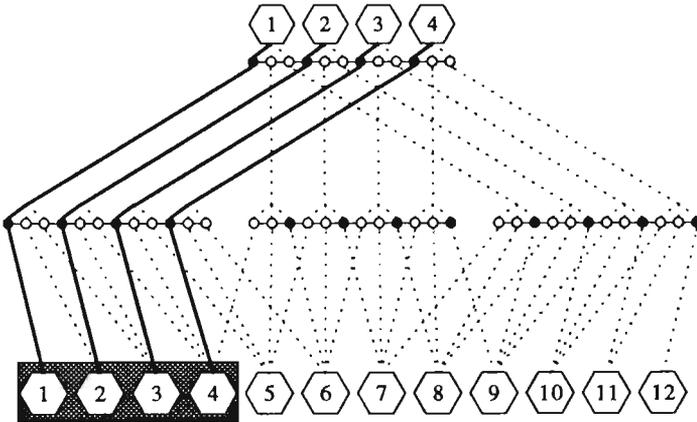


Figure 5.3: *The data part of SCAN.*

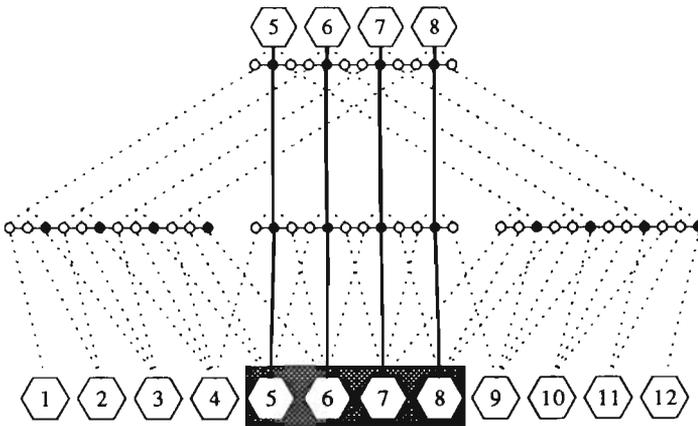
nine inputs. The magnitude of a control signal determines how strongly it attracts the attentional searchlight. The subimage with the largest control signal is the attended pattern A that is channelled to the output O.

To allow for the gating of control signals, the gating lattices are expanded with triplets of gates, denoted  $g_A^{tri}$ ,  $g_B^{tri}$ , and  $g_C^{tri}$ . Each of these gates “summarizes” the state of the sublattice specified by its subscript. When all gates on sublattice A are open,  $g_A$  is open too, and both  $g_B$  and  $g_C$  are closed. The connections between the gating lattice and its gate triplet (the box containing three circles) are illustrated in Figure 5.5.

Figure 5.6 shows the control part of SCAN. (The connections between gating lattices and their gate triplets are omitted for clarity of presentation.) The square boxes at the bottom represent the control signals  $v(i)$  ( $i \in \{1, 2, \dots, 9\}$ ) that feed into the gating lattices at the base level of the gating network. In SCAN, a control signal  $v(i)$  is defined as the Euclidian distance of the expect-



(a)



(b)

Figure 5.4: Two examples of pattern channelling in the data part of SCAN: (a) selection of  $A(1)$  and (b) selection of  $A(5)$ .

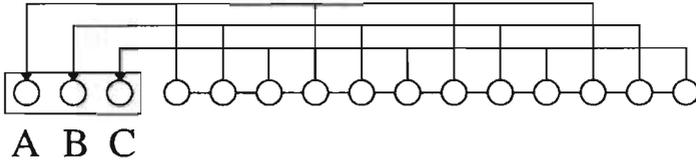


Figure 5.5: *The connections between the gating lattice and its gate triplet.*

tation pattern  $E$  with the to-be-attended subimage  $A(i)$ . As the magnified insets (large dashed circles) show, the control part of the gating network involves the control inputs ( $h$ ) of the gates in the gating lattices and the data inputs ( $in$ ) of the gates in the triplets. Control signals feed into the gates on one sublattice and also into the corresponding gate of the triplet. The output of the control part of the gating network is the largest control signal  $v_{max}$ .

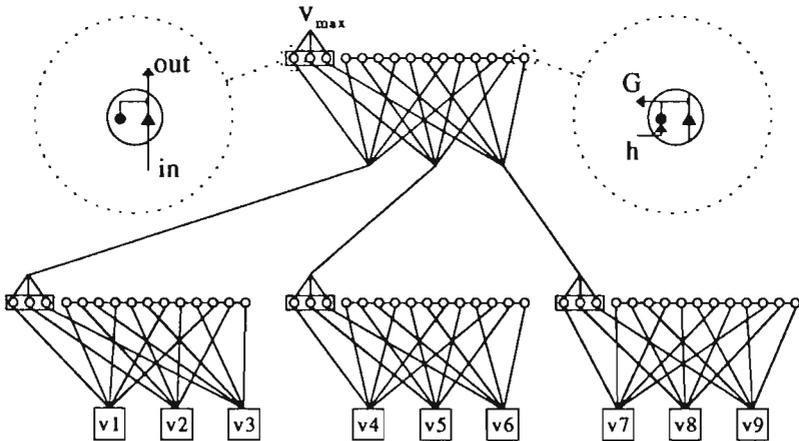


Figure 5.6: *The control part of SCAN.*

A gating lattice opens the sublattice that receives the largest of three control signals. Simultaneously, the corresponding gate of the triplet opens too, so that the control signal is gated to the subsequent level. This is to ensure that

the control signal associated with the selected subimage is channelled alongside the subimage.

Figure 5.7 illustrates the flow of control signals associated with the states shown in Figure 5.4. Figure 5.7a shows the selection of control signal  $v_1$  (shaded box) and its channelling towards the top of the gating network (solid lines). Figure 5.7b depicts the selection and channelling of control signal  $v_5$ .

### 5.2.2 The Classifier Network

The classifier network used in SCAN (see Figure 5.2) is a modified and simplified ART2 network (Carpenter and Grossberg, 1987b). The main reason for choosing this particular network is that it generates expectation patterns which, when matched with an input pattern, determine whether that input pattern will be stored into a new pattern class or recognized as a member of an existing pattern class. In SCAN, the expectation pattern  $E$  generated by the classifier network is matched in parallel at all locations  $i$  of the input image  $I$ . As a result, a control signal  $v(i)$  is obtained for each subimage  $A(i)$  selectable for gating.

Selecting ART as the classifier network underlines a possible top-down control of attentional selection (cf. Section 2.6.6). A bottom-up control of attentional selection, in which conspicuous image features deliver the control signals, is also possible but is not investigated here for two reasons. First, whether the control signals are derived from the matching with an expectation template or from image features, is of no significance for the workings of the gating network. Second, by focusing on top-down control we explicitly address the integration of the gating network with the ART network. Other choices for the classifier network are possible, (e.g., Rumelhart, Hinton and Williams, 1986; Murre, Phaf, and Wolters, 1992), but these require alternative ways of generating expectation patterns.

In the classifier network, discussed below, many computational shortcuts are taken. In particular, we assume simplified rules instead of the original dynamics. The classifier network consists of two layers of processing elements: the input/output layer  $F1$ , that is composed of an input part  $F1_{in}$  and an output part  $F1_{out}$ , and the category layer  $F2$ . Moreover, a threshold element  $T$  is added (see Figure 5.8).

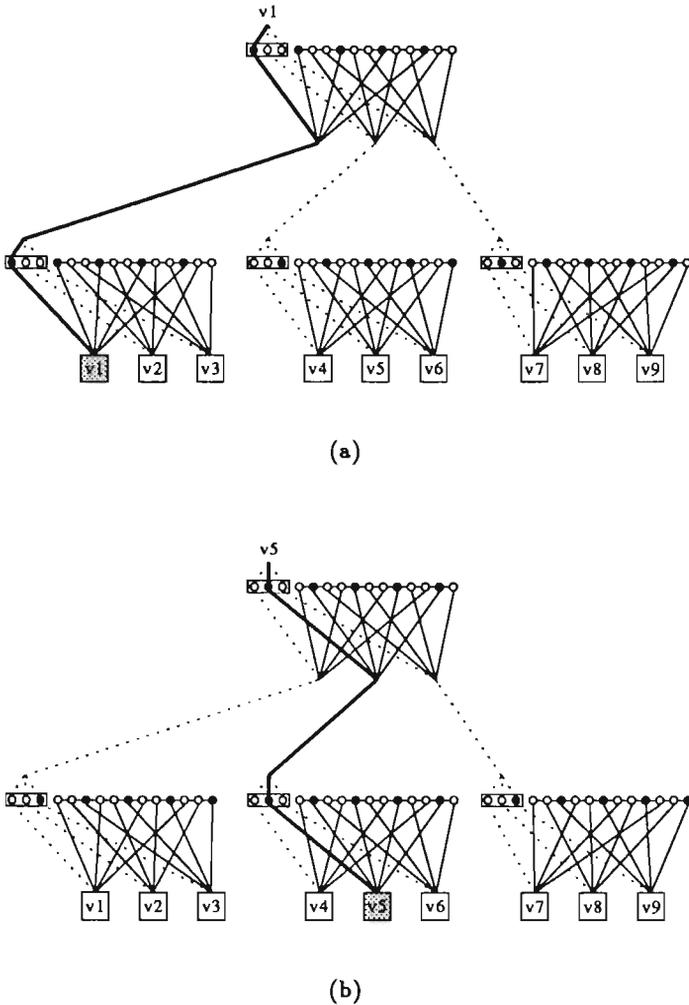


Figure 5.7: Two examples of signal channelling in the control part of SCAN: (a) selection of  $v(1)$  and (b) selection of  $v(5)$ .

The output pattern  $O$  of the gating network feeding into  $F1_n$  is transformed by the bottom-up connection matrix  $w^{bu}$  into an input pattern to  $F2$ . The dynamics of  $F2$  are of the winner-take-all type yielding a single active node representing class membership. An active  $F2$  element generates an expectation

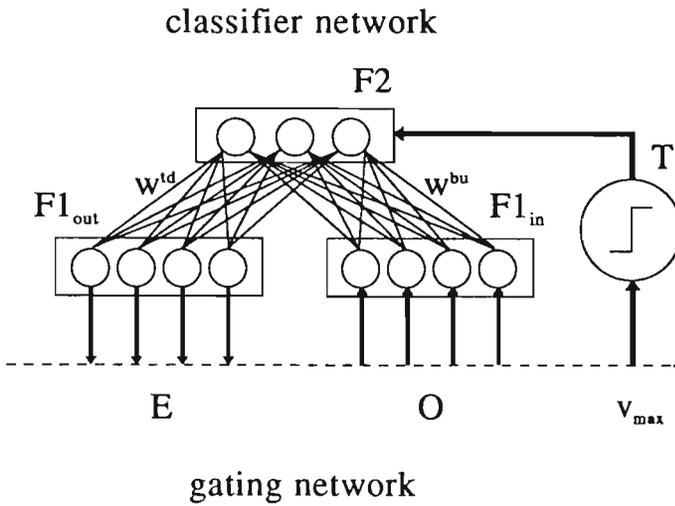


Figure 5.8: *The classifier network (based on Carpenter and Grossberg (1987b)).*

pattern  $E$  at  $F1_{out}$  via the top-down adaptive links  $w^{td}$ .

The threshold element  $T$  decides whether there is a mismatch between the expectation pattern and the input pattern. It takes the maximum control signal  $v_{max}$  as its input and compares it to an internal threshold value. In case of a mismatch, i.e., if the control signal is too small, the active (winning) category node is set to inactive. As a result, a new category node becomes active and generates a new expectation pattern  $E$ . The parallel matching of  $E$  with the subimages  $A(i)$  of  $I$  yields a new output pattern  $O$  and associated control signal  $v_{max}$ . The process continues until  $T$  detects a match (i.e.,  $v_{max}$  exceeds the threshold) and the active  $F2$  element is not suppressed. By changing the weights of the links connecting active elements, the patterns  $O$  and  $E$  are encoded in  $w^{bu}$  and  $w^{td}$ , respectively.



the gating network, the states of the gating lattices lead to the channelling of the best-matching subimage (shaded rectangle in the data part) towards the classifier input  $F1_{in}$ .

The output of the control part of the gating network feeds into the threshold element. The control signal is large enough to activate the threshold element  $T$  (solid circle). This element, in turn, keeps the  $F2$  element activated. Consequently, the patterns  $O$  and  $E$  are encoded in the bottom-up and top-down weights, respectively. In case the largest control signal is not large enough to activate  $T$ , the active  $F2$  node is suppressed, and another  $F2$  node becomes active. A new expectation is generated and compared to all subimages. The process repeats until a subimage is classified.

## 5.3 Formal Description of SCAN

In our formal specification of SCAN, we treat all patterns (i.e.,  $I$ ,  $A$ ,  $O$ , and  $E$ ) as one-dimensional vectors. Accordingly, a single index is employed to refer to the position of a gate on a gating lattice. The transformation to the two-dimensional case is discussed in Section 5.4.2.

### 5.3.1 Formal Description of the Gating Network

In a general gating network there are  $3^{L-1}$  gating lattices at level  $l$  ( $l \in \{1, 2, \dots, L\}$ ), with  $L$  the number of levels. Because each gating lattice processes three subimages, an  $L$ -level gating network can select one out of  $3^L$  subimages.

The width of the attentional beam, i.e., the number of elements in  $A$ ,  $O$ , and  $E$ , follows from the number of gates in a gating lattice. If there are  $3N$  gates in a gating lattice ( $N$  gates per sublattice), the attentional beam contains  $N$  "pixels".

The elements of input image  $I$  are denoted by  $I(i)$  ( $i \in \{1, 2, \dots, 3^L\}$ ,  $I(i) \in [0, 1]$ ). The subimages  $A(i)$  ( $i \in \{1, 2, \dots, 3^L - N + 1\}$ ) contain the  $N$  elements  $I(i), I(i+1), \dots, I(i+N-1)$ . The gates of a gating lattice at level  $l$  of the gating network are denoted by  $g_x(k, l, m)$ , where  $x \in \{A, B, C\}$  refers to the sublattice of which the gate is part,  $k \in \{1, 2, \dots, N\}$  specifies the location of

the gate on the gating lattice,  $l$  is the level, and  $m \in \{1, 2, \dots, 3^{l-1}\}$  is an index of the gating lattice.

### Data Part

The data input of gate  $g_x(k, l, m)$ , denoted  $in_x(k, l, m)$ , is defined as follows. If the gate is part of a gating lattice at the base level  $L$ , then

$$in_x(k, L, m) = I(x_{num} - 1 + k + 3[m - 1]), \quad (5.1)$$

where  $x_{num}$  is a numeric value associated with sublattice  $x \in \{A, B, C\}$  defined as

$$x_{num} = \begin{cases} 1 & \text{if } x = A \\ 2 & \text{if } x = B \\ 3 & \text{if } x = C \end{cases} \quad (5.2)$$

At levels  $1 \leq l < L$  of the gating network, the data input is

$$in_x(k, l, m) = \sum_{y \in \{A, B, C\}} out_y(k, l + 1, x_{num} + 3[m - 1]), \quad (5.3)$$

where  $out_x(k, l, m)$  is the data output of gate  $g_x(k, l, m)$  that is defined as

$$out_x(k, l, m) = \left[ 1 - \frac{1}{2} (G_x(k, l, m) + 1) \right] in_x(k, l, m). \quad (5.4)$$

$G_x(k, l, m) \in \{-1, +1\}$  is the gate state of  $g_x(k, l, m)$  (cf. equation (4.1)).

The elements  $O(i)$  ( $i \in \{1, 2, \dots, N\}$ ) of output pattern  $O$  are defined as

$$O(i) = \sum_{y \in \{A, B, C\}} out_y(i, 1, 1). \quad (5.5)$$

### Control Part

The inputs of the control part of the gating network are the control signals  $v(i)$ . They are defined as

$$v(i) = 1 - \frac{2}{N} \sum_{j=1}^N |I(i+j-1) - E(j)|. \quad (5.6)$$

Where the  $E(j)$  ( $E(j) \in [0, 1]$ ) are the elements of the expectation pattern  $E$ . It should be noted that  $v(i) \in [-1, 1]$ .

The control input of gate  $g_x(k, l, m)$ , denoted  $h_x(k, l, m)$ , are defined as follows. At the base level  $L$ , the control inputs are supplied by the control signals  $v(i)$  ( $i \in \{1, 2, \dots, 3^{L-1}\}$ ):

$$h_x(k, L, m) = v(x_{num} + 3[m-1]), \quad (5.7)$$

for all  $k \in \{1, 2, \dots, N\}$ . At levels  $1 \leq l < L$  the control inputs are defined as

$$h_x(k, l, m) = \sum_{y \in \{A, B, C\}} out_y^{tri}(l+1, x_{num} + 3[m-1]), \quad (5.8)$$

where  $out_x^{tri}(l, m)$  is the data output of gate  $g_x^{tri}(l, m)$ , i.e., the gate in the triplet that is associated with sublattice  $x$  of gating lattice  $m$  at level  $l$ .

The data input of a gate in a triplet,  $in_x^{tri}(l, m)$ , equals the control input of the gates on the sublattice it represents:

$$in_x^{tri}(l, m) = h_x(k, l, m). \quad (5.9)$$

The control input of a gate in a triplet,  $h_x^{tri}(l, m)$ , is provided by all gates on the sublattice it summarizes (cf. Figure 5.5):

$$h_x^{tri}(l, m) = \frac{1}{N} \sum_{k=1}^N G_x(k, l, m), \quad (5.10)$$

where  $G_x(k, l, m)$  is the gate state of gate  $g_x(k, l, m)$ . The gate state of a gate in a triplet equals its control input:

$$G_x^{tri}(l, m) = h_x^{tri}(l, m). \quad (5.11)$$

Finally, the data output of gate  $g_x^{tri}(l, m)$  is defined as

$$out_x^{tri}(l, m) = \left[ 1 - \frac{1}{2} (G_x^{tri}(l, m) + 1) \right] in_x^{tri}(l, m). \quad (5.12)$$

### 5.3.2 Formal Description of the Classifier Network

Although, in our simulation studies we do not study the dynamics of the classifier network, for completeness we present a formal description to demonstrate its integration with the gating network.

The classifier network has three layers,  $F1_{in}$ ,  $F1_{out}$ , and  $F2$ , whose processing elements (i.e., neurons) are denoted by  $p_{in}(i)$  ( $1 \leq i \leq N$ ),  $p_{out}(i)$  ( $1 \leq i \leq N$ ), and  $p_{class}(i)$  ( $1 \leq i \leq C$ ), respectively. The state of a processing element, called the *activation*, is defined for the three layers as follows. The activation of processing element  $p_{in}(i)$  of  $F1_{in}$ ,  $a_{in}(i)$  is defined as

$$a_{in}(i) = O(i). \quad (5.13)$$

The input of processing element  $p_{class}(i)$  of layer  $F2$  is

$$in_{class}(i) = T(i) \sum_{j=1}^N w^{bu}(i, j) a_{in}(j), \quad (5.14)$$

where  $T(i) \in \{0, 1\}$  is a variable whose value is set by the threshold element  $T$  during processing (see below) and  $w^{bu}(i, j)$  is the weight (or strength) of the connection from  $p_{in}(j)$  to  $p_{class}(i)$ . The activation of a processing element in  $F2$  is

$$a_{class}(i) = \begin{cases} \delta & \text{if } in_{class}(i) = \max_{i' \in \{1, \dots, C\}} \{in_{class}(i')\} \\ 0 & \text{otherwise} \end{cases} \quad (5.15)$$

with  $\delta$  a constant ( $\delta \in \langle 0, 1 \rangle$ ). Consequently, whenever an  $F2$  node receives an input that is the maximum of all  $F2$  inputs, it switches into an active state.

Finally, the activation of processing element  $p_{out}(i)$  of layer  $F1_{out}$  is defined as

$$a_{out}(i) = \sum_{j=1}^C w^{td}(i, j) a_{class}(j), \quad (5.16)$$

with  $w^{td}(i, j)$  the weight of the connection from  $p_{class}(j)$  to  $p_{out}(i)$ .

The state of the threshold element  $T$ , denoted  $a_T \in \{0, 1\}$  is

$$a_T = \begin{cases} 1 & \text{if } in_T > \theta \\ 0 & \text{otherwise} \end{cases} \quad (5.17)$$

The constant  $\theta \in \langle 0, 1 \rangle$  is a threshold that determines the granularity of classification. The input of the threshold element,  $in_T$  is defined as

$$in_T = v_{max} = \sum_{y \in \{A, B, C\}} out_y^{tri}(1, 1). \quad (5.18)$$

In case the largest control signal  $v_{max}$  does not exceed threshold  $\theta$ , the active  $F2$  node is suppressed for a fixed period of time. This is modelled by a threshold variable  $T(i)$  that is associated with each  $F2$  node. Initially,  $T(i) = 1$  for all  $i$ . If  $a_T = 0$  and  $a_{class}(i) = 1$ , then  $T(i)$  is set to 0.

The weights of the links connecting the three layers are assumed to be set instantaneously (see Simpson, 1990). We assume that the patterns  $O$  and  $E$  are *static*. Given that these patterns do not vary over time, the weights are set according to

$$w^{bu}(i, j) = a_{class}(j) a_{in}(i) \quad (5.19)$$

and

$$w^{td}(i, j) = a_{out}(i)a_{class}(j). \quad (5.20)$$

With these weight dynamics the patterns  $O$  and  $E$  are directly encoded in the weights. The reader is referred to Carpenter and Grossberg (1987a, 1987b) for more elaborate definitions of the dynamics.

## 5.4 Simulation Studies

The behaviour of the SCAN model as defined in the previous sections has been studied using simulations. The classifier network is not included in our simulations. Instead, we present the gating network with control signals  $v(i)$  directly. In Section 5.4.1 we examine the effect of varying the global intrinsic-noise level on the attentional orienting towards a predefined target. Then, in Section 5.4.2, SCAN's performance on an input distribution obtained by cross-correlating an object-template with a natural image is studied. Finally, in Section 5.4.3, the performance of SCAN with time-varying intrinsic noise is investigated.

### 5.4.1 Performance as a Function of Intrinsic Noise

As argued in the previous chapter, the value of the noise parameter  $T$  determines the speed and quality of gating in the gating lattices. Moreover, it has become clear that the optimal value of the noise parameter, depends on the magnitudes of the control signals that feed into the sublattices. In this simulation, we vary the global value of  $T$  and examine the effect on the attentional orienting of SCAN's beam towards a predefined target.

#### Procedure

Starting point for the simulations is a five-level ( $L = 5$ ) gating tree containing  $\sum_{l=1}^L 3^{l-1} = 121$  gating lattices. The gating lattices have periodic boundary conditions and contain  $3N = 99^2$  gates each. A single iteration is defined as the asynchronous Glauber updating of  $3N$  gates of all gates in the gating lattices. After each iteration, all gates in the triplets are updated according to (5.11). The control signals  $v(i)$  are set as follows. The *target*  $v_{max}$  is set to 0.2 and the remaining values are uniformly distributed over the interval  $[-0.1, 0.1]$ . These values are arbitrary, but fall in the range within which the competitive interactions of the gating lattice are effective.

At the end of each iteration, the SCAN order parameter  $m_{beam}(t, l)$  is recorded. For level  $l$  this order parameter is defined as the state of the gate in the triplet at level  $l$  that is part of the attentional beam. Initially, all gating lattices are

set to randomly selected winning states. For the lattices that are part of the beam the initial state does not correspond to the target state. Simulation runs of 1000 iterations are performed for four different intrinsic-noise values:  $T = 1.3, 1.4, 1.5,$  and  $1.6$ .

## Results

The graphs in Figure 5.10, show the development of  $m_{beam}(t, l)$  for the five levels of SCAN. In all graphs the curves from left to right represents  $m_{beam}(t, 5)$  to  $m_{beam}(t, 1)$ . For  $T = 1.3$ , shown in Figure 5.10a,  $m_{beam}(t, 5)$  and  $m_{beam}(t, 4)$  reach their maximum (convergence) value within 1000 iterations. Figure 5.10b shows the results for  $T = 1.4$ . For all levels convergence is established at  $t < 650$ . Larger intrinsic-noise values lead to even faster convergence, as can be seen from Figures 5.10c and d, that show the results for  $T = 1.5$  and  $T = 1.6$ , respectively. Associated with the increase of  $T$  is a decrease in the convergence value of  $m_{beam}(t, l)$ . For  $T = 1.3$ , the convergence values are approximately equal to 0.95, while for  $T = 1.6$ , the convergence values range from 0.75 ( $l = 1$ ) to 0.40 ( $l = 5$ ).

## Discussion

From the results of the first simulation, we conclude that SCAN is capable of orienting its attentional searchlight to a predefined target. The order-parameter response functions  $m_{beam}(t, l)$  associated with the five levels are shifted in time. In Figure 5.10a, for example, the response function  $m_{beam}(t, 4)$  (second curve from the left) starts to increase only after  $m_{beam}(t, 5)$  has converged. This behaviour illustrates the sequential nature of processing in SCAN. The channelling of patterns proceeds from the base to the top level.

The time  $t$  at which the order parameter reaches its maximum value decreases with increasing  $T$ . By varying the intrinsic-noise value, the response latency of the attentional searchlight can be controlled. However, with increasing  $T$ , the quality of gating, i.e., the convergence value of  $m_{beam}(t, l)$  decreases, as is evident from the small convergence values for  $T = 1.5$  and  $T = 1.6$  in Figure 5.10. A decreased gating quality at one level impairs the gating quality at all subsequent levels. The result is a strong deterioration of gating quality with decreasing  $l$  (see, e.g., Figure 5.10d). By keeping  $T$  small enough, the

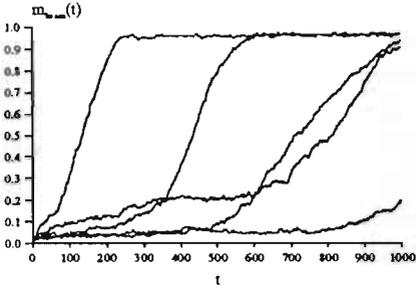
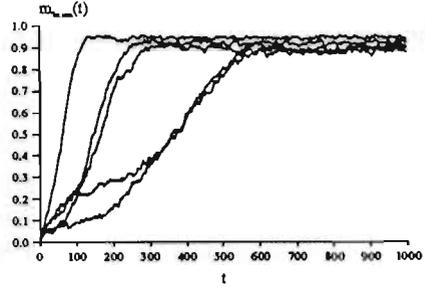
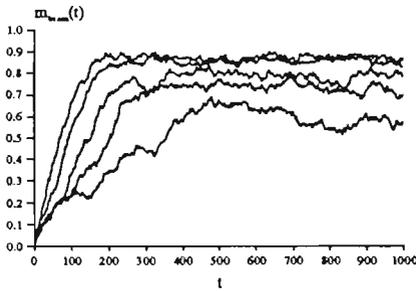
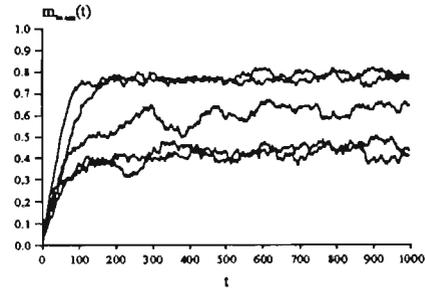
(a)  $T = 1.3$ (b)  $T = 1.4$ (c)  $T = 1.5$ (d)  $T = 1.6$ 

Figure 5.10: *The SCAN order parameter  $m_{beam}(t)$  as a function of  $t$  for four intrinsic-noise values: (a)  $T = 1.3$ , (b)  $T = 1.4$ , (c)  $T = 1.5$ , and (d)  $T = 1.6$ .*

deterioration is kept to a minimum. Alternatively, the value of  $T$  may be varied with  $t$  (see below).

### 5.4.2 Scanning a Natural Image

The first simulation was based on randomly distributed control signals. We now set out to perform a more realistic simulation in which the input image  $I$  is defined as a natural image. Control signals for each image position are derived by matching an expectation pattern  $E$  as generated by the classifier network to all image locations in parallel. The resulting distribution of control signals is expected to provide a better approximation to a realistic distribution than the one used in the previous simulation.

Applying SCAN to natural images requires the spatial structure of the gating network be specified in detail. A recursively-defined (fractal) structure, called the *triangular gating network*, accommodates the two-dimensional spatial sampling at the base level (Postma, Van den Herik, and Hudson, 1994b). It is defined as follows:

At any level  $l < L$  of the tree, a tri-furcation is defined as three branches of fixed length that emerge from a common centrepoint and have orientations that are rotated with respect to each other by 120 degrees, e.g., 0, 120, and 240 degrees. The end points of these branches serve as the centrepoints of the tri-furcations at level  $l + 1$ . At level  $l + 1$ , new tri-furcations emerge from the new centrepoints, after the reference frame has been rotated by 90 degrees and the length scale is reduced by a factor  $\sqrt{3}^{-1}$ .

Starting at a centrepoint at  $l = 1$ , recursive application of the above definition generates the triangular gating structure. Each tri-furcation represents a gating lattice. The end points of the tri-furcations at the base level  $L$  of the network form a triangular sampling grid. The triangular gating network is illustrated in Figure 5.11.

#### Procedure

An  $L = 10$  triangular gating tree is used to sample a grey-value input image. The input image used in the simulation is part of a hand-scanned ( $931 \times 931$  8-bit pixels) photograph of the cover of the Beatles album *Sgt. Pepper* taken from Evans (1984) (see the left part of Figure 5.12).

A straightforward mapping scheme is employed to map the input image onto a triangular grid resulting in an input image  $I$  of  $3^{10} = 59049$  pixels. The

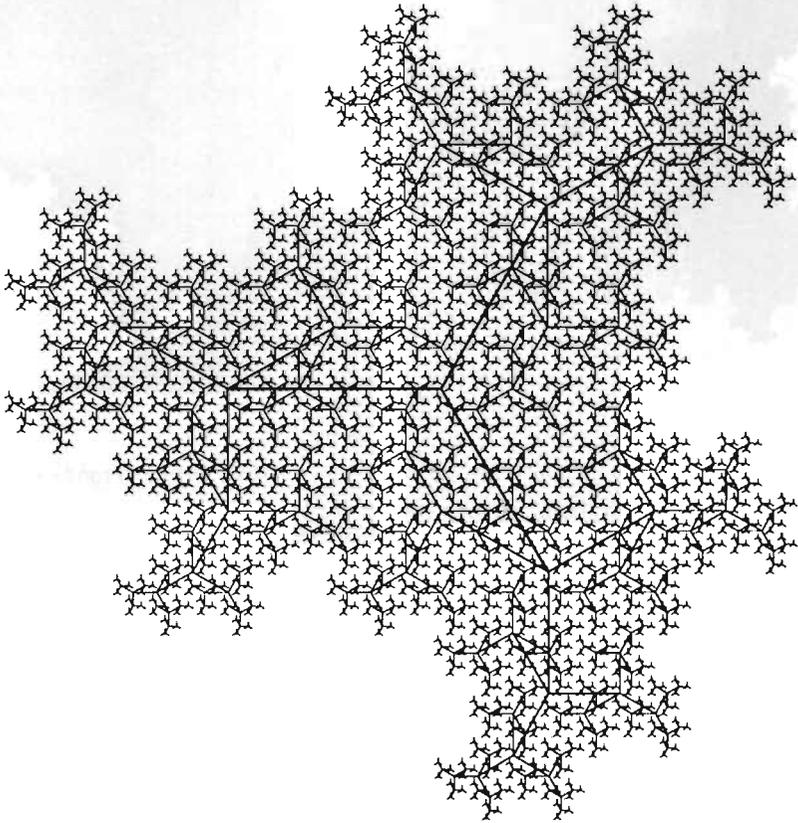


Figure 5.11: *The triangular gating network ( $L = 8$ ).*

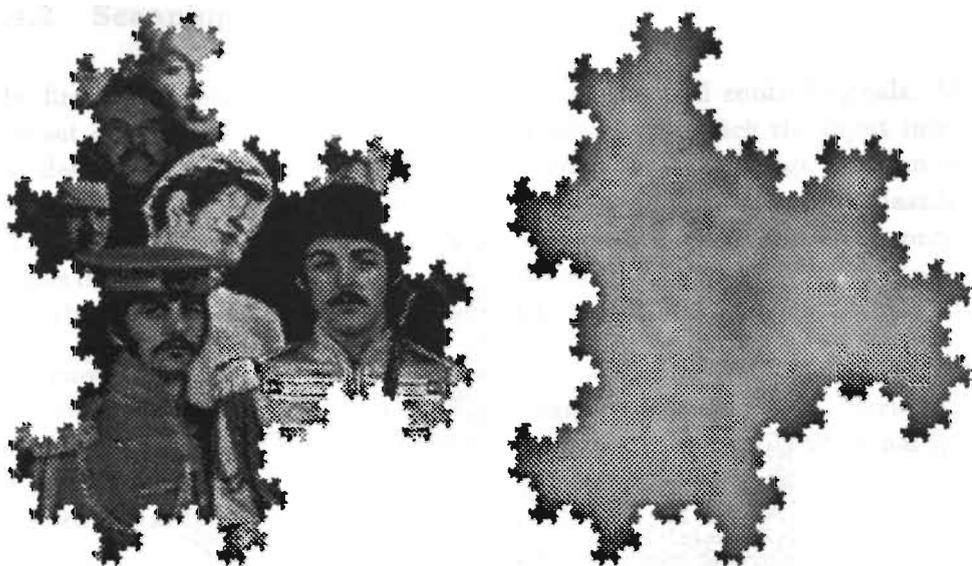


Figure 5.12: *The data input (left) and control input (right) used in the simulation.*

expectation pattern  $E$  is assumed to have an approximate square shape and contains part of Paul McCartney's face (see Figure 5.13).



Figure 5.13: *The expectation pattern used in the simulation.*

The expectation pattern is matched against all pixel positions according to (5.6) yielding a distribution of control signals composed of gray values in the

range from 0 to 255 mapped onto the interval  $[-1, 1]$ . The maximum value  $v_{max} = 1.0$  corresponds to the location of the subimage containing Paul McCartney's face (see the right part of Figure 5.12). The gray-value histogram is shown in Figure 5.14. The inset shows an enlarged view on the maximum value (255) and its nearest competitors.

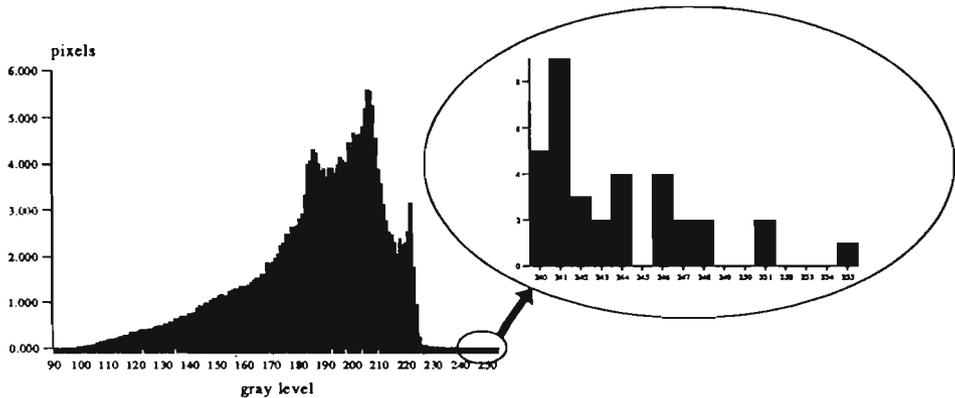


Figure 5.14: *Gray-value histogram of the match-value distribution.*

The normalised distribution forms the input of the gating network that is composed of 29524 gating lattices containing  $3N = 33^2$  gates each. Gate updating is based on activity-conserving dynamics. The intrinsic-noise parameter is set to  $T = 1.4$ .

## Results

Figure 5.15 shows the result of the simulation. Convergence of the network is achieved at  $t \approx 300$ . Except for the bottom two levels (i.e., levels 10 and 9), the convergence time grows with descending  $l$ . The convergence value of the base level,  $m_{beam}(t, 10)$ , is rather small:  $\approx 0.7$ . At subsequent levels the convergence value increases steadily to a value of  $\approx 0.95$  at the top level.

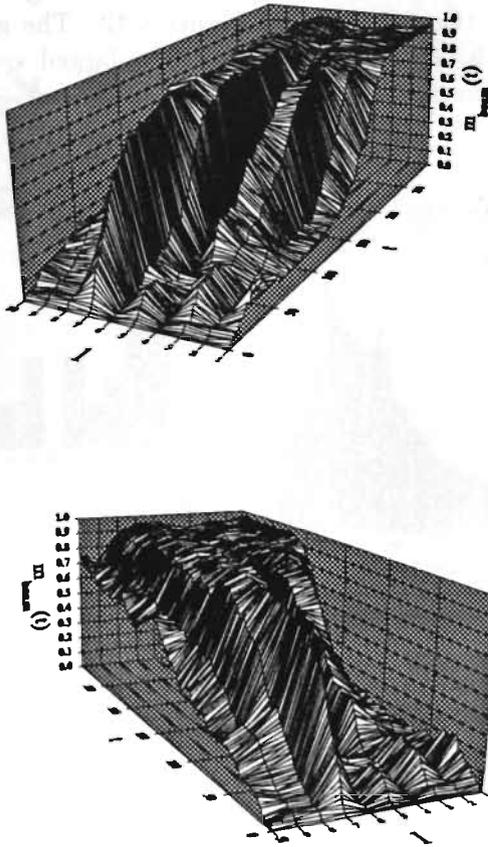


Figure 5.15: *Two perspectives on the SCAN order parameter as a function of  $t$  and  $l$ .*

## Discussion

Our results show that SCAN can cope with a natural image of a realistic size. Within 300 iterations, the target pattern is channelled to the top level. Although, the convergence values of the bottom two levels are rather small,

they do not impair performance at subsequent levels. This is due to the concentration of large control signals near the target. Taking into account the multiplicative effect of the gating lattices making up the attentional beam, the overall gating quality  $m_{beam}$  is:

$$m_{beam} = \prod_{l=1}^L m_{beam}(t, l). \quad (5.21)$$

The overall gating quality of our simulation is  $m_{beam} \approx 0.30$ . This implies that more than two thirds of the attended pattern is lost during channelling. Such an information loss hampers the application of the SCAN model to realistic problems.

### 5.4.3 Performance with Time-Varying Intrinsic Noise

The gating quality may be improved significantly when the intrinsic noise is varied over time. Strictly speaking, intrinsic noise is a property of the gating element that is not controllable. We assume, however, the presence of an external noise source that superimposes noise onto a “resting level” of intrinsic noise.

We model time-varying intrinsic noise as follows: initially, a sudden rise of the intrinsic noise associated with stimulus presentation facilitates the destruction of the initial ordered state. Subsequently, a gradual lowering of the intrinsic noise allows the gating lattices to converge to a completely ordered state. This idea was examined in a simulation employing time-varying intrinsic noise.

#### Procedure

The same procedure is followed as in the first simulation, except that in this simulation Glauber dynamics are used and the intrinsic noise is varied as a function of time. At  $t = 0$ , the intrinsic noise is set to a value  $T(0) (> T_c)$ . This value was maintained for a fixed period of time, called the *sustain* time  $t_{sust}$ . Then, at  $t = t_{sust} + 1$ , the intrinsic noise decays according to  $T(t) = dT(t - 1)$ , where  $d (\in \langle 0, 1 \rangle)$  is a decay parameter. After a decay time  $t_{decay}$  that depends on the value of  $d$ , the intrinsic noise value reaches a minimal value  $T_{min}$ . In the present simulation, at  $t = 0$ , the intrinsic noise of all gating lattices is set

to  $T(0) = 2.0$ . This value of  $T(0)$  has been verified to yield a disordered state within 10 iterations. The decay parameter is set to  $d = 0.990$  for which it has been verified that the gating lattice converges to an ordered state when presented with the target signal and its two distractors. For this value of  $d$ , the minimal value  $T_{min} = 0.01$  is reached within 100 iterations. For all gating lattices at level  $l$ , the sustain time is defined as:  $t_{sustain}(l) = 100(l - 1) + 10$ . The time-dependent curves of  $T(t, l)$  are shown in Figure 5.16.

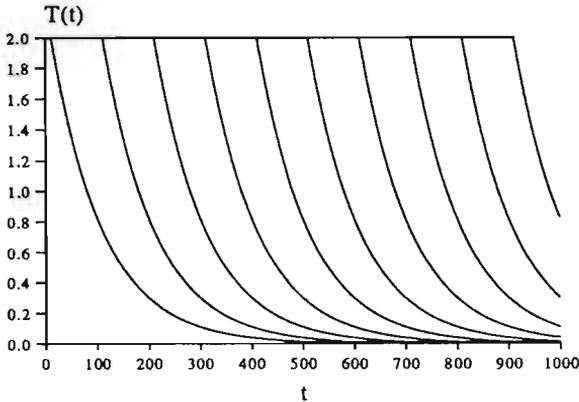


Figure 5.16: *Intrinsic noise as a function of time for levels  $l = 10$  to  $l = 1$  (from left to right).*

## Results

Figure 5.17 displays the results obtained with time-varying intrinsic noise. All lattices within the beam reach a perfect gating quality  $m_{beam}(t, l)$  within approximately 100 iterations after  $t_{sustain}(l)$ , i.e., the time that the intrinsic-noise starts to decay. The rapid increase of  $m_{beam}(t, l)$  is the result of the phase transition induced by the transition through  $T_c$ . After 1000 iterations, the overall gating quality is perfect (i.e.,  $m_{beam} = 1.0$ ).

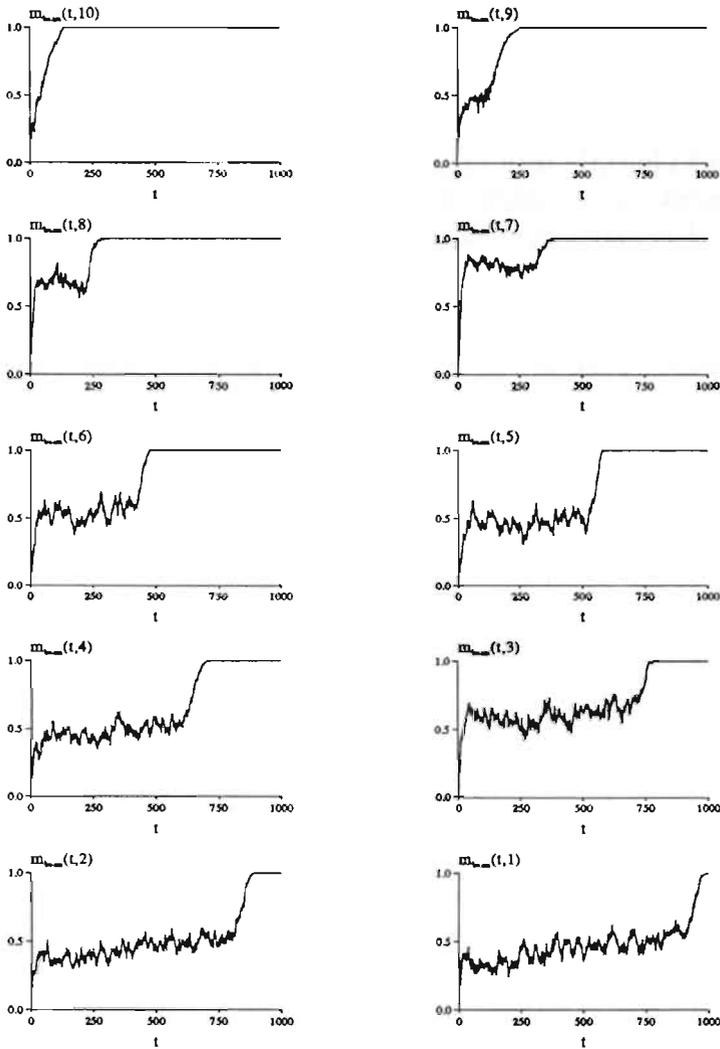


Figure 5.17: Development of  $m_{beam}(t, l)$  for  $l = 10$  to  $l = 1$ .

## Discussion

The perfect results obtained show that SCAN scales well to a problem of a realistic size. The price to be paid for the perfect gating quality is the

additional connectivity required to control the intrinsic noise of the gating elements. Assuming a single time-varying noise source per level, each noise source has to make connections to all gates in the gating lattices at that level. Although the total number of additional connections needed is large, this number grows linearly with the size of the gating network. Alternatively, each gating element may control its intrinsic noise locally in response to changes in the input. The sustain period, or the latency of the onset, should then be a function of the level of which the gate is part.

Time-varying intrinsic noise serves two goals in our simulation. First, the initial sustained peak at  $t = 0$  restores the symmetry that was broken by a previous settling of the attentional searchlight. Second, the exponential decay is a form of *simulated annealing* (Kirkpatrick, Gelatt, and Vecchi, 1983) that enhances the speed and quality of gating. The main parameter of the intrinsic-noise dynamics is the decay. The rate with which the intrinsic noise decreases determines the quality of the ordered state. The slower the decay the better the quality of the solution.

It is interesting to note that an analogue of the time-varying noise employed in our simulation may exist in biology. In response to stimulus presentation, the typical neural cell starts with a transient peak after which it decays exponentially to a resting level (Albrecht, Farrar, and Hamilton, 1984). It has been argued that this response characteristic of neural cells serves a dual purpose. The initial response signals the presence of a stimulus, whereas the decay response carries the stimulus information (Dinse, Best, and Krüger, 1990; Rybak *et al.*, 1991). Additional evidence for the importance of time-varying noise comes from the study of Happel and Murre (1994). They used a genetic algorithm to create modular neural network for digit recognition and obtained network behaviour in which an initial noise burst was followed by a reduced noise level. Although it is not clear if and how our intrinsic noise relates to neural activity, one could interpret the neural-response profile as a means of restoring symmetry combined with rapid annealing.

#### 5.4.4 Improving SCAN: Scanning Scale Space

The expectation pattern (template) used in this simulation represents a rather unrealistic aspect. An exact expectation is only rarely available. A good strategy to follow when no exact expectations are available is to perform a

hierarchical search in scale space (Koenderink, 1990). SCAN can be extended to incorporate selection of scale and to be capable of scale-invariant object recognition by adding a large gating lattice at the base level that selects one of a number of scales. Initially, a general hypothesis or expectation is formed, e.g., “select bright spots”. Subsequently, more detailed expectations are matched against image data, e.g., “oval shapes”. And finally, a detailed expectation leads to a small number (or single) of candidates, e.g., “human face (of Paul McCartney)”. The causal structure of scale space (e.g., Koenderink, 1990) facilitates the localization of objects in large patterns. Tracing the “object fingerprint” at a coarse resolution towards one of its possible causes at the level of fine spatial details represents an efficient visual search procedure (Witkin, 1983; Witkin, Terzopoulos, and Kass, 1987; Olshausen, 1994).

#### 5.4.5 Related Work

As an extension of their earlier work (Anderson and Van Essen, 1987), Olshausen, Anderson and Van Essen (1993; see also Anderson and Van Essen, (1992) and Olshausen (1994)) proposed a dynamic-routing model that shares many similarities with our work. Their routing model achieves translation and scale invariance by means of a special type of gating network in which routing and Gaussian blurring are combined. The gating elements in their model are links that differ from our gates in that their states (weights) are defined by continuous variables. The values of the weights are set by control neurons that are part of a Winner-Take-All network that is based on Hopfield (1984). The inputs of the control neurons are similar to our control signals. Just as in SCAN, these control signals may be derived from cross-correlating an object representation (the output of an associative memory) with part of the input.

The routing model differs from SCAN in several respects. First, in the routing model the attentional beam is controlled through localised competition, whereas in SCAN distributed competition is employed. Second, the routing model allows for scale-invariant pattern processing, whereas SCAN does not. (SCAN can be extended in a straightforward way to include scale invariance, see Section 6.2.) Third, the routing model employs deterministic mean-field dynamics, whereas SCAN is based on stochastic dynamics. The main difference between the routing model and the SCAN model, however, is a difference in emphasis. While in the routing model biological plausibility is stressed, in our model the emphasis is on scalability. Simulations of the routing model

have been limited to two-level networks (Olshausen, 1994). It is therefore not clear if the routing model scales well to large images.

## 5.5 General Discussion

The simple way in which the problems of translation invariance and attentional selection are solved in SCAN (essentially through the use of cascaded parallel switches), contrasts sharply with the more complicated computational schemes used in other approaches (e.g., Fukushima, 1989; von der Malsburg, 1990).

These alternative approaches all rely on an initial decomposition of the visual image into elementary features. Since each feature can be detected at each position, objects at different positions in the visual image do not change the pattern of detected features. The main problem associated with feature decomposition is however, the binding of the features into a single object representation. This binding problem (cf. Chapter 1) represents a major problem for vision architectures based on feature decomposition. In SCAN, the binding problem does not arise, since patterns are not decomposed (cf. Anderson and Van Essen, 1992). Instead of throwing away information on spatial relations and form at the first stage of processing, SCAN preserves this information to present it to a centralized pattern classifier.

The idea of using neural “switches”, originates from Anderson and Van Essen (1987) and Baron (1987). It lies at the basis of SCAN, but we consider the characteristics summarized in Section 2.6 essential for its architecture. In particular, the characteristics of stochasticity, profligacy, and sparse connectivity, yielded the building block of SCAN, a parallel switch in which attentional selection is distributed. The characteristic of hierarchy led to the formulation of SCAN as a hierarchical structure in which the attentional-selection task is spatially decomposed. The characteristic of invariance is evident in SCAN; at the base and subsequent levels, many gating lattices operate concurrently to reformat localized regions of the visual input.

This chapter provided a detailed specification of SCAN’s structure and dynamics. Simulation studies revealed the capacity of SCAN to select salient or expected parts of a large input image. In Chapter 6, the SCAN model is evaluated as a biological or psychological model of attentional selection. It will appear that, at a global level, the architecture of SCAN is consistent with

---

qualitative and quantitative aspects of the biological visual system. Moreover, the biological example provides hints as to how SCAN can be modified and improved.



## Chapter 6

# Evaluating SCAN

This chapter discusses how the structure and behaviour of the SCAN model relate to the structural and behavioural characteristics of the mammalian visual system. Evaluating SCAN from a biological point of view helps to identify its weaknesses; simultaneously, it provides suggestions for extending and improving the original model.

In Section 6.1, we discuss the available evidence for gating in the brain. Then, in Section 6.2 quantitative considerations are presented indicating the number of components needed to build a SCAN model on a biologically realistic scale. Section 6.3 treats the behavioural aspect of perceptual stability. In Section 6.4 some incompatibilities between SCAN and the visual system are highlighted; they are taken as starting points for improving the model. Finally, Section 6.5 concludes the chapter.

### 6.1 Gating in the Brain

The biological plausibility of SCAN as a model of covert attention depends on the existence of biological analogues of its main components, i.e., the gating element, the gating lattice, the gating network, and the classifier network. This section discusses the interneuron, the interneuron network, and the temporal pathway of the visual system, respectively, as possible neural realisations of SCAN's main components.

### 6.1.1 Gating by the Interneuron

The gating element is the basic element of the SCAN model. Assuming that a gating operation plays a central role in the covert-attention process in the brain, some neural basis for it should exist. Marin-Padilla (1990) hypothesized that cortical *interneurons* modulate signal flow through pyramidal cells (cf. Section 2.3.2). In particular, the axoaxonic or *chandelier cell* plays an important role in controlling pyramidal output (Somogyi and Martin, 1985; Marin-Padilla, 1987; Van Essen and Anderson, 1990). Our gating element (cf. Figure 4.3) realises this hypothesis; the signal element represents a pyramidal cell whose output is modulated by an interneuron (the control element). The gating element is unlikely to represent a realistic model of the full pyramidal-interneuron interaction. Real pyramidal and chandelier neurons are certainly more complex than the binary signal and control element that make up our gating element (see e.g., Shepherd, 1990). Moreover, it has been argued by Douglas and Martin (1990a) that the output of pyramidal cells can not be shut off completely by inhibitory inputs. However, recent work by Bush and Sejnowski (1993) reached the conclusion that

...strong cortical inhibition is able to prevent the firing of even strongly driven pyramidal cells... (p. 99).

### 6.1.2 Gating by the Interneuron Network

Considering the control element of a gate as an (inhibitory) interneuron, the gating lattice models the *interneuron network*, a lattice of cells assumed to modulate signal flow through pyramidal (projection) neurons (cf. Postma, Van den Herik, and Hudson, 1994a). The axoaxonic cells contact each other and the initial segment of the axon of pyramidal neurons in the superficial layers of cerebral cortex and are therefore believed to control pyramidal output to other cortical areas. In the words of Somogyi and Martin (1985):

...given the appropriate input, the axoaxonic cell could provide a most effective control of the output of projection neurons. If this control were applied selectively to specific cortical regions, it could provide a basis for the mechanism of selective attention (p. 517).

It appears, then, that the pool of excitatory signal-carrying pyramidal neurons are interspersed with inhibitory interneurons that shape the distribution of the outgoing signals. A two-dimensional gating lattice is probably too simple a model for a cortical interneuron network; a three-dimensional gating lattice with a large number of sublattices seems more appropriate. For each cortical area, a cerebral Gating Lattice might control its rewiring to other areas.

Speculating, having about 300 projection cells contacted by each axoaxonic interneuron (Douglas and Martin, 1990b) allows one to derive certain consequences for the complexity of cerebral gating lattices. The maximum number of sublattices in cerebral cortex may, accordingly, also be estimated as 300. Obviously, these pyramidal cells need not all be located on distinct sublattices, so that the actual number of sublattices is probably considerably smaller. If we assume that each sublattice gates a spatial-frequency band or a spatial position, then attentional selection could be reflected by the opening of all gates on a single sublattice in the cerebral gating lattice. Such a mechanism could subserve the construction of invariant representations at higher visual areas as modelled by SCAN.

### 6.1.3 Gating Network in the Temporal Pathway

In Section 2.3.4 we interpreted the temporal pathway as a hierarchically structured system where positional dependency is gradually removed (see for a discussion: Plaut and Farah, 1990). Our hypothesis - that routing as performed by the gating lattices in SCAN may be analogous to the action of the interneuron network - fits nicely with the interpretation of the temporal pathway as a hierarchical pathway for identification. Patterns leaving cortical area V1 towards V2 are gated by projection neurons whose outputs are modulated by the interneuron network so that the input of V2 is under dynamic control of the V1 interneuron network. A similar scheme, laid out over the respective interneuron networks of V1, V2 and V4, accommodates routing that extends over the entire retinal sampling area. The SCAN model may thus be interpreted as a model, albeit a crude one, of the temporal pathway or identification pathway. As such it shares a number of global features with its biological counterpart. The increasing size of the input dealt with by gating lattices at higher levels of SCAN corresponds with the gradual increase of receptive-field size observed when ascending the temporal pathway. The selective action of the gating lattices is in accordance with the receptive-field

modulation observed at the higher levels of the temporal pathway. The lack of convincing evidence for attentional modulation at lower levels (V1 and V2) may be due to the relatively small magnitude of the induced shifts at these levels. There has been a report that states that the proportion of cells exhibiting attentional effects increases when going from V1 to V4. These proportions are 0.20 for V1, 0.50 for V2, and 0.50 for V4 (Motter (1988) cited in Colby and Duhamel (1991); Motter, 1993). For the inferotemporal areas, this proportion is about 0.67 (Colby and Duhamel, 1991).

#### 6.1.4 Classification in the Inferotemporal Cortex

Placing a classification network at the top of the SCAN hierarchy is consistent with the alleged function of the inferotemporal areas as being involved in the matching of stored object prototypes against visual data (e.g., Li, Miller, and Desimone, 1993). Section 2.3.4 discussed the neurophysiological evidence that such a matching function is performed in the inferotemporal areas.

#### 6.1.5 Parallel Matching at V1 and V2

In SCAN, the actual matching process also involves the base level of the gating network. Clearly, an explicit template matching as used in SCAN is highly unrealistic. It is more likely, however, that object representations exert their influence on the base level in a more subtle way, e.g., by changing the response characteristics of striate cells (Tsotsos, 1991). This leads to the hypothesis that, in the visual system, such matching is partly performed right down in the striate cortex. Consequently, to allow such a process to take place, representations stored at the level of the inferotemporal areas need to be able to exert their influence at the level of V1 and V2. There are several lines of evidence that indicate that this is indeed the case in the visual system.

First, as already discussed in Chapter 2, for virtually every feedforward pathway in the brain there exists a rich feedback pathway (DeYoe and Van Essen, 1988); the hardware required for top-down matching at V1 is available. Mignard and Malpeli (1991) state that:

...areas of cortex thought to be carrying out low-order analysis may be able to extract considerable information from higher order

areas (p. 1249).

Second, several studies indicate that visual processes known to occur early in the cortical hierarchy can be affected by higher, more abstract, processes (Gilchrist, 1979; Knill and Kersten, 1991). This implies that representations that are active at the level of semantic areas, such as the inferotemporal areas, modulate processes in the peripheral areas.

Finally, the behavioural studies performed by Biederman and Cooper (1991; 1992) indicate that visual expectations, activated by prior presentation of visual images, generalize over retinal position and spatial scale.

Viewing areas V1 and V2 as the sites where stored representations are matched against environmental stimuli may explain why these area forms such a large part of the visual system (more than 40% according to Felleman and Van Essen, (1991)). The neural machinery needed to perform the parallel matching of stored representations across the whole retinal scale space certainly involves a large amount of neural "hardware". Nevertheless, major savings can be achieved by applying an optimized representation of visual data (see below).

## 6.2 Quantitative Considerations

The feasibility of the SCAN model from a biological or engineering perspective is dependent on the resources (e.g., number of gates) required for actual realisation. Basing ourselves on the dimensions of Koenderink and Van Doorn's (1978) Stack model, we assume an input image comprising  $4 \times 10^4$  sampling points. The attentional window is taken to have  $N$  sampling points.

The gating network, thus dimensioned, is capable of routing a contiguous subpattern (of  $N$  elements) from any of the  $4 \times 10^4$  locations. Assuming a triangular gating tree this requires approximately  $L = 10$  layers ( $3^{10} \approx 6 \times 10^4$ ). The total number of gating lattices in such a tree is

$$\sum_{l=0}^9 3^l = 1 + \frac{3^{10} - 3}{2} = 29524. \quad (6.1)$$

Each of these gating lattices is associated with a single triplet of gates. There

are therefore  $3 \times 29524 = 88569$  such gates. Additionally, each lattice contains  $3N$  gates. A total number of  $88569(1 + 3N)$  gates are therefore needed to achieve complete translation invariance for recognition of patterns defined by  $N$  contiguous sampling points.

A straightforward extension of SCAN leads to scale-invariant pattern processing. A variable-size attentional window may be realized by placing a (Potts) gating lattice with  $\approx 40$  sublattices just prior to the base level of the tree. This gating lattice selects one of 40 sampling scales (i.e., one of the levels of the stack model in Figure 2.3). The output of the gating lattice is then a scaled pattern of  $3^{10}$  elements containing an image ranging from a wide-angle coarse-resolution view to a small-angle high-resolution view. Because 40 gates are needed for each sampling point, the gating lattice contains  $1.6 \times 10^6$  gates. Consequently, a SCAN model of biologically plausible dimensions requires  $1.6 \times 10^6 + 88569(1 + 3N)$  gates.

In the SCAN model a single gating lattice is associated with each input position. Therefore,  $3N$  pyramidal-interneuron pairs should be allocated for each retinal input. The size of the attentional window  $N$  has been estimated to be on the order of  $10^2$  elements (Nakayama, 1990) to  $30^2$  elements (Olshausen, 1994). We take the largest estimate to obtain the total number of gates required:  $1.6 \times 10^6 + 88569(1 + 3N) \approx 81$  million gates.

Recalling from Chapter 2 that for each of the  $1.5 \times 10^6$  LGN inputs there are about 100 to 1000 V1 neurons available, the total number of these first-stage neurons ranges from 1.5 billion to 15 billion. Apparently, the number of first-stage neurons outnumbers the total amount of neural elements needed in the biologically plausible SCAN. We therefore conclude that there are ample neural elements in the temporal pathway areas V1, V2, V4, and IT for SCAN to be realised.

### 6.3 Stable Perception

The SCAN model provides an explanation for the behavioural finding that object identification can take place independently of the retinal position and the size of the object. Its attentional beam operates on an object-centred frame of reference because it is controlled by signals linked to the shape of the object. In the brain the perceived stability of the visual scene over mul-

multiple fixations can be achieved in an analogous way. A peak in the match distribution, signifying a good match between the top-down expectation and bottom-up visual data, shifts to a new position with each saccadic eye movement. The attentional searchlight follows the movement of the peak with a delay inversely proportional to the saliency of the visual data (height of the peak). At any moment the searchlight's position represents the origin of the coordinate system for further image analysis. The attended object serves as an anchoring point that is stable across eye, head, or body movements. A problem arises, however, when the object being attended to leaves the field of view as a result of a large saccade. This problem can be mitigated by maintaining a multiscale representation of the attended object, one encompassing the range from a small-field high-resolution view to a large-field low-resolution view. Although, after a large saccade, the attended object inside the small-field view may be lost, (part of) its context may still be available. We note here that the perception of a stable world is based on an even more complex interplay of information sources (see for a discussion: Bridgeman, Van der Heijden, and Velichkovsky, 1993).

SCAN does not address the more general, but higher-level, problem of maintaining and updating a representation of spatial relations in and between visual objects (see Hinton (1981) for a discussion, Hinton and Lang (1985) for a modelling approach, and Yeshurun and Schwartz (1989) for an algorithmic approach).

## 6.4 Incompatibilities

There are also some apparent incompatibilities between the SCAN model and its biological counterpart. We discuss two of them in detail and demonstrate how they can guide extension and improvement of the model.

### 6.4.1 Representation of Visual Data

A major point of disagreement between SCAN and the visual system concerns the representation of visual data. In Chapter 5, the explicit assumption is made that the input to SCAN is formed by a representation of the visual data corresponding to the retinal sampling pattern. The attended part of the

image is assumed to be routed upward in its original form. It is generally appreciated, however, that such a representation is not very parsimonious (Daugman, 1988; Koenderink and Van Doorn, 1990). The standard image format appears to be highly redundant for natural image data. Suppose, for instance, that we examine the inputs to two adjacent retinal receptors as they are exposed to natural images. Whenever one of the two receptors samples a visual signal of a certain intensity, there is a good chance that its neighbour will sample a signal with the same intensity. Inspection of the pixel values of any digitized natural image reveals that many proximate pixels have the same gray value. The reason being, of course, that the objects making up the image are often composed of relatively large uniformly illuminated regions. In other words, the inputs of adjacent receptors are not statistically independent. The implication is that the representation of visual images as used in SCAN, i.e., two-dimensionally structured intensity distributions, can be enhanced.

It appears that one of the main functions of striate cortex is the removal of such redundancy inherent in the retinal image. The sensitivity of striate cells to orientation and spatial frequency (see e.g., Daugman, 1988) accords with this function. The selective response of striate cells to oriented stimuli at several spatial scales reflects an emphasis on *changes* in illumination rather than on absolute illumination values, the former being far more informative than the latter. The transformation of the visual data sampled by the retina into a more economical format present enormous savings on the wiring and resources at further cortical stages. Sanger (1989) showed how a representation of natural images in terms of its principal components can lead to a drastic decrease in the bits-per-pixel rate. Application of his decorrelation algorithm (also formulated as a neural network) to an 8 bit-per-pixel coded image (256 grey levels) yields a new representation with an information rate of 3.4 bit-per-pixel.

The performance of the SCAN model can be improved in a straightforward way by adopting the strategy of reducing redundancy at the initial level of the identification hierarchy. Instead of using the raw image data for matching and routing, a non-redundant image representation can be used equally well. One implication is that due to the reduction in information rate per pixel, smaller gating lattices are needed to achieve the same attentional-window size. Another implication is that the match-value distribution obtained by the parallel matching of an expectation pattern with image parts may actually more effective with transformed image data (Mallat, 1989; Casasent, Smokelin, and

Ye, 1992).

### 6.4.2 Randomness of Individual Cortical Connections

The connectivity pattern in SCAN obeys rather strict rules with respect to topography to ensure proper gating. The pattern of connections between two adjacent gating lattices forms a bijective mapping. Although the composite mapping formed by an attentional beam, i.e., the connected sequence of  $L - 1$  such bijective mappings, does not need to preserve the topography of the input image (any permutation of the connections is allowed because classifier networks do not employ the absolute spatial positions of its inputs), it has to be identical for all  $3^L$  possible sequences.

This *consistent-mapping* requirement is at odds with the observed pattern of individual cortical connections. While at a macroscopic scale cortical areas are interconnected in an orderly way (see Section 2.4.1), at the level of individual connections there appears to be a considerable degree of randomness in their termination pattern. It seems therefore highly unlikely that neural connections within a cortical pathway are “labelled”, in the sense that the position of a target neuron is of significance for the analysis of its activity at further stages. A plausible line of thinking is based on the idea that the way in which neural signals co-vary carries all the information necessary for object identification (Koenderink, 1984; MacKay, 1985). Covariational analysis implies that, since neural signals are labelled intrinsically by the simultaneity of their activity pattern, no stringent connectivity requirements are needed. At some convergence site, e.g., area 46 (Young, 1992), all signals that co-vary within a small time window can be taken to have a common cause.

If the temporal dimension of visual signals is not exploited, an alternative solution has to be found to relax the consistent-mapping requirement. Let us focus for the moment on a gating network in which the requirement is *not* fulfilled. Clearly, its output will be a permutation of the attended part of the input that is different for each input location. The problem to solve is to make the permutation consistent for all input locations. This can be done as follows. Let us denote the bijective mapping from a gating lattice in layer  $l - 1$  to one in layer  $l$  by  $\mathcal{M}(l)$ . The composite mapping formed by a connected sequence of linear bijective mappings can now be written as  $\mathcal{M}(L) \circ \mathcal{M}(L - 1) \dots \circ \mathcal{M}(1)$ . Any concatenation of linear mappings can be rewritten as a single

mapping. The composite mapping can be written as  $C(i)$ , a single mapping specifying the permutation of signals along the entire sequence. Since the permutation is different for each input position, the composite mapping is indexed by  $i \in \{1, 2, \dots, 3^L\}$  (see Chapter 5). Our objective is to make all mappings  $C(i)$  consistent. To do this we form a composite mapping  $C(i) \circ D(i)$  for each input position  $i$ . In SCAN the mapping  $D(i)$  can be implemented by adding an intermediate level between the base-level gating lattice and its inputs for each input position. The idea is now to obtain a consistent mapping for all input positions by choosing the mappings  $D(i)$  in such a way that they “translate” their associated mapping  $C(i)$  into a uniform composite mapping  $C$ , i.e.,  $C(i) \circ D(i) = C$  for all  $i$ .

## 6.5 Concluding Remarks

This chapter evaluated the SCAN model from a biological and psychological point of view. Comparison of the model with its biological example led us to propose several extensions of the original model. However, we did not discuss extensions that endow SCAN with a knowledge-based scanning procedure. Stored spatial knowledge may indicate which locations are candidates for attentional inspection. Such spatial templates lead to attentional routines analogous to the visual routines discussed by Ullman (1984). Inspected locations may be suppressed for some time to prevent the attentional searchlight from visiting the same location twice (Milanese, 1993). In fact, there is biological and psychological evidence for such “inhibition of return” effects (Posner and Presti, 1987). Klein (1988) reported behavioural evidence indicating that visited locations were tagged adaptively to enhance the attentional inspection of unvisited locations. SCAN can be endowed with similar spatial operations by extending the model with a second pathway analogous to the parietal pathway (see Chapter 2).

## Chapter 7

# Evaluation and Conclusion

In this final chapter, we evaluate the results of our approach. With respect to our model we conclude that SCAN succeeds in combining active vision, translation invariance, and scalability. As to our approach, we conclude that viewing the brain as a neural machine incorporating a solution to the scaling problem represents a fruitful line of research to pursue.

### 7.1 Evaluation of the Interdisciplinary Approach

The SCAN model is the result of using biological and psychological ideas in developing a scalable model of covert attention. We phrase these ideas in the language of statistical mechanics. Our computer-science perspective leads to an emphasis on scalability; the SCAN model provides a blueprint for the realisation of a parallel distributed active-vision machine applicable to problems of a real-world size. The combination of elements from four disciplines satisfies the first two objectives stated in Section 1.3: *specifying a scalable model of covert attention inspired by a biological example* and *formalising the model in terms of statistical mechanics*. The advantage of our interdisciplinary approach is that it provides a new perspective on the subject of visual attention. Biological ideas, when considered in isolation, do not provide an unequivocal identification of the underlying mechanisms and processes. However, the searchlight metaphor of covert attention, used in psychology, helps to interpret the biological findings. Adding the computer-science perspective

to both disciplines leads to an appreciation of the problems associated with building scalable parallel distributed computers. The biological characteristics of sparse connectivity and profligacy make sense from this perspective. Finally considering the statistical mechanics of lattices of stochastic processing elements leads to a detailed specification of the strengths and limitations of sparsely-connected neural networks.

## Evaluation of the Biological and Psychological Characteristics

In this Section, we evaluate the implications of the five characteristics listed in Section 2.6: *stochasticity*, *sparse connectivity*, *profligacy*, *hierarchy*, and *invariance*. Additionally we evaluate the implications of adopting the statistical-mechanics formalism to realise the characteristics in a model.

*Stochasticity:* The stochastic dynamics of the gates reflect the noisy neuron dynamics. The use of stochastic processing elements enabled the formalisation of the neural and gating lattices as Ising lattices. The properties of these lattices depend critically on the presence of intrinsic noise in their constituent processing elements. There is a tendency in neural-network research to reformulate stochastic models in terms of mean-field theory. The Boltzmann machine (Ackley, Hinton, and Sejnowski, 1985), originally based on stochastic processing elements, has been reformulated in terms of a set of deterministic mean-field equations by Peterson and Anderson (1987, see also Hinton, 1989). The deterministic equations lead to an improved solution quality and speed. However, averaging out stochastic fluctuations, as in the mean-field approach, disregards the problem of how to deal with noise intrinsic to biological processing elements (e.g., Kosko, 1989). Realisation of mean-field approaches in analogue hardware (see, e.g., Hopfield, 1990) still has to deal with the intrinsic noise of operational amplifiers. We believe that instead of avoiding noise it should be exploited. The stochastic nature of processing in the gating lattice relaxes the requirement for noiseless components in a hardware realisation. In fact, noise may help rather than hinders efficient processing.

*Parsimony:* The parsimony in wiring observed in the brain led us to define the gating lattice as a locally interconnected network. Although only nearest-neighbouring gates are connected, the effect of this pattern of connectivity may spread out over the entire lattice surface. This property of locally interconnected networks deserves more attention, both in biology and

in neural-network research (cf. Kruglyak and Bialek, 1993; Henseler, 1993). The prevailing emphasis on densely-connected networks results from the analytical advantage of the mean-field approach. Parsimony is, of course, directly related to the scaling problem (Minsky and Papert, 1988). Upward scaling of SCAN to real-world-problem sizes, through the addition of levels, leads to only a moderate decrease in solution quality. In contrast, the performance of densely-connected networks breaks down much sooner (Hopfield and Tank, 1985; Minsky and Papert, 1988). Although the computational tasks faced by densely-connected networks are different from those of the SCAN model, the main point is that the development of the former is not constrained by the requirement of scalability.

*Profligacy.* The availability of a large number of processing elements in the brain led us to consider large lattices. Increasing the number of gates in a gating lattice has a positive effect on its robustness to noise. Increasing the number of stochastic gates not only helps to reduce the stochastic fluctuations in the order parameter, but also amplifies the cooperative properties of the lattice because the susceptibility or responsiveness of the gating lattice increases with the number of gates (Binder and Landau, 1984).

*Hierarchy.* The hierarchical organization of the visual system and, in particular, of the temporal pathway, led us to define SCAN as a hierarchical neural network. As a result, the channelling problem of how to route contiguous parts of the retinal image towards a centralized pattern recognizer could be solved with moderate wiring costs. As discussed in Section 4.6, the characteristic of hierarchy may also be applied to gating lattices.

*Invariance.* The interpretation of the covert-attention searchlight as a means of achieving translation invariance enabled a straightforward solution to the problem of removing positional dependency. In combination with the hierarchical network structure, application to large-scale problems became feasible.

Each of these five characteristics, based on well-established biological and psychological findings, contributes to the performance of the SCAN model as a whole. If we ignore stochasticity, and assume deterministic processing elements instead, the result would be a rigid cellular automaton failing to cope with extrinsic noise and lacking the flexibility to respond rapidly to changing inputs. As to parsimony, we may state that, although a high resistance to extrinsic noise is obtained with fully-connected networks, they may exhibit poor flexibility. Moreover, the costs in wiring do not outweigh the gain in

robustness (cf. Chapter 3 ).

Neglecting the profligate assignment of processing elements to computational problems would degrade gating behaviour and resistance to noise. Finally, without the characteristics of hierarchy and invariance, it would have been difficult to realise a scalable model of the covert-attention process.

### **Evaluation of Statistical Mechanics as a Formalism**

We have defined our model as an Ising lattice to enable a statistical-mechanics formulation of its properties. We could therefore utilize the large body of theoretical and experimental (Monte-Carlo) results on the Ising lattice. These results provided important insights into the emergence of cooperativity among the processing elements in the neural and gating lattices. As demonstrated in Chapter 3, the effect of connection density on the performance measures of robustness and flexibility could be readily related to the critical intrinsic-noise levels of the underlying Ising lattices. Through the notions of dynamic scaling and critical exponents, we were able to characterise the dynamic behaviour of the gating lattice near the critical intrinsic noise level. To establish formal equivalence with the Ising lattice, we were forced to assume a regular lattice structure in the gating lattice. According to the universality concept, the properties of the gating lattice may generalize to more realistic models with a less regular lattice structure belonging to the same universality class.

### **Conclusion on the Interdisciplinary Approach**

We conclude that biological and psychological characteristics play a pivotal role in the realisation of a scalable model in which attentional selection and translation invariance are combined. Furthermore, we conclude that formulating SCAN in terms of the statistical mechanics of Ising lattices yields a precise characterisation of its behaviour.

## 7.2 Evaluation of SCAN

The SCAN model has been subjected to attentional tasks in Chapter 5. Although a strict initialization procedure is used in the simulation studies, a flexible response to changing input distributions is still observed. SCAN thus exhibits the flexibility of the attentional searchlight. Since attentional orienting in SCAN is linked to the channelling of the attended pattern to a centralized pattern classifier, we achieved translation invariance as stated in our third objective in Section 1.3. The evaluation of SCAN from a biological and psychological perspective in Chapter 6 (i.e., the fourth objective in Section 1.3), revealed that SCAN agrees to a large extent with biological and psychological data. We therefore conclude that SCAN represents a scalable model of covert attention that combines active selection with translation-invariant pattern processing.

## 7.3 General Evaluation and Conclusion

Minsky and Papert (1988) stated that

...our analysis does show that parallel networks are, in general, subject to serious scaling phenomena. Consequently, researchers who propose such models must show that, in their context, those phenomena do not occur (p. 264).

Our approach takes the scaling problem seriously and makes use of the biological paradigm that uses neural networks to deal with large-scale problems. The problems that occur when SCAN is scaled up have been successfully solved by employing transient noise. Furthermore, formulating scalable neural models requires restricting the range of possible architectures. The main restriction concerns the density of connections. Although for certain tasks (e.g., pattern storage) fully-connected networks are the only solution, for many tasks a more parsimonious use of connections is appropriate. The channelling of patterns to a centralized pattern classifier in SCAN confines the expensive dense connectivity pattern of the classifier to a single location, while a much sparser connectivity pattern suffices for the spatially distributed channelling operation. We believe that our approach deserves further study to allow neural

networks to be applied to real-world problems. There are only slight penalties attached to the architecture chosen when it is required to face large problems.

In Chapter 1, our problem statement was formulated as a question: is it possible to construct a working model of covert attention as described by psychology which satisfies the constraints as set by our knowledge of biology?

Biological knowledge has served to constrain the architecture of SCAN, requiring sparse connectivity and hierarchical structure involving many simple elements. Biology also required us to confront the issue of noise. Surprisingly, noise turned out to be advantageous. This contrasts with the mean-field approaches which are mathematically rather than biologically inspired. Psychological knowledge required us to design a system which could successfully implement the searchlight metaphor as well as being sensitive to both top-down and bottom-up influences. The SCAN architecture allows expectation to steer the searchlight toward known objects but also allows salient aspects of the environment to capture attention. In addition, the computer-science approach, essentially an engineering one, complements our biological and psychological emphases. The requirement that a system should be operational led to a cooperative computational model which naturally maps onto the Ising model from statistical physics as well as onto the known neuroanatomy.

# References

- Ackley, D.H., Hinton, G.E., and Sejnowski, T.J. (1985). A learning algorithm for Boltzmann machines. *Cognitive Science*, **9**, 147-169.
- Adrian, E.D. (1946). *The physical background of perception*. Oxford: Clarendon Press
- Aharony, A. (1978). Tricritical points in systems with random fields. *Physical Review*, **B 18**, 3318-3327.
- Akutsu, N. and Akutsu, Y. (1990). Critical curve of two-dimensional Ising antiferromagnet in a magnetic field. *Journal of Magnetism and Magnetic Materials*, **90/91**, 296-298.
- Albrecht, D.G., Farrar, S.B., and Hamilton, D.B. (1984). Spatial contrast adaptation characteristics of neurones recorded in the cat's visual cortex. *Journal of Physiology*, **347**, 713-739.
- Allport, A. (1989). In M.I. Posner (Ed.), *Foundations of cognitive science* (pp. 631-682). Cambridge, MA: MIT Press.
- Amit, D.J. (1989). *Modeling brain function: The world of attractor neural networks*. Cambridge, MA: Cambridge University Press.
- Andelman, D., Orland, H., and Wijewardhana, L.C.R. (1984). Lower critical dimension of the random-field Ising model: a Monte-Carlo study. *Physical Review Letters*, **52**, 145-148.
- Andersen, R.A. (1987). Inferior parietal lobule function in spatial perception and visuomotor integration. In V.B. Mountcastle, F. Plum, and S.R. Geiger (Eds.), *Handbook of physiology, Section I: The Nervous System, Volume V*.

- Higher Functions of the Brain, Part 2* (pp. 483-518). Bethesda, MA: American Physiological Society.
- Andersen, R.A., Essick, G.K., and Siegel, R.M. (1985). Encoding of spatial locations by posterior parietal neurons. *Science*, **230**, 546-548.
- Anderson, C.H. and Van Essen, D.C. (1987). Shifter circuits: A computational strategy for dynamic aspects of visual processing. *Proceedings of the National Academy of Sciences USA*, **84**, 6297-6301.
- Anderson, C.H. and Van Essen, D.C. (1992). Dynamic Neural Routing Circuits. In D. Brogan (Ed.), *Proceedings of the Second International Conference on Visual Search* (pp. 311-319). London: Taylor & Francis.
- Anstis, S.M. (1974). A chart demonstrating the variations in acuity with retinal position. *Vision Research*, **14**, 589-592.
- Ballard, D.H. (1991). Animate vision. *Artificial Intelligence*, **48**, 57-86.
- Barlow, H.B. (1972). Single units and sensation: A neuron doctrine for perceptual psychology. *Perception*, **1**, 371-394.
- Barlow, H.B. (1981). Critical limiting factors in the design of the eye and visual cortex. *Proceedings of the Royal Society B*, **212**, 1-34.
- Baron, R.J. (1987). *The Cerebral Computer. An introduction to the computational structure of the brain*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Baxter, R.J. (1982). *Exactly solved models in statistical mechanics*. San Diego, CA: Academic Press.
- Baylis, G.C. and Driver, J. (1993). Visual attention and objects: evidence for hierarchical coding of location. *Journal of Experimental Psychology: Human Perception and Performance*, **19**, 451-470.
- Berker, A.N. (1984). Ordering under random fields: renormalization-group arguments. *Physical Review*, **B 29**, 5243-5245.
- Biederman, I. and Cooper, E.E. (1991). Evidence for complete translational and reflectional invariance in visual object priming. *Cognitive Psychology*, **23**, 393-419.
- Biederman, I. and Cooper, E.E. (1992). Size invariance in visual object prim-

- ing. *Journal of Experimental Psychology: Human Perception and Performance*, **18**, 121-133.
- Binder, K. (Ed.) (1976). Monte Carlo investigations of phase transitions and critical phenomena. In C. Domb and M.S. Green (Eds.), *Phase transitions and critical phenomena. Volume 5b* (pp. 1-105). New York, NY: Academic Press.
- Binder, K. (Ed.) (1987). *Applications of the Monte Carlo method in statistical physics (second edition)*. Berlin: Springer-Verlag.
- Binder, K. and Heermann, D.W. (1988). *Monte Carlo simulation in statistical physics*. Berlin: Springer-Verlag.
- Binder, K. and Landau, D.P. (1984). Finite-size scaling at first-order phase transitions. *Physical Review*, **B 30**, 1477-1485.
- Binder, K. and Stoll, E. (1973). Scaling theory for metastable states and their lifetimes. *Physical Review Letters*, **31**, 47-51.
- Bonda, E. and Thorpe, S.J. (1993). Identification of natural images. II. The role of attention in the processing of multi-image arrays. *International Journal of Psychology*, **72** (3/4), p. 42.
- Bridgeman, B., Van der Heijden, A.H.C., and Velichkovsky, B.M. (1993). A theory of visual stability across saccadic eye movements. Draft of target article to appear in *Behavioral and Brain Sciences*.
- Broadbent, D.E. (1958). *Perception and communication*. London: Pergamon Press.
- Brown, J.W. (Ed.) (1989). *Neuropsychology of visual perception*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, **97**, 523-547.
- Bundesen, C. and Pedersen, L.F. (1983). Color segregation and visual search. *Perception & Psychophysics*, **33**, 487-493.
- Burkhardt, H. (1989). Homogeneous structures for position-invariant feature extraction. In J.C. Simon (Ed.), *From Pixels to Features* (pp.165-173). Ams-

terdam: North-Holland.

Bush, P.C. and Sejnowski, T.J. (1993). Simulations of synaptic integration in neocortical pyramidal cells. In F.H. Eeckman and J.M. Bower (Eds.), *Computation and neural systems* (pp. 97-101). Boston, MA: Kluwer Academic Publishers.

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 category recognition codes for analog input patterns. *Applied Optics*, **26**, 4919-4930.

Caramazza, A. and Hillis, A.E. (1990). Spatial representation of words in the brain implied by studies of a unilateral neglect patient. *Nature*, **346**, 267-269.

Casasent, D.P., Smokelin, J-S., and Ye, A. (1992). Wavelet and Gabor transforms for detection. *Optical Engineering*, **31**, 1893-1898.

Chandler, D. (1987). *Introduction to modern statistical mechanics*. Oxford: Oxford University Press.

Chaudhuri, A. (1990). Modulation of the motion aftereffect by selective attention. *Nature*, **344**, 60-62.

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.

Colby, C.L. and Duhamel, J-R. (1991). Heterogeneity of extrastriate visual areas and multiple parietal areas in the Macaque monkey. *Neuropsychologia*, **29**, 517-537.

Corbetta, M., Miezin, F.M., Dobmeyer, S., Shulman, G.L., and Petersen, S.E. (1990). Attentional modulation of neural processing of shape, color, and velocity in humans. *Science*, **248**, 1556-1559.

Crick, F. (1984). Function of the thalamic reticular complex: the searchlight hypothesis. *Proceedings of the National Academy of Sciences USA*, **81**, 4586-4590.

- Crick, F. and Anasuma, C. (1986). Certain aspects of the anatomy and physiology of the cerebral cortex. In J.L. McClelland, D.E. Rumelhart, and the PDP Research Group (Eds.), *Parallel Distributed Processing, Explorations in the microstructure of cognition, Vol. II, Psychological and biological models* (pp. 333-371). Cambridge, MA: MIT Press.
- Daugman, J.G. (1988). Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **36**, 1169-1179.
- Desimone, R. (1992). Neural circuits for visual attention in the primate brain. In G.A. Carpenter and S. Grossberg (Eds.), *Neural networks for vision and image processing* (pp. 343-364). Cambridge, MA: MIT Press.
- Desimone, R., Moran, J., and Spitzer, H. (1989). Neural mechanisms of attention in extrastriate cortex of monkeys. In M.A. Arbib and S. Amari (Eds.), *Dynamic interactions in neural networks: Models and data* (pp. 169-182). New York, NJ: Springer-Verlag.
- Desimone, R. and Ungerleider, L.G. (1989). Neural mechanisms of visual processing in monkeys. In F. Boller and J. Grafman (Eds.), *Handbook of Neuropsychology, Vol.II* (pp. 267-299). Dordrecht: Elsevier Science Publishers B.V. (North-Holland).
- Desimone, R., Wessinger, M., Thomas, L., and Schneider, W. (1990). Attentional control of visual perception: cortical and subcortical mechanisms. *Cold Spring Harbor Symposia on Quantitative Biology, Volume LV* (pp. 963-971). Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
- Deutsch, J.A. and Deutsch, D. (1963). Attention: some theoretical considerations. *Psychological Review*, **70**, 80-90.
- DeYoe, E.A. and Van Essen, D.C. (1988). Concurrent processing streams in monkey visual cortex. *Trends in Neurosciences*, **11**, 219-226.
- Dinse, H.R.O., Best, J., and Krüger, K. (1990). A new dynamic model of receptive-field organization. In R. Eckmiller, G. Hartmann, and G. Hauske (Eds.), *Parallel processing in neural systems and computers* (pp. 61-64). Dordrecht: Elsevier Science Publishers B.V. (North-Holland).
- Douglas, R.J. and Martin, K.A.C. (1990a). Control of neuronal output by

- inhibition at the axon initial segment. *Neural Computation*, **2**, 283-292.
- Douglas, R.J. and Martin, K.A.C. (1990b). Neocortex. In G.M. Shepherd (Ed.), *The synaptic organization of the brain (third edition)* (pp. 389-438). Oxford: Oxford University Press.
- Dowling, J.E. (1987). *The retina. An approachable part of the brain*. Cambridge, MA: Harvard University Press.
- Driver, J. and Halligan, P.W. (1991). Can visual neglect operate in object-centred co-ordinates? An affirmative single-case study. *Cognitive Neuropsychology*, **8**, 475-496.
- Duncan, J. and Humphreys, G.W. (1989). Visual search and stimulus similarity. *Psychological Review*, **96**, 433-458.
- Durbin, R. and Mitchison, G. (1990). A dimension reduction framework for understanding cortical maps. *Nature*, **343**, 644-647.
- Evans, M. (1984). *The Art of the Beatles*. London: Muller, Blond and White Ltd.
- Fang, M. and Häusler, G. (1990). A shift invariant network utilizing the GRT. In R. Eckmiller, G. Hartmann, and G. Hauske (Eds.), *Parallel processing in neural systems and computers* (pp. 339-342). Dordrecht: Elsevier Science Publishers B.V. (North-Holland).
- Farah, M.J., Brunn, J.L., Wong, A.B, Wallace, M.A., and Carpenter, P.A. (1990). Frames of reference for allocating attention to space: evidence from the neglect syndrome. *Neuropsychologia*, **28**, 335-347.
- Farah, M.J., Wallace, M.A., and Vecera, S.P. (1993). "What" and "Where" in visual attention: evidence from the neglect syndrome. In I.H. Robertson and J.C. Marshall (Eds.), *Unilateral neglect: clinical and experimental studies* (pp. 123-137). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Feldman, J.A. and Ballard, D.H. (1982). Connectionist models and their properties. *Cognitive Science*, **6**, 205-254.
- Felleman, D.J. and Van Essen, D.C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, **1**, 1-47.

- Ferrera, V.P., Nealey, T.A., and Maunsell, J.H.R. (1992). Mixed parvocellular and magnocellular geniculate signals in visual area V4. *Nature*, **358**, 756-758.
- Fujita, I., Tanaka, K., Ito, M., and Cheng, K. (1992). Columns for visual features of objects in monkey inferotemporal cortex. *Nature*, **360**, 343-346.
- 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. (1989). Analysis of the process of visual pattern recognition by the Neocognitron. *Neural Networks*, **2**, 413-420.
- Georgopoulos, A.P., Taira, M., and Lukashin, A. (1993). Cognitive neurophysiology of the motor cortex. *Science*, **260**, 47-52.
- Gilchrist, A.L. (1979). The perception of surface blacks and whites. *Scientific American*, **240**, 112-124.
- Glauber, R.J. (1963). Time-dependent statistics of the Ising model. *Journal of Mathematical Physics*, **4**, 294-307.
- Glünder, H. (1990). Sigma-Pi networks for motion and invariant form analysis. In R. Eckmiller, G. Hartmann, and G. Hauske (Eds.), *Parallel processing in neural systems and computers* (pp. 117-120). Dordrecht: Elsevier Science Publishers B.V. (North-Holland).
- Goldberg, M.E. and Colby, C.L. (1989). In F. Boller and J. Grafman (Eds.), *Handbook of Neuropsychology, Vol.II*. (pp. 267-299). Dordrecht: Elsevier Science Publishers B.V. (North-Holland).
- Goodale, M.A. and Milner, A.D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, **15**, 20-25.
- Green, M. (1991). Visual search, visual streams, and visual architectures. *Perception & Psychophysics*, **50**, 388-403.
- Gross, C.G. and Mishkin, M. (1977). The neural basis of stimulus equivalence across retinal translation. In S. Harnad, R. Doty, J. Jaynes, L. Goldstein and G. Krouthamer (Eds.), *Lateralization in the nervous system*. New York: Academic Press.

- Gross, C.G., Rocha-Miranda, C.E., and Bender, D.B. (1972). Visual properties of neurons in inferotemporal cortex of the macaque. *Journal of Neurophysiology*, **35**, 96-111.
- Grossberg, S. (1973). Contour enhancement, short term memory, and constancies in reverberating neural networks. *Studies in Applied Mathematics*, **LII**, 213-257.
- Grossberg, S. (Ed.) (1982) *Studies of Mind and Brain: Neural Principles of Learning, Perception, Development, Cognition, and Motor Control*. Boston, MA: Reidel Press.
- Happel, B.L.M. and Murre, J.M.J. (1994). Design and evolution of modular neural network architectures. *Neural Networks*, **7**, (in press).
- Heemskerk, J.N.H., Hoekstra, J., Murre, J.M.J., Kemna, L.H.J.G., and Hudson, P.T.W. (1994). The BSP400: A Modular Neurocomputer. *Microprocessors and Microsystems*, **18** (2), 67-78.
- Heermann, D.W. (1990). *Computer simulation methods in theoretical physics (second edition)*. Berlin: Springer-Verlag.
- Heilman, K.M., Watson, E., and Goldberg, M.E. (1987). Attention: behavior and neural mechanisms. In V.B. Mountcastle, F. Plum, and S.R. Geiger (Eds.), *Handbook of physiology, Section I: The Nervous System, Volume V. Higher Functions of the Brain, Part 2* (pp. 461-481). Bethesda, MA: American Physiological Society.
- Henseler, H.J. (1993). *Connections, Neurons and Activation*. Ph.D. thesis, University of Limburg.
- Hertz, J., Krogh, A., and Palmer, R.G. (1991). *Introduction to the theory of neural computation*. Reading, MA: Addison-Wesley.
- Hillis, A.E. and Caramazza, A. (1991). Deficit to stimulus-centered, letter shape representations in a case of "unilateral neglect". *Neuropsychologia*, **29**, 1223-1240.
- Hinton, G.E. (1981). Shape representation in parallel systems. *Proceedings of the 7th IJCAI*, Vancouver, 1088-1096.
- Hinton, G.E. (1989). Deterministic Boltzmann learning performs steepest

- descent in weight-space. *Neural Computation*, **1**, 143-150.
- Hinton, G.E. and Lang, K.J. (1985). Shape recognition and illusory conjunctions. *Proceedings of the 9th IJCAI, Vol. 1* (pp. 252-259). Los Angeles.
- Hinton G.E. and Anderson, J.A. (Eds.) (1981). *Parallel models of associative memory*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Hopfield, J.J. (1982). Neural networks and physical systems with emergent collective computational properties. *Proceedings of the National Academy of Sciences U.S.A.*, **79**, 2554-2558.
- Hopfield, J.J. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences U.S.A.*, **81**, 3088-3092.
- Hopfield, J.J. (1990). The effectiveness of analogue 'neural network' hardware. *Network*, **1**, 27-40.
- Hopfield, J.J. and Tank, D.W. (1985). "Neural" computation of decisions in optimization problems. *Biological Cybernetics*, **52**, 141-152.
- Hubel, D.H. and Wiesel, T.N. (1974). Uniformity of monkey striate cortex: A parallel relationship between field size, scatter and magnification factor. *Journal of Comparative Neurology*, **158**, 295-306.
- Hurlbert, A. and Poggio, T. (1985). Spotlight on attention. *Trends in Neurosciences*, **8**, 309-311.
- Imry, Y. and Ma, S.K. (1975). *Physical Review Letters*, **35**, 1399.
- Ising, E. (1925). Beitrag zur theorie des ferromagnetismus. *Zeitschrift für Physik*, **31**, 253-258. °
- Jacobs, R.A. and Jordan, M.I. (1992). Computational consequences of a bias toward short connections. *Journal of Cognitive Neuroscience*, **4**, 323-336.
- Katz, B. (1966). *Nerve, muscle and synapse*. New York, NJ: McGraw-Hill.
- Kawasaki, K. (1972). Kinetics of Ising models. In C. Domb and M.S. Green (Eds.), *Phase transitions and critical phenomena. Volume 2* (pp. 443-499). New York, NY: Academic Press.

- Kinzel, W. and Schick, M. (1981). Phenomenological scaling approach to the triangular Ising antiferromagnet. *Physical Review*, B **23**, 3435-3441.
- Kirkpatrick, S., Gelatt, C.D. Jr., and Vecchi, M.P. (1983). Optimization by simulated annealing. *Science*, **220**, 671-680.
- Klein, R. (1988). Inhibitory tagging system facilitates visual search. *Nature*, **334**, 430-431.
- Knill, D.C. and Kersten, D. (1991). Apparent surface curvature affects lightness perception. *Nature*, **351**, 228-230.
- Koch, C. and Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, **4**, 219-277.
- Koenderink, J.J. (1984). Simultaneous order in nervous nets from a functional standpoint. *Biological Cybernetics*, **50**, 35-41.
- Koenderink, J.J. (1990). *Solid Shape*. Cambridge, MA: MIT Press.
- Koenderink, J.J. and Van Doorn, A.J. (1978). Visual detection of spatial contrast: influence of location in the visual field, target extent and illuminance level. *Biological Cybernetics*, **30**, 157-167.
- Koenderink, J.J. and Van Doorn, A.J. (1990). Receptive field taxonomy. In R. Eckmiller (Ed.), *Advanced Neural Computers* (pp. 295-301). Amsterdam: North-Holland.
- Kohonen, T. (1977). *Associative memory: A system theoretical approach*. New York, NJ: Springer-Verlag.
- Kosko, B. (1989). Unsupervised learning in noise. In *Proceedings of the International Joint Conference on Neural Networks, Volume I* (pp. 7-17). San Diego, CA: IEEE Neural Network Committee.
- Kosslyn, S.M., Flynn, R.A., Amsterdam, J.B., and Wang, G. (1990). Components of high-level vision: A cognitive neuroscience analysis and accounts of neurological syndromes. *Cognition*, **34**, 203-277.
- Kreyszig, E. (1970). *Introductory mathematical statistics. Principles and methods*. New York, NY: John Wiley and Sons.
- Kruglyak, L. and Bialek, W. (1993). Statistical mechanics for a network of

- spiking neurons. *Neural Computation*, **5**, 21-31.
- Kuffler, S.W., Nicholls, J.G., and Martin, A.R. (1984). *From Neuron to Brain (second edition)*. Sunderland, MA: Sinauer Associates.
- LaBerge, D. (1990). Thalamic and cortical mechanisms of attention suggested by recent positron emission tomographic experiments. *Journal of Cognitive Neuroscience*, **2**, 358-372.
- Lawrie, I.D. (1990). *A unified grand tour of theoretical physics*. New York, NY: Adam Hilger.
- Lenting J.H.J. (1992). On the virtue of unlimited parallelism in a neural network. In W. Krug and A. Lehmann (Eds.), *Proceedings of the 1992 European Simulation Symposium ESS'92* (pp. 278-284). San Diego, CA: Simulation Councils, Inc.
- Li, L., Miller, E.K., and Desimone, R. (1993). The representation of stimulus familiarity in anterior inferior temporal cortex. *Journal of Neurophysiology*, **69**, 1918-1929.
- Little, W.A. (1974). The existence of persistent states in the brain. *Mathematical Biosciences*, **19**, 101-120.
- Little, W.A. and Shaw, G.L. (1975). A statistical theory of short and long term memory. *Behavioral Biology*, **14**, 115-133.
- MacKay, D.M. (1985). The significance of 'feature sensitivity'. In D. Rose and V.G. Dobson (Eds.), *Models of the Visual Cortex* (pp. 47-53). New York, NY: John Wiley and Sons.
- Mackeben, M. and Nakayama, K. (1993). Express attentional shifts. *Vision Research*, **33**, 85-90.
- Mallat, S.G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **11**, 674-693.
- Marin-Padilla, M. (1987). The chandelier cell of the human visual cortex. A Golgi study. *Journal of Comparative Neurology*, **265**, 61-70.
- Marin-Padilla, M. (1990). The pyramidal cell and its local-circuit interneu-

- rons: A hypothetical unit of the mammalian cerebral cortex. *Journal of Cognitive Neuroscience*, **2**, 180-194.
- Maris, H.J. and Kadanoff, L.P. (1978). Teaching the renormalization group. *American Journal of Physics*, **66**, 652-657.
- McCulloch, W.S. and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, **5**, 115-133.
- McIlwain, J.T. (1976). Large receptive fields and spatial transformations in the visual system. In R. Porter (Ed.), *International Review of Physiology, Neurophysiology II*, **10**, 223-247.
- Merigan, W.H. and Maunsell, J.H.R. (1993). How parallel are the primate visual pathways? *Annual Review of Neuroscience*, **16**, 369-402.
- Mesulam, M.-M. (1990). Large-scale neurocognitive networks and distributed processing for attention, language and memory. *Annals of Neurology*, **28**, 597-611.
- Mignard, M. and Malpeli, J.G. (1991). Path of information flow through visual cortex. *Science*, **251**, 1249-1251.
- Milanesi, R. (1993). *Detecting salient regions in an image: from biological evidence to computer implementation*. Ph.D. thesis, University of Geneva.
- Minsky, M.L. and Papert, S.A. (1988). *Perceptrons: An introduction to computational geometry*. Cambridge, MA: MIT Press.
- Mishkin, M. and Appenzeller, T. (1987). The anatomy of memory. *Scientific American*, **256**(6), 62-71.
- Miyashita, Y. (1993). Inferior temporal cortex: where visual perception meets memory. *Annual Review of Neuroscience*, **16**, 245-263.
- Motter, B.C. (1988). Responses of visual cortical neurons during a focal attention task. *Society for Neuroscience Abstracts*, **14**, 10.
- 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.
- Mozer, M.C. (1991). *The perception of multiple objects. A connectionist ap-*

*proach*. Cambridge, MA: MIT Press.

Müller-Hartmann, E. and Zittartz, J. (1977). Interface free energy and transition temperature of the square-lattice Ising antiferromagnet at finite magnetic field. *Zeitschrift für Physik B*, **27**, 261-266.

Murre, J.M.J. (1992). *Learning and categorization in modular neural networks*. Hemel-Hempstead: Harvester Wheatsheaf.

Murre, J.M.J., Phaf, R.H., and Wolters, G. (1992). CALM: Categorizing and Learning Module. *Neural Networks*, **5**, 55-82.

Nakayama, K. (1990). The iconic bottleneck and the tenuous link between early visual processing and perception. In C. Blakemore (Ed.), *Vision: coding and efficiency* (pp. 411-422). Cambridge: Cambridge University Press.

Nelson, M.E. and Bower, J.M. (1990). Brain maps and parallel computers. *Trends in Neurosciences*, **13**, 403-408.

Nienhuis, B. and Nauenberg, M. (1975). Renormalization-group calculation for the equation of state of an Ising ferromagnet. *Physical Review*, **B 11**, 4152-4155.

Olshausen, B.A. (1994). *Neural routing circuits for forming invariant representations of visual objects*. Ph.D. thesis, California Institute of Technology.

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. *The Journal of Neuroscience*, **13**, 4700-4719.

Onsager, L. (1944). Crystal statistics: I. A two-dimensional model with an order-disorder transition. *Physical Review*, **65**, 117-149.

Perrett, D.I., Mistlin, A.J., and Chitty, A.J. (1987). Visual neurons responsive to faces. *Trends in Neurosciences*, **10**, 358-364.

Peterson C. and Anderson, J.R. (1987). A mean field theory learning algorithm for neural networks. *Complex Systems*, **1**, 995-1019.

Plaut, D.C. and Farah, M.J. (1990). Visual object representation: interpreting neurophysiological data within a computational framework. *Journal of Cognitive Neuroscience*, **2**, 320-343.

- Plischke, M. and Bergersen, B. (1989). *Equilibrium statistical mechanics*. Englewood Cliffs, NJ: Prentice-Hall.
- Posner, M.I. (1978). *Chronometric explorations of mind*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Posner, M.I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, **32**, 3-25.
- Posner, M.I. and Dehaene, S. (1994). Attentional networks. *Trends in Neurosciences*, **17**, 75-79.
- Posner, M.I. and Petersen, S.E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, **13**, 25-42.
- Posner, M.I. and Presti, D.E. (1987). Selective attention and cognitive control. *Trends in Neurosciences*, **10**, 13-17.
- Postma, E.O. (1993). Towards scalable neurocomputers. In H. Hüning, S. Neuhauser, M. Raus, and W. Ritschel (Eds.), *Proceedings of the Workshop on Neural Networks Aachener Beiträge zur Informatik, Band 3* (pp. 63-68). Aachen: Verlag der Augustinus Buchhandlung.
- Postma, E.O., Van den Herik, H.J., en Hudson, P.T.W. (1992a). Dynamic Selection through Gating Lattices. *Proceedings International Joint Conference on Neural Networks (IJCNN'92)*, Baltimore.
- Postma, E.O., Van den Herik, H.J., en Hudson, P.T.W. (1992b). A connectionist model of covert attention. *International Journal of Psychology*, **27** (3/4), p. 23.
- Postma, E.O., Van den Herik, H.J., and Hudson, P.T.W. (1993a). The gating lattice: a neural substrate for dynamic gating. In F.H. Eeckman and J.M. Bower (Eds.), *Computation and neural systems* (pp. 221-225). Boston, MA: Kluwer Academic Publishers.
- Postma, E.O., Van den Herik, H.J., and Hudson, P.T.W. (1993b). Activity-conserving dynamics for neural networks. In S. Gielen and B. Kappen (Eds.), *Proceedings of the International Conference on Artificial Neural Networks, ICANN'93* (pp. 539-544). London: Springer-Verlag.
- Postma, E.O., Van den Herik, H.J., and Hudson, P.T.W. (1994a). Attentional

- scanning. In A. Cohn (Ed.), *ECAI 94, 11th European Conference on Artificial Intelligence* (pp. 173-177). New York: John Wiley and Sons.
- Postma, E.O., Van den Herik, H.J., and Hudson, P.T.W. (1994b). Adaptive rewiring in the brain. In R. Cotterill (Ed.), *Proceedings of the Academy of Sciences, Denmark* (in press).
- Ratliff, F. (1965). *Mach bands: Quantitative studies on neural networks in the retina*. San Francisco: Holden-Day.
- Rumelhart, D.E., Hinton, G.E., and Williams R.J. (1986). Learning representations by back-propagating errors. *Nature*, **323**, 533-536.
- Rumelhart, D.E., McClelland, J.L. and the PDP research group (Eds.) (1986). *Parallel Distributed Processing. Explorations in the Microstructure of Cognition. Vol. 1. Foundations*. Cambridge, MA: MIT Press.
- Rybak, I.A., Shevtsova, N.A., Podladchikova, L.N., and Golovan, A.V. (1991). A visual cortex domain model and its use for visual information processing. *Neural Networks*, **4**, 3-13.
- Sanger, T.D. (1989). Optimal unsupervised learning in a single-layer linear feedforward neural network. *Neural Networks*, **2**, 459-473.
- Schick, M. Walker, J.S., and Wortis, M. (1976). Antiferromagnetic triangular Ising model. *Physics Letters*, **58A**, 479-480.
- Schick, M. Walker, J.S., and Wortis, M. (1977). Phase diagram of the triangular Ising model: renormalization-group calculations with application to absorbed monolayers. *Physical Review*, **B 16**, 2205-2219.
- Sejnowski, T.J. (1986). Open questions about computation in cerebral cortex. In J.L. McClelland, D.E. Rumelhart, and the PDP Research Group (Eds.), *Parallel Distributed Processing, Explorations in the microstructure of cognition, Vol. II, Psychological and biological models* (pp. 372-389). Cambridge, MA: MIT Press.
- Shepherd, G.M. (Ed.) (1990). *The synaptic organization of the brain (third edition)*. Oxford: Oxford University Press.
- Shepherd, G.M. and Koch, C. (1990). In G.M. Shepherd (Ed.), *The synaptic organization of the brain (third edition)* (pp. 3-31). Oxford: Oxford University

Press.

- Shulman, G.L. (1990). Relating attention to visual mechanisms. *Perception & Psychophysics*, **47**, 199-203.
- Simpson, P.K. (1990). *Artificial Neural Systems: Foundations, Paradigms, Applications, and Implementations*. New York, NY: Pergamon Press.
- Somogyi, P. and Martin, K.A.C. (1985). Cortical circuitry underlying inhibitory processes in cat area 17. In D. Rose and V.G. Dobson (Eds.), *Models of the Visual Cortex* (pp. 514-523). New York, NY: John Wiley and Sons.
- Stanley, H.E. (1971). *Introduction to phase transitions and critical phenomena*. Oxford: Oxford University Press.
- Stevens, C.F. (1979). The Neuron. In Scientific American (Ed.), *The Brain* (pp. 14-25). San Francisco, CA: Freeman.
- Stryker, M.P. (1990b). Summary: the brain in 1990. *Cold Spring Harbor Symposia on Quantitative Biology, Volume LV* (pp. 1049-1067). Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
- Sparks, D.L., Lee, C., and Rohrer, W.H. (1990). Population coding of the direction, amplitude, and velocity of saccadic eye movements by neurons in the superior colliculus. *Cold Spring Harbor Symposia on Quantitative Biology, Volume LV* (pp. 805-811). Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
- Squire, L.R. and Zola-Morgan, S. (1991). The medial temporal lobe memory system. *Science*, **253**, 1380-1386.
- Tanaka, K. (1993). Neuronal mechanisms of object recognition. *Science*, **262**, 685-688.
- Tanaka, K., Saito, H-A., Fukada, Y., and Moriya, M. (1991). Coding visual images of objects in the inferotemporal cortex of the Macaque monkey. *Journal of Neurophysiology*, **66**, 170-189.
- Thorpe, S.J. and Imbert, M. (1988). Biological constraints on connectionist modelling. In R. Pfeifer, Z. Schreter, F. Fogelman-Soulié and L. Steels (Eds.), *Connectionism in Perspective* (pp. 63-92). Dordrecht: Elsevier Science Publishers.

- Treisman, A.M. (1982). Perceptual grouping and attention in visual search for features and objects. *Journal of Experimental Psychology: Human Perception and Performance*, **8**, 194-214.
- Treisman, A.M. (1986). Features and objects in visual processing. *Scientific American*, **255**, 106-115.
- Treisman, A.M. (1989). Properties, parts, and objects. In K.R. Boff, L. Kaufmann, and J.P. Thomas (Eds.), *Handbook of perception and human performance Vol. 2*, Chapter 35. New York, NJ: John Wiley and Sons.
- Treisman, A. and Sato, S. (1990). Conjunction search revisited. *Journal of experimental psychology: Human perception and performance*, **16**, 459-478.
- Tsotsos, J.K. (1991). *Localizing stimuli in a sensory field using an inhibitory attention beam*. Technical Report, RBCV-TR-91-37, Department of Computer Science, University of Toronto.
- Ullman, S. (1984). Visual Routines. *Cognition*, **18**, 97-159.
- Ullman, S. (1986). Artificial Intelligence and the brain: computational studies of the visual system. *Annual Review of Neuroscience*, **9**, 1-26.
- Vallar, G. (1993). The anatomical basis of spatial hemineglect in humans. In I.H. Robertson and J.C. Marshall (Eds.), *Unilateral neglect: clinical and experimental studies* (pp. 27-59). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Van der Heijden, A.H.C. (1992). *Selective attention in vision*. London: Routledge.
- Van der Heijden, A.H.C. (1993). The role of position in object selection in vision. *Psychological Research*, **56**, 44-58.
- Van Essen, D.C. and Anderson, C.H. (1990). Information processing strategies and pathways in the primate retina and visual cortex. In S.F. Zornetzer, J.L. Davis, and C. Lau (Eds.), *An introduction to neural and electronic networks* (pp. 43-72). San Diego, CA: Academic Press.
- Van Essen, D.C. and Maunsell, J.H.R. (1983). Hierarchical organization and functional streams in the visual cortex. *Trends in Neurosciences*, **6**, 370-375.
- Von der Malsburg, C. (1988). Pattern recognition by labeled graph matching.

*Neural Networks*, 1, 141-148.

Von der Malsburg, C. (1990). Considerations for a visual architecture. In R. Eckmiller (Ed.), *Advanced Neural Computers* (pp. 303-312). Amsterdam: North-Holland.

Walley, R.E. and Weiden, T.D. (1973). Lateral inhibition and cognitive masking: a neuropsychological theory of attention. *Psychological Review*, 80, 284-302.

Watson, A.B. and Ahumada, A.J. (1989). A hexagonal orthogonal-oriented pyramid as a model of image representation in visual cortex. *IEEE Transaction on Biomedical Engineering*, 36, 97-106.

Wilson, K.G. (1979). Problems in physics with many scales of length. *Scientific American*, 241, 140-157.

Wilson, H.R., Levi, D., Maffei, L., Rovamo, J., and De Valois, R. (1990). The perception of form: retina to striate cortex. In L. Spillmann and J. S. Werner (Eds.), *Visual perception: the neurophysiological foundations* (pp. 231-272). London: Academic Press.

Wise, S.P. and Desimone, R. (1988). Behavioral neurophysiology: insights into seeing and grasping. *Science*, 242, 736-741.

Witkin, A.P. (1983). Scale-space filtering. *Proceedings of IJCAI-83*, Karlsruhe, West-Germany, 1019-1022.

Witkin, A.P., Terzopoulos, D., and Kass, M. (1986). Signal matching through scale space. *International Journal of Computer Vision*, 1, 133-144.

Wurtz, R.H., Goldberg, M.E., and Robinson, D.L. (1982). Brain mechanisms of visual attention. *Scientific American*, 246, 100-107.

Yantis, S. and Johnston, J.C. (1990). On the locus of visual selection: Evidence from focused attention tasks. *Journal of Experimental Psychology: Human perception and Performance*, 16, 135-149.

Yarbus, A.L. (1967). *Eye movements and vision*. New York: Plenum Press.

Yeomans, J.M. (1992). *Statistical mechanics of phase transitions*. Oxford: Clarendon Press.

- Yeshurun, Y. and Schwartz, E.L. (1989). Shape description with a space-variant sensor: algorithms for scan-path, fusion, and convergence over multiple scans. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **11**, 1217-1222.
- Young, M.P. (1992). Objective analysis of the topological organization of the primate cortical visual system. *Nature*, **358**, 152-155.
- Zeki, S. (1993). *A vision of the brain*. Oxford: Blackwell Scientific Publications.
- Zipser, D. and Andersen, R.A. (1988). A back propagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature*, **331**, 679-684.
- Zohary, E. and Hochstein, S. (1989). How serial is serial processing in vision? *Perception*, **18**, 191-200.
- Zucker, R.S. (1989). Short-term synaptic plasticity. *Annual Review of Neuroscience*, **12**, 13-31.



# Summary

While fixating the eyes on one location of the visual scene one is still able to identify objects at other locations. The process underlying this ability is called covert attention. The covert-attention process can be likened to a searchlight that selects part of a visual scene by illuminating it. The searchlight metaphor of covert attention is at the base of the SCAN model described in this thesis.

In Chapter 1, the problem statement of this study reads: *is it possible to construct a working model of covert attention as described by psychology which satisfies the constraints set by biology?*

A multidisciplinary approach is pursued in order to answer this question. The computer-science approach is central in that it emphasizes the formulation of a *working* model. In addition, psychological knowledge about the functional aspects of the attentional searchlight are combined with biological knowledge about its realisation in neural structures. Finally, models and insights from statistical physics are used in formulating and studying the behaviour of the model.

Scalability to problems of a realistic size is an important requirement for neural-network models. Minsky and Papert had already pointed out, in their book *Perceptrons*, that the performance of neural networks often remains limited to small-sized toy problems. Given that a working model has to be applicable to problems of a realistic size, it is important to take scalability into account when formulating the architecture. Biological knowledge about the neural realisation of brain function plays a pivotal role in this process. Characteristics of biological hardware may be formulated as constraints for scalable neural machines.

Chapter 2 studies the biological basis of visual perception and visual attention.

At the level of neurons and their interconnections three characteristics appear to be of importance for scalability: (1) the stochasticity of neurons, (2) the sparseness of connections, (3) and the profligacy of neurons. At the level of the visual system, two characteristics of vision are: (4) the hierarchical structure of the visual system, and (5) the invariant representation of visual objects in the identification part of the visual system. The characteristics (1) to (3) are used in the definition of the neural lattice in Chapter 3 and in defining the gating lattice in Chapter 4. The characteristics (4) and (5) are used in the SCAN model described in Chapter 5.

Chapter 3 studies the effect of connection density on neural-network performance. Because the scalability of a neural network decreases with increasing connectivity, it is important to keep the number and length of connections small. A model called the neural lattice is introduced and used to study the effect of connection density. The neural lattice is a parallel distributed neural network that detects an above-threshold input signal and is equivalent to the ferromagnetic Ising lattice from statistical mechanics. Although the neural lattice is not applied in the SCAN model, it is strongly related to the gating lattice, the building block of SCAN that is described in Chapter 4.

The structure and dynamics of the neural lattice obey the characteristics (1) to (3) of Chapter 2: the lattice consists of a large number of stochastic elements that communicate through nearest-neighbour connections. Varying the lattice dimensionality leads to three distinct models that differ in their connection density: the one-, two-, and three-dimensional neural lattices. The behaviour of the three lattices is compared with the behaviour of a lattice in which all  $N$  elements ( $N \gg 6$ ) are interconnected: the " $N$ -dimensional" neural lattice.

Through Monte-Carlo simulations the effect of lattice dimensionality on the robustness and flexibility of above-threshold detection are assessed. The robustness of a neural lattice is defined as: the degree to which the quality of detection is maintained with increasing noise in the input signal. The flexibility of a neural lattice is its independency from the initial state.

Simulation results show that the robustness increases with lattice dimensionality, whereas flexibility decreases with lattice dimensionality. This leads to the conclusion that optimal robustness in combination with optimal flexibility can be obtained for a two- or three-dimensional lattice.

Chapter 4 introduces the gating lattice, that serves as the building block for

SCAN in Chapter 5. The gating lattice is equivalent with the antiferromagnetic Ising lattice from statistical mechanics and is, therefore, related to the neural lattice. The conclusions about the neural lattice also apply, therefore, to the gating lattice. As a result the gating lattice may be defined as a two- or three- dimensional lattice. The gating lattice is defined as a two-dimensional lattice in order to be able to use existing knowledge of the equivalent antiferromagnetic Ising lattice.

The gating lattice is a parallel distributed switch with three switching directions. It is capable of selecting one of three patterns without affecting the structure of the pattern. The selection of a switching direction is based on the relative magnitudes of three control signals. The distributed competition in the gating lattice amplifies small differences among the control signals. As a result, the sublattice with the largest control signal opens. When the control signals are somehow derived from the pattern to be gated (e.g., its saliency), a mechanism is obtained in which a pattern (via its saliency) itself may determine whether it is gated or not.

A formulation of the gating lattice as an Ising model enables an exact computation of a lattice composed of nine elements. Monte-Carlo simulations are used for the study of larger lattices. Combining these results with knowledge of the critical behaviour of Ising models reveals that the optimal gating quality is achieved when the intrinsic-noise value is equal to a critical value. In contrast, the gating speed decreases when approaching the critical value. This leads to the conclusion that an intrinsic-noise value near the critical value yields the optimal combination of gating quality and gating speed.

Chapter 5 introduces the SCAN (Signal Channelling Attentional Network) model. Apart from the three first characteristics, SCAN incorporates the characteristics (4) and (5): it is a hierarchical network that yields a translation-invariant representation. SCAN achieves invariance by selecting contiguous subpatterns (object-patterns) from an input pattern. The gating network, a hierarchical structure of gating lattices, channels subpatterns towards a centralized neural-network classifier. The classifier used in SCAN is based on the ART network of Carpenter and Grossberg. The main feature of the ART network that is of relevance for our purposes is that it generates an expectation pattern when confronted with an input pattern. In SCAN, the expectation pattern is compared with all subpatterns of the input pattern. The degree to which the expectation and a subpattern match yields a control signal; the bet-

ter the match, the larger the control signal. Because a control signal controls the sublattice that gates the accompanying subpattern, the subpattern that has the best match with the expectation is channelled towards the classifier.

Simulations with random control signals show how the speed and quality of gating in SCAN behave as a function of the intrinsic-noise value and the number of levels in the gating tree. In order to enable the selection of patterns with an intrinsic two-dimensional structure (e.g., visual images), the gating lattice is defined as a fractal structure (see the cover). At each node in this structure, a gating lattice selects one of three subpatterns from a two-dimensional plane. When presented with a realistic image of  $3^{10}$  pixels and the associated control signals, SCAN is capable of channelling the target pattern with a reasonable overall gating quality. Perfect gating quality is obtained by varying the intrinsic noise in time.

It is concluded that the SCAN model constitutes a scalable realisation of the attentional searchlight. The method used in SCAN to obtain translation invariance differs radically from standard neural-network approaches. Generally, input images are decomposed in invariant features. Such a decomposition leads to the problem of how features belonging to a single object should be glued together, the so-called binding problem. In SCAN the binding problem does not arise because no decomposition takes place. Moreover, SCAN combines translation invariance with selection enabling and active scanning of visual images.

In Chapter 6, the structure and functioning of the SCAN model are evaluated by comparing it to the biological example. The plausibility of SCAN as a biological model is discussed. Two inconsistencies lead to suggestions for extending and improving the model.

Chapter 7 concludes that in SCAN a model is realised that is in agreement with the requirements set out in the problem statement by combining biology (the characteristics of a scalable architecture), psychology (the functionality of the attentional searchlight), statistical physics (the behaviour of a large number of locally interconnected stochastic elements), and computer science (the scalability of parallel distributed neurocomputers).

# Samenvatting

Wanneer de ogen gefixeerd worden op één lokatie in de visuele omgeving is het niettemin mogelijk objecten op andere lokaties te identificeren. Het proces dat hiervoor verantwoordelijk is wordt selectieve aandacht (*covert attention*) genoemd. Het visuele aandachtsproces kan tot op zekere hoogte vergeleken worden met een zoeklicht dat een deel van de omgeving selecteert door het te belichten. De notie van selectieve aandacht als zoeklicht ligt ten grondslag aan het in dit proefschrift beschreven SCAN-model.

In hoofdstuk 1 wordt de probleemstelling van deze studie geformuleerd. De probleemstelling luidt: *is het mogelijk een werkend model van visuele aandacht te formuleren dat in overeenstemming is met de functionaliteit van het aandachtszoeklicht zoals beschreven in de psychologie en dat voldoet aan de voorwaarden voor een succesvolle fysieke realisatie zoals beschreven in de biologie?*

Om deze vraag te beantwoorden is voor een multidisciplinaire aanpak gekozen. Centraal in deze aanpak staat de informatica-benadering waarin de nadruk valt op het formuleren van een *werkend* model. Daarnaast wordt psychologische kennis van de functionaliteit van het aandachtszoeklicht gecombineerd met biologische kennis over zijn realisatie in neurale structuren. Tenslotte worden modellen en inzichten uit de statistische fysica gebruikt voor de beschrijving en bestudering van het model.

Een belangrijk vereiste van een neuraal-netwerk model is dat het schaalbaar is naar problemen van een realistische omvang. Zoals reeds onderstreept door Minsky en Papert in hun boek *Perceptrons*, blijft het bereik van neurale netwerken vaak beperkt tot problemen van een kleine omvang. Aangezien een werkend model toepasbaar moet zijn op problemen van een realistische

omvang, is het van belang om bij het specificeren van de architectuur rekening te houden met de schaalbaarheid. Biologische kennis over de neurale realisatie van hersenfuncties speelt hierbij een belangrijke rol. Karakteristieke eigenschappen van de biologische hardware kunnen worden geformuleerd als voorwaarden voor schaalbare neurale machines.

In hoofdstuk 2 wordt de biologische basis van de visuele waarneming en de visuele aandacht bestudeerd. Op het niveau van neuronen en hun verbindingen blijken drie karakteristieke eigenschappen van belang te zijn voor schaalbaarheid: (1) de stochasticiteit van neuronen, (2) de zuinigheid van verbindingen, en (3) de veelheid van neuronen. Op het niveau van het visueel systeem komen twee karakteristieke eigenschappen naar voren die specifiek van toepassing zijn op de visuele waarneming: (4) de hiërarchische structuur van het visueel systeem en (5) de invariantie van object-representaties in het deel van het visueel systeem dat verantwoordelijk is voor object-identificatie. De eigenschappen (1) tot en met (3) worden gebruikt in de definitie van het neurale rooster in hoofdstuk 3 en in de definitie van het schakelrooster in hoofdstuk 4. De eigenschappen (4) en (5) worden gebruikt in het in hoofdstuk 5 gedefinieerde SCAN-model.

In hoofdstuk 3 wordt een studie verricht naar het effect van de mate van connectiviteit op het gedrag van een neurale netwerk. Aangezien de schaalbaarheid van een neurale netwerk afneemt met toenemende connectiviteit, is het van belang het de lengte van verbindingen en het aantal verbindingen klein te houden. Uitgangspunt voor de in dit hoofdstuk beschreven studie vormt het neurale rooster (*neural lattice*), een parallel gedistribueerd neurale netwerk dat een bovendrempelig invoersignaal detecteert. Het neurale rooster is gelijkwaardig aan het ferromagnetisch Ising rooster uit de statistische fysica. Hoewel het neurale rooster niet toegepast wordt in het SCAN-model, is het wel sterk verwant aan het in hoofdstuk 4 beschreven schakelrooster dat als bouwsteen van SCAN fungeert.

De structuur en dynamica van het neurale rooster voldoen aan de karakteristieken (1) tot en met (3) uit hoofdstuk 2: het rooster is opgebouwd uit een groot aantal stochastische elementen die door naaste-buur verbindingen communiceren. Variatie van de dimensionaliteit van het rooster leidt tot drie verschillende modellen die zich onderscheiden in hun mate van connectiviteit: het één-, twee-, en drie-dimensionale neurale rooster. Het gedrag van de drie roosters worden vergeleken met het gedrag van een rooster waarin alle  $N$  ele-

menten ( $N \gg 6$ ) met elkaar verbonden zijn: het " $N$ -dimensionale" rooster.

Door middel van Monte-Carlo simulaties wordt het effect van de dimensionaliteit van het rooster op de robuustheid en flexibiliteit van detectie bepaald. Onder de robuustheid van het neurale rooster wordt verstaan: de mate waarin de detectiekwaliteit behouden blijft bij toenemende ruis in het invoersignaal. De flexibiliteit van het neurale rooster is de mate waarin de detectiekwaliteit onafhankelijk is van de begintoestand van het rooster.

Uit de simulatieresultaten blijkt dat de robuustheid toeneemt met de dimensionaliteit van het rooster en de flexibiliteit afneemt met de dimensionaliteit van het rooster. Dit leidt tot de conclusie dat voor optimale robuustheid in combinatie met optimale flexibiliteit, een twee- of drie-dimensionaal rooster de voorkeur verdient boven het één- of  $N$ -dimensionaal rooster.

In hoofdstuk 4 wordt het schakelrooster (*gating lattice*) ingevoerd. Het schakelrooster is gelijkwaardig met het antiferromagnetisch Ising rooster en daarmee verwant aan het neurale rooster. De conclusies over het neurale rooster zijn derhalve tevens van toepassing op het schakelrooster. Als uitvloeisel van de conclusie van hoofdstuk 3, kan het schakelrooster gedefinieerd worden als een twee- of drie-dimensionaal rooster. Het schakelrooster wordt gedefinieerd als een twee-dimensionaal rooster zodat bestaande kennis van het gelijkwaardig antiferromagnetisch Ising rooster kan worden gebruikt.

Het schakelrooster vormt een parallel gedistribueerde drie-standen schakelaar die één van drie patronen selecteert waarbij de structuur van het patroon intact blijft. De keuze van de schakelrichting wordt bepaald door drie stuursignalen. De gedistribueerde competitie in het schakelrooster versterkt kleine verschillen tussen de stuursignalen zodat het rooster met het grootste stuursignaal zich opent. Wanneer de stuursignalen op een of andere wijze afgeleid zijn van de door te laten patronen (het stuursignaal drukt b.v. de opvallendheid van een patroon uit), ontstaat er een mechanisme waarbij het patroon zelf (via zijn opvallendheid) bepalend is voor het al of niet doorlaten ervan.

Een formulering van het schakelrooster als het Ising model maakt een exacte berekening mogelijk van een rooster bestaande uit negen elementen. Voor de bestudering van grotere roosters wordt gebruik gemaakt van Monte-Carlo simulaties. In combinatie met kennis over het kritische gedrag van het Ising model laten deze simulaties zien hoe de schakelkwaliteit en schakelsnelheid afhangen van de intrinsieke ruiswaarde. Het blijkt dat optimale schakel-

kwiteit bereikt wordt wanneer de intrinsieke ruis een waarde heeft die gelijk is aan de kritische waarde. De schakelsnelheid neemt daarentegen af naarmate de intrinsieke ruis de kritische waarde dichter nadert. Dit leidt tot de conclusie dat door een waarde van de intrinsieke ruis nabij de kritische waarde een optimale combinatie van schakelkwiteit en schakelsnelheid verkregen wordt. Het schakelrooster vormt de bouwsteen van het in hoofdstuk 5 beschreven SCAN-model.

In hoofdstuk 5 wordt het SCAN (Signal Channelling Attentional Network) model geïntroduceerd. In het SCAN-model zijn, naast de eerste drie karakteristieke eigenschappen, de eigenschappen (4) en (5) verwerkt: het vormt een hiërarchisch netwerk dat leidt tot een translatie invariante representatie. SCAN bereikt invariantie door aaneengesloten deelpatronen (objectpatronen) te selecteren uit een invoerpatroon. Dit geschiedt door middel van een schakelnetwerk (*gating network*), een hiërarchische structuur van schakelroosters, die deelpatronen in baan brengt naar een gecentraliseerde neuraal-netwerk patroonclassificator. De in SCAN gebruikte classificator is gebaseerd op het ART-netwerk van Carpenter en Grossberg. De voor deze toepassing voorname eigenschap van het ART-netwerk is dat het een verwachting genereert voor een voor classificatie aangeboden invoerpatroon. In SCAN wordt het door de classificator gegenereerde verwachtingspatroon vergeleken met alle deelpatronen in het invoerpatroon. De mate van overeenkomst van een deelpatroon met het verwachtingspatroon levert de waarde van het stuursignaal: hoe beter de overeenkomst, des te hoger het stuursignaal. Aangezien dit stuursignaal gekoppeld is aan het subrooster dat het betreffend deelpatroon als invoer heeft, wordt het deelpatroon dat het meest met de verwachting overeenkomt het invoerpatroon van de classificator.

Simulaties met willekeurige stuursignalen laten zien hoe SCAN zich gedraagt, afhankelijk van de waarde van de intrinsieke ruis, en hoe de schakelkwiteit en schakelsnelheid afhangen van het aantal lagen in het schakelnetwerk. Om selectie uit invoerpatronen met een intrinsiek twee-dimensionale structuur (b.v. een foto) mogelijk te maken, wordt het schakelnetwerk gedefinieerd als een fractale structuur (zie omslag). Op ieder knooppunt van deze structuur selecteert een schakelrooster één van drie deelpatronen uit het twee-dimensionale vlak. Gegeven een realistisch visueel beeld bestaande uit  $3^{10}$  beeldelementen en de bijbehorende stuursignalen, is SCAN in staat om het deelpatroon met een redelijke schakelkwiteit in baan te brengen naar de classificator. Een perfecte schakelkwiteit wordt bereikt door de intrinsieke ruis in de tijd te

variëren.

De conclusie van hoofdstuk 5 is dat het SCAN-model een schaalbare realisatie is van het aandachtszoeklicht. De in SCAN gehanteerde methode om translatie-invariante patroonherkenning te bereiken wijkt radicaal af van de gangbare neurale-netwerk methoden. Doorgaans wordt namelijk uitgegaan van een ontbinding van het beeld in invariante kenmerken. Een dergelijke ontbinding leidt tot het bindingsprobleem (*binding problem*), oftewel, het probleem hoe de kenmerken behorend bij een object verbonden worden tot een invariante objectrepresentatie. In SCAN doet het bindingsprobleem zich niet voor, omdat er geen ontbinding plaatsvindt. Bovendien combineert SCAN translatie-invariantie met selectie waardoor een actieve *scanning* van visuele beelden mogelijk wordt.

In hoofdstuk 6 worden de structuur en het gedrag van het SCAN-model geëvalueerd door een vergelijking met het biologische voorbeeld. De aanvaardbaarheid van SCAN als biologisch model wordt besproken. Twee inconsistenties leiden tot suggesties voor de wijze waarop het SCAN-model uitgebreid en verbeterd kan worden.

In hoofdstuk 7 wordt geconcludeerd dat in SCAN een model verwezenlijkt is dat door de combinatie van biologie (de karakteristieke kenmerken van een schaalbare architectuur), psychologie (de functionaliteit van het aandachtszoeklicht), statistische fysica (het gedrag van een groot aantal stochastische lokaal communicerende elementen), en informatica (de schaalbaarheid van parallel gedistribueerde neurocomputers) in overeenstemming is met de in de probleemstelling genoemde eisen.



# Curriculum Vitae

Eric Onno Postma werd op 28 december 1961 te Nijmegen geboren. In 1982 behaalde hij het diploma Atheneum B aan de Nijmeegse Scholen Gemeenschap. In datzelfde jaar begon hij een studie Natuurkunde aan de Katholieke Universiteit te Nijmegen (KUN). Deze studie bleek echter niet geheel in overeenstemming met zijn interesses. Daarom werd in 1983 gestart met de studie psychologie, in het bijzonder de Cognitiewetenschap (eveneens aan de KUN). Deze studie werd in 1989 afgesloten met een doctoraal examen. Van oktober 1989 tot oktober 1993 was hij als Assistent In Opleiding verbonden aan de vakgroep Informatica van de Rijksuniversiteit Limburg te Maastricht. Tijdens deze aanstelling verrichtte hij onderzoek onder begeleiding van Prof. dr H.J. van den Herik en Prof. dr P.T.W. Hudson. Het onderzoek resulteerde in een aantal publicaties en dit proefschrift. Momenteel is Eric Postma als Universitair Docent verbonden aan genoemde vakgroep.

