

1995

# Towards the Development of a Model of Vision: An Investigation into the Architectures and Mechanisms of Visual Perception

TROUP, LUCY JANE

<http://hdl.handle.net/10026.1/1924>

---

<http://dx.doi.org/10.24382/1339>

University of Plymouth

---

*All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.*

**Towards the Development of a Model of Vision:  
An Investigation into the Architectures and Mechanisms of  
Visual Perception.**

by

**LUCY JANE TROUP**

A thesis submitted to the University of Plymouth  
in partial fulfilment for the degree of

**DOCTOR OF PHILOSOPHY**

School of Computing  
Faculty of Technology

July 1995

UNIVE	U.S. PLYMOUTH
Item No.	900 2442638
Date	09 OCT 1995
Class No.	T 152-14 TR0
Contl. No.	X703136894
LIBR. SERVICES	

90 0244263 8



REFERENCE ONLY

# **DEDICATION**

I would like to dedicate this work to my parents;

Barty Hotchkiss

and

Audrey Hotchkiss

*"You cannot teach a man anything;*

*you can only help him to find it within himself."*

Galileo Galilei.

# **ABSTRACT**

## **Towards the Development of a Model of Vision: An Investigation into the Architectures and Mechanisms of Visual Perception.**

by

**LUCY JANE TROUP**

A conceptual model of visual perception has been developed using a multi-disciplinary approach which combines both top-down and bottom-up descriptions of vision. Top-down psychological theories of visual perception have been investigated resulting in the development of a theory of perception which combines the best of existing accounts. Perception is defined in terms of a combination of “data driven” and “concept driven” explanations. Bottom-up neurophysiological descriptions have also been investigated to provide possible descriptions of structure and function for the development of a conceptual model based upon the theory. An attempt is made to provide a “complete” account of visual perception through the development of both the theory and conceptual model. Further it is envisaged that the development of such a model will provide new insight into the development of artificial vision systems and new algorithms for perceptual function in such systems.

# **CONTENTS**

Copyright Statement .....	1
Title Page .....	2
Dedication.....	3
Abstract .....	4
List of contents.....	5
List of illustrations .....	7
Nomenclature .....	9
Acknowledgement.....	10
Authors declaration.....	11
 CHAPTER 1. Introduction: An Overview Of The Development Of A Model Of Visual Perception.....	 12
INTRODUCTION .....	12
A THEORETICAL FRAMEWORK FOR MODELLING VISION. ....	 21
INTRODUCTION TO A MODEL OF VISUAL PERCEPTION..	25
SYNOPSIS OF THE FOLLOWING 5 CHAPTERS.....	34
 CHAPTER 2. The visual system: An overview of the neurophysiology of visual perception. ....	 36
UNDERSTANDING VISION. ....	36
INTRODUCTION TO THE NEUROPHYSIOLOGY OF THE VISUAL SYSTEM. ....	 43
 CHAPTER 3. Theories of visual perception: Neurophysiological and psychological theories of vision. ....	 74
INTRODUCTION .....	74
NEUROPHYSIOLOGICAL THEORIES OF VISION.....	76
PSYCHOLOGICAL THEORIES OF VISION. ....	103
 CHAPTER 4. A model of visual perception: An outline of a model of perception from a top-down and bottom-up approach. ....	 115
A THEORY OF VISUAL PERCEPTION. ....	115
A CONCEPTUAL MODEL OF VISUAL PERCEPTION.....	119
 CHAPTER 5. Support for theory and model: additional support for both the theory and model .....	 144
INTRODUCTION .....	144
VISION AS BOTH A TOP-DOWN AND BOTTOM-UP PROCESS. ....	 145
SUPPORT FOR MECHANISMS UNDERLYING PERCEPTION. ....	 156

<b>CHAPTER 6. Summary: a discussion of the achievements of the model and suggestions for further work. ....</b>	<b>173</b>
<b>INTRODUCTION. ....</b>	<b>173</b>
<b>HAS THE THESIS ACHIEVED WHAT IT SET OUT TO ACHIEVE?.....</b>	<b>174</b>
<b>A MULTI-DISCIPLINARY APPROACH.....</b>	<b>180</b>
<b>HOW THE CONCEPTUAL MODEL IS DIFFERENT FROM OTHER MODELS .....</b>	<b>183</b>
<b>SPECIFIC IMPLICATIONS OF THE CONCEPTUAL MODEL.....</b>	<b>190</b>
<b>GENERAL IMPLICATIONS OF THE CONCEPTUAL MODEL....</b>	<b>201</b>
<b>SUGGESTIONS FOR FUTURE WORK.....</b>	<b>207</b>
<b>CONCLUSIONS. ....</b>	<b>213</b>
<b>Glossary.....</b>	<b>214</b>
<b>List of references.....</b>	<b>215</b>

# LIST OF ILLUSTRATIONS

<b>Figure 1.</b> A schematic diagram of the visual processing areas of the brain including the primary and secondary visual pathways. ....	41
<b>Figure 2.</b> The basic structure of the retina. ....	44
<b>Figure 3.</b> The receptive fields of retinal ganglion cells. ....	45
<b>Figure 4.</b> “On centre” and “off centre” retinal ganglions. ....	46
<b>Figure 5.</b> The primary visual pathway from the retinal to the dLGN showing the X, Y and W ganglion classification. ....	48
<b>Figure 6.</b> A schematic diagram showing the basic structure of the striate cortex (V1). ....	56
<b>Figure 7.</b> A schematic diagram of the functional architecture of the striate cortex (V1). ....	59
<b>Figure 8.</b> A schematic diagram of the extrastriate and association areas of the visual cortex. ....	63
<b>Figure 9.</b> The secondary, reflexive visual pathways. ....	70
<b>Figure 10.</b> Single cell feature detection theory. ....	79
<b>Figure 11.</b> A schematic diagram illustrating Maunsell and Newsome’s two stream theory of visual processing. ....	86
<b>Figure 12.</b> A schematic diagram illustrating Zeki’s four stream theory of visual processing. ....	89
<b>Figure 13.</b> A simple outline of the structure of the model. ....	121
<b>Figure 14.</b> The HOW and WHAT pathways. ....	123
<b>Figure 15.</b> A simple outline of feedforward functional processing. .	125
<b>Figure 16.</b> A simple outline of feedback functional processing. ....	128
<b>Figure 17.</b> A summary of the model so far. ....	130
<b>Figure 18.</b> A simple representation of visual processing in the model. ....	134



**Figure 19.** A schematic diagram of a possible implementation of the model. .... 140

**Figure 20.** Examples of visual illusions: The goblet/two faces illusion and Kanizas triangle. .... 149

**Figure 21.** A summary of Gray and Singers findings..... 158

**Figure 22.** Gray and Singer’s 4 experiments. .... 160

**Figure 23.** Oscillatory activity between groups showing no correlation. .... 163

# **NOMENCLATURE**

AI.....	Artificial Intelligence.
AIT .....	Anterior Inferotemporal area.
ANN .....	Artificial Neural Network.
ART .....	Adaptive Resonance Theory.
GLI .....	Grey Level Image.
IT.....	Inferior Temporal area.
LFP .....	Local Field Potential.
LGN .....	Lateral Geniculate Nucleus.
dLGN.....	Dorsal Lateral Geniculate Nucleus.
vLGN.....	Ventral Lateral Geniculate Nucleus.
MST .....	Medial Superior Temporal area.
MT .....	Medial Temporal area.
MTL .....	Medial Temporal Lobe.
MRI.....	Magnetic Resonance Imaging.
MUA .....	Multi-Unit Activity.
NOT .....	Nucleus Optic Tract.
ONPT .....	Olivary Nucleus of Pre-tectum.
PDP.....	Parallel Distributed Processing.
PET .....	Positron Emission Topography.
PFC .....	Pre Frontal Cortex.
PIT.....	Posterior Inferotemporal area.
PMLS .....	Posteromedial Lateral Suprasylvian area.
PN.....	Pulvinar Nucleus.
SC.....	Superior Colliculus.
SCFD.....	The Single Cell Feature Detection theory.
SCN.....	Suprachiasmatic Nucleus.
SNN .....	Standard Neural Network.
STG.....	Superior Temporal Gyrus.
STPp.....	Posterior Superior Poly Sensory areas.
STPa.....	Anterior Superior Poly Sensory areas.
dTAON .....	Dorsal Terminal Accessory Optic Nucleus.
lTAON.....	Lateral Terminal Accessory Optic Nucleus.
mTAON .....	Medial Terminal Accessory Optic Nucleus.
V1 .....	Primary visual cortex.
V2 .....	Visual processing area 2 (Part of the extra striate areas).
V3 .....	Visual processing area 3 (Part of the extra striate areas).
V3a .....	Visual processing area 3a (Part of the extra striate areas).
V4 .....	Visual processing area 4 (Part of the extra striate areas).
V5 (MT).....	Visual processing area 5 (Also known as MT -The Medial Temporal area).

# **ACKNOWLEDGEMENTS**

Firstly I would like to thank my director of studies, Professor Mike Denham for all his support and encouragement over the last three years. I would also like to thank my second supervisor Dr. Chris Hindle. In addition I would like to thank the School of Computing for the financial support that enabled me to carry out my research. I would also like to express my thanks to all the members of the Neurodynamics Research Group, especially Faye O' Shaughnessy and Dr Guido Bugmann who not only provided me with a platform for discussion but also kept me sane during the summer of 1994. Special thanks go to Faye, for providing me with invaluable insight into the possible role of "memory" in my work and to Quentin Hunt for photocopying services beyond the call of duty.

Finally I would like to thank all my family, especially Simon, Toby, Esther, and John, and my friends, especially Dr. Rob Simpson. Lastly, but by no means least I would like to thank my husband Tim, quite simply for being there.

# **AUTHOR'S DECLARATION**

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other University award.

## **Publications:**

**DENHAM, M. J., PATEL, S., TROUP, L. J. & NORMAN, M. 1991.** Oscillatory Neural Networks and their Application to Sensory-Motor Co-ordination and Control in Adaptive Robots. *In Proceedings of the IEE 2nd International Conference on Artificial Neural Networks. IEE Publishing, London, pp. 328-332.*

**DENHAM, M. J. & TROUP, L. J. 1992.** Coherent Oscillatory Responses as a Feature Linking Mechanism in an Artificial Neural Network Model of Visual Perception. *In Proceedings of NEURO NIMES: The 5th International Conference on Neural Networks and their Applications, Nimes, France. Ec2, Paris, pp. 505-516.*

## **Presentations and Conferences Attended:**

IEE 2nd International Conference on Artificial Neural Networks.  
Bournemouth International Centre, U.K.  
November 18-20th, 1991.

Presentation to the Mathematics Group, Centre for Neural Networks, Kings College, London.  
"Beyond Bramblescape: Towards a Biologically Plausible Model Of Visual Perception".  
Wednesday 24th February, 1993.

Signed.....*Lucy Troup*.....  
Date.....*31<sup>st</sup> July 1995*.....

# **CHAPTER 1**

## **INTRODUCTION:**

### **AN OVERVIEW OF THE DEVELOPMENT OF A MODEL OF VISUAL PERCEPTION.**

*This introductory chapter aims to provide an outline of the work that follows. An initial overview is provided, then the chapter has been divided into four sections. The first is an introduction to perception through an overview of previous attempts at explaining vision. The second is an overview of the theoretical framework in which the model is to be based, introducing the approach that has been taken to modelling vision. The third section is an introduction to the model itself in relation to the previous two sections. Finally there is a summary of the following five chapters.*

## **INTRODUCTION.**

### **AIMS OF THE THESIS.**

Before embarking on a description of the thesis it is important to explain the motivations behind the work that follows. The title of the thesis begins with the words; "Towards the development of a model of vision..." This is extremely important because as it suggests this work in no way constitutes a definitive account of a model of biological vision, neither does it describe a model implemented in software for use in the development of artificial vision applications. The aim then is quite simply twofold.

The first aim is to provide a better understanding of visual perception through the development of both a theory and conceptual model of perception. One of the key issues that will be addressed in the thesis is that it is important to have a good conceptual model based on sound theory before you attempt any implementation. There is a danger that pressure to simulate any model developed distracts from the insight gained in understanding what is to be modelled.

The second aim of the thesis, is to provide the foundations for a model of visual perception that could be easily implemented as an artificial vision system, by formulating new more efficient algorithms as a result of studying in detail real neuronal behaviours and architectures.

It is reasonable to assume from attempts at developing artificial vision systems that the closer those systems are to the brain the more efficient they might be. An example of this can be seen in limitations experienced with expert systems approaches to the development of artificial vision systems. Early attempts at artificial vision systems such as "blocks world" (Roberts, 1965; Waltz, 1975) were at best limited to the domain in which they were developed. The advent of artificial neural networks (ANN's) which had their foundations in a simplified understanding of how the brain was thought to function, as outlined in Rumelhart & Mc Clelland's 10 characteristics of brain function (Rumelhart & Mc Clelland, 1986), were more successful in dealing with the problem of restricted domains. The key to ANN's was that they could learn, so could adapt to deal with new information, rather than depend on representations already stored in the

system. However the networks developed were inclined to be restricted to subsets of visual processing. For example successful neural network implementations tended to be built to deal with the problems of low level vision or high level vision, a distinction to levels of visual processing first popularised by Marr (Marr, 1982), separately. One of the reasons suggested for the inadequacies of ANN's in terms of explaining how the brain functioned was that the algorithms, for example "back-propagation" and architectures they ran on were not brain like at all (Massaro, 1988; Ratliff, 1990). If systems developed under these principles do not function like the brain, it would seem reasonable to suggest that in the development of a truly intelligent machine their performance could be bettered. Through the formulation of a conceptual model based closely on the brain it is hoped that systems can be developed that produce much more effective behaviours.

#### DIFFERENT SORTS OF MODELS.

It seems that to progress in the field of understanding any aspect of the brain involves a relationship with computers in some form. This relationship might just be to analyse data. However at the other end of the scale, the development of computer models of neuronal behaviour has proved to be extremely useful in developing the understanding that we have of the brain today. Implementing models of brain behaviour in software provides support for both reasons for modelling the brain. The sort of system that is developed determines how much detail is necessary in both the theory and conceptual model. For example, it is possible to define two broad categories of artificial system development with different aims and

objectives. These two sorts of system correspond with the two reasons for modelling perception that were previously outlined. Firstly there are those systems that will eventually be marketed as being capable of performing some task that was otherwise performed by a human being. Such systems include fault detection devices, robots, chess playing games, etc. The development of such systems is not necessarily concerned with how the system works its concern is producing systems that can perform certain tasks that seem to be examples of "intelligent behaviour". These systems are concerned with the apparent behaviours they produce rather than the underlying mechanisms of those behaviours. However generalisations of the underlying mechanisms of the brain provide valuable artificial mechanisms that mean the performance of the artificial system is superior. Usually one has a task in mind and a system is developed that performs that task, this is referred to as a "top-down" approach to artificial intelligence.

Secondly there are those systems that are developed to demonstrate behaviours but which also rely on architectures and mechanisms that have some kind of biological realism. These systems are built initially to work from a "bottom-up" approach. They start at the lowest level of function and build up on that in the hope that the eventual emergent behaviours will be realistic. The underlying mechanisms are similar to those of the brain. At their purest level such systems are used by neurophysiologists to investigate how the brain actually works. through developing detailed models of neuronal behaviours. This approach also has bearing on the development of systems that are behaviour driven. Such models can provide information about new mechanisms that could be simplified and used in the development



of systems that are designed to perform a given task. However there is a certain amount of overlap between the two sorts of systems developed and it is not always the case that only one particular approach, either top-down or bottom-up, is applied in the development of one or other type of system.

#### A MULTI-DISCIPLINARY APPROACH.

The idea of modelling the brain to understand its behaviour is not new. A distinguishing feature of this thesis is the aim to provide a multi-disciplinary model of visual perception, that encompasses both psychological theory and an understanding of neuronal behaviours. To understand neuronal behaviours, mathematical analysis and computer simulation can provide invaluable tools. This means the involvement of at least four different disciplines, neurophysiology, psychology, computer science and mathematics in the modelling process.

Such a model can take advantage of the understanding of vision that has been developed from each of these perspectives. There are however very few theories of vision that are not related specifically to one discipline. There are psychological, neurophysiological, and philosophical theories of how we see. It is a relatively simple task to develop models of vision from detailed neurophysiological research. It is when there is no specific physiological information that it becomes difficult to develop models, and psychological theories and behavioural information from psychological experiments become central to understanding (Bruce & Green, 1990). However reconciling neurophysiological data with psychological theory is not an easy task. The eventual outcome is that you have several theories

some derived from neurophysiology and others derived from psychology. In turn this means that most artificial intelligence models have been developed through concentrating on one of these particular theories.

Links between what we know about vision and how we can create machines that can perform the same tasks have been dominated by theories of vision originating from cognitive psychology. The implications of this are that artificial vision systems tend to be all developed based on the same theories. Arguably it is only through major paradigm shifts, or the advent of new theories generally linked to the failure of existing theories that changes in the models take place. Thus if a theory is dominant for any length of time, the models remain unchanged regardless of developments from other disciplines. For example the expert systems approach was dominant for a long while in artificial intelligence even though new approaches to intelligent systems development were apparent. Neural network theory that eventually usurped the dominant expert systems approach had its foundations in 1940's neurophysiology, although it was the early 1980's before artificial neural network models began to be recognised in the development of artificial intelligence models.

Through looking at the capabilities of artificial vision systems to date it becomes obvious that there is a need to improve on their performance greatly. They are generally restricted to single tasks or "domains" and compared to the human visual system they are greatly inferior. The insight artificial systems applications have provided has been through their inability to perform rather than through actually demonstrating real visual behaviours. A classic example being the previously mentioned "blocks

world" program. The blocks world system was unable to process artificial geometric scenes, being unable to generalise to object recognition in natural visual images (Roberts, 1965; Waltz, 1975).

#### PROBLEMS OF MULTI-DISCIPLINARY APPROACH.

There are many problems associated with multi-disciplinary approaches to the study of any subject. One such problem is that by combining information from different disciplines, in this case in research into vision, there is a danger that any discoveries about the nature of the problem may seem to be general rather than detailed. Detailed understanding evolves from looking extremely closely at a particular aspect of perception. A specific problem, or area of the brain is studied in depth, usually through research in a particular discipline such as neurophysiology or psychology. This often means that any insight gained is dependent on the research methods associated with that discipline. For example psychological interpretations of vision often centre on behavioural experimentation to support a theory, whereas neurophysiological research provides detailed accounts of how the brain actually functions in terms of the mechanisms of neuronal behaviour. To make things even more complex what is assumed to be a singular discipline comprises again of a whole set of sub-disciplines. Arguably neuroscience itself is a multi-disciplinary research area, involving the work of anatomists, biochemists, and physiologists working either in clinical neuroscience or experimental neuroscience (Bloom & Lazerson, 1988), all using their own particular research methods, all working on different aspects of the same problem. The same applies to psychological

approaches to understanding vision. Within the discipline of psychology there are developmental psychologists, cognitive psychologists, psychobiologists, etc., all looking at vision from their own perspective. This leads to another very real problem, how information is communicated within and across disciplines. It is not surprising that there is little or no relationship between a model developed for example from psychological research and a model developed from neurophysiological research as they are both working at very different levels of explanation. The descriptions they provide about vision could be seen as being written in different languages. This means that if you understand one particular language, you will be able to understand what the research in that discipline is saying. For example if you are a neurophysiologist you will explain vision in terms of its underlying neuronal processes and relate this information in a specific language used to describe these mechanisms. If however you are a psychologist the information will not hold the same meaning, any understanding will be limited compared to that of a neuroscientist. If a clear understanding is difficult to communicate at the theoretical level there is little chance of that information being of any use in developing artificial intelligence models.

If the development of a better understanding of vision or the development of new artificial systems capable of perception is to become a reality there has to be inter-disciplinary communication. It seems reasonable to suggest that an understanding of vision in its fullest sense is the first step to developing any sort of model for whatever reason. Even if your intention is to model a very small part of vision, or even if you are only concerned with

modelling vision at its highest level, it is essential to have the most complete understanding possible. It seems clear through looking at the development of artificial vision systems that the demands of being able to implement a model in software shape the kind of model that is developed. This leads to things being added or left out purely because of the need to code the model. Eventually you may lose important features of the model just so you can produce the desired behaviours. It is therefore my aim not to emphasise the development of a simulation of the model. If a strong theoretically based conceptual model is developed first then maybe its eventual simulation will retain more of the delicate intricacies underlying vision in translation. Further, if significant changes have to be made to enable simulation, there is always the conceptual model to refer to for a more detailed description.

To be able to fully understand information from other disciplines you need to have some understanding of those disciplines. One possible solution to the problem of understanding the diversity of relevant information provided by several different disciplines is to translate everything into a single language that can be understood by everyone through a common factor. Initially this would seem to be a good idea, but there are problems with this approach. One obvious problem is that information is lost in translation. Some types of information have to be represented in certain ways and could not be translated, or else attempting to translate them becomes a central issue distracting from understanding the information itself. A crucial aspect of any attempt at multi-disciplinary research is to overcome these problems. It is envisaged that by using a strong theoretical framework through

combining a top-down and bottom-up approach the problems of multi-disciplinary research can be overcome.

## **SUMMARY.**

Through the combination of research into visual perception from varying disciplines it is hoped that a conceptual model can be developed that will provide a more complete understanding of vision. The aims of the conceptual model are both to form a more comprehensive understanding of vision, and to aid the development of more sophisticated artificial vision systems. There are many diverse explanations of various aspects of visual perception, both neurophysiological and psychological. Through uniting such a plethora of information it is hoped that progress in understanding the brain will be made. In turn this progression will influence the development of artificial systems. As noted previously this work does not aim to provide an account of an implementation of a model of vision. It aims to provide a strong theoretical framework for a conceptual model that could ultimately be developed in software.

## **A THEORETICAL FRAMEWORK FOR MODELLING VISION.**

### **TOP-DOWN VERSUS BOTTOM-UP.**

Churchland addresses the issue of “top-down” versus “bottom-up” processing as a theoretical framework with which to study the structure and function of the brain. (Churchland, 1986). This distinction between top-

down and bottom-up in principle is much the same as discussed previously, although Churchland specifically defines the terms in relation to the study of the brain. Churchland refers to top-down and bottom-up as follows.

A top-down approach involves starting with a theory then explaining the underlying mechanisms of that theory, for example taking a psychological approach to the study of visual perception. This would involve having a theory of visual perception and then going out to look for evidence that supports that theory. A bottom-up approach on the other hand starts by working up from the lowest level of explanation to develop a theory, e.g. taking a neurophysiological approach to the study of visual perception. This would involve trying to understand how the brain works at its lowest level, for example at the neuronal level, then using that information to develop a theory of perception. However if you only work at one level you miss out on important information that would become evident if you were working at the other. So in the case of vision if it was studied merely in terms of low level neuronal behaviour we would not progress towards an understanding of perception. The same is true if we start with a theory of higher cognitive function and did not look at possible underlying mechanisms for that theory. By disregarding the elements of neuronal behaviour the theory of perception might never be complete. It may miss out on subtleties that an explanation at a lower level might provide. Churchland goes on to illustrate the importance of a unification of top-down and bottom-up approach to understanding the brain. As previously discussed this theme provides the basis for the conceptual model of vision that is developed here. The aim is to provide a theory of visual perception that is easily reduced to its

neurophysiological basis, the rationale being that so called "higher level" cognitive functions must have underlying neuronal foundations. The aim is thus to explain and develop a conceptual model of visual perception by looking at the underlying architecture and mechanisms of vision, i.e. the brain, what it is and how it works.

Churchland goes on to explain that the majority of our understanding of the brain has been developed through adopting either of these two approaches. For example there are neurophysiologists working on the lower levels based on a bottom-up approach. In contrast there are psychologists developing "grand theories" of how the brain works from a top-down perspective. What is clearly evident according to Churchland is that there is a serious lack of communication between the two approaches. Not only that, but there are also many disciplines using very different tools of explanation that could be relevant to understanding the brain, an example being the discipline of mathematics. Using the tools of mathematical analysis could provide essential insight into the behaviour of low level neuronal behaviour. Although there are many reasons for this lack of communication leading to a lack of progression in our understanding of the brain Churchland suggests that one important factor is the subject matter itself.

#### **BRAIN EQUALS BRAMBLESCAPE.**

Churchland describes the brain as "bramblescape", an unruly mass of wild and tangled structures. This makes it difficult to study as the fibres bringing information into the cortex and fibres taking out are tangled together. It leads on from this that types of information, both incoming and outgoing



are therefore difficult to characterise. So relating a top-down theory of the brain to a bottom-up theory becomes an almost impossible task. In the case of developing a theory of visual perception clearly the problem is making sense of the tangled mass of fibres in an attempt to try and develop an understanding of what actually constitutes perception at its lowest level before any theory can be developed. We have to go beyond bramblescape, and the key to this and ultimately the key to developing a conceptual model of visual perception is to make sense of the brain through the emergence of structure from mayhem. The problem of understanding visual perception becomes the problem of understanding the information processing structures and functions of the brain. By looking at visual processing structures and functions, the lowest level of behaviour, a theory of visual perception will evolve. The two are reliant on each other, in that understanding processing provides information about perception, and ultimately how the brain works. Although processing can be explained in isolation from perception it has no real value in being able to explain anything about the brain as a whole unless it is linked to a theory of perception. (Churchland, 1986).

Through developing models of the brain it is possible to develop a unification of both bottom-up and top-down approaches. The sorts of models developed even at the neurophysiological level can be categorised into two types of model, "simplifying brain models", where the generalised neurophysiology is used (Sejnowski, Koch & Churchland, 1988), or finer more detailed models (Wehmeier, Dong, Koch & van Essen, 1989). They suggest that the essential differences between these two sorts of models can be likened to the distinction that Chomsky makes between "competence"

and “performance” in models of language understanding (Chomsky, 1965). Competence models are similar to simplifying brain models whereas performance models are related to detailed models of structure and function of the brain. You may get the same behaviours with the two but the latter is more realistic. The former, however, provide an excellent overall picture from which to build more detailed models (Wehmeier, Dong, Koch & van Essen, 1989).

## **INTRODUCTION TO A MODEL OF VISUAL PERCEPTION.**

### **THE IMPORTANCE OF VISION.**

Vision could be considered as being our most important sensory system, making up over forty per cent of all sensory input (Dowling, 1992).

Intelligence can be defined as having the ability to adapt to our constantly changing environment. Within this fluctuating environment we encounter and have to solve many problems enabling us to act upon the world in which we exist in order to maintain that existence (Beer, 1989; Gould, 1981).

Information about our environment arrives through our sensory receptors in differing forms, e.g. light waves (eyes), chemicals (mouth and nose), sound waves (ears), etc. This information is then directed to the relevant area in the brain to be processed to enable incoming information to be interpreted and acted upon. Although all of our senses are important to us it could be argued that visual information is the most crucial to our performance in our environment. Without hearing we can still function well enough to survive, the same can be said about smell and taste. Visual impairment on the other

hand can have devastating effects on our survival. This then suggests that explaining how this system works is a key issue not just in developing artificially intelligent systems, but also in understanding real biological intelligence:

*"Our thinking is to a large extent visual imagery. All navigation and manipulation is guided directly by vision or by imagined geometry. Our intelligence is curiously restricted to things that can be formulated as scenes, which are largely visual in substance. Language only makes sense by continuous reference to visual imaginations. Understanding vision very likely will mean understanding the brain and mind."*

von der Malsburg (1990).

## WHAT IS VISION?

To attempt to define vision calls for an explanation of all the component parts that go to make up an understanding of what "seeing" involves. Vision is a general term which is used to describe both the mechanisms and apparatus that are involved in the processing of visual information, as well as the actual activity of "seeing" in its global sense, i.e. vision as a cognitive function, as an interpretation rather than purely a physical representation. There is a clear distinction between explaining "seeing" in terms of mechanisms and apparatus, and explaining "seeing" as cognition in the psychological sense. The distinction is perhaps best described as the distinction between visual processing and visual perception.

## VISUAL PERCEPTION OR VISUAL PROCESSING?

Visual perception is not the same as visual processing. There is a clear difference between the two, although the latter is the basis for the former.

Visual processing involves understanding seeing in terms of light falling on the retina, reflected from objects in the visual scene, and the mechanisms and structures in the brain that process that information. The level at which the information about the visual image is described is purely neuronal. That is, it is understood in terms of the electrochemical reactions that occur in relevant neurons in specific areas of the brain dedicated to visual processing. The study of visual processing might involve tracing the path of information through the brain, identifying the various areas involved in processing visual information. It also might involve describing neuronal behaviour in those particular structures, and making some sense of how visual information is represented in the brain. Visual processing therefore, is about looking solely at how light falling on the retina is processed and encoded by the component parts of the brain that make up the visual system, and how the brain combines all the relevant information about the patterns of light falling on the retina to provide an overall representation of the visual scene. Visual processing is thus best understood through neurophysiological theory.

On the other hand, visual perception is not just about the processing that takes place when the brain is representing visual information. Visual perception is about interpreting and making some form of evaluation of that information. Rather than just describing the structures and mechanisms underlying the encoding of the visual scene or an image it involves acting upon that information, using it to perform transformations upon the world. Perception is crucial to the notion of intelligence as adaptive behaviour, our ability to solve problems in an environment that is constantly changing, in

order to survive. (Beer, 1989). A possible definition of visual perception is as follows:

*Extracting from the retinal image a meaningful description of the world that can be used to carry out necessary actions on that world.*

Bearing in mind that the image that is projected onto the retina is static and meaningless, whereas the world we see is constantly changing and meaningful, giving an image meaning involves going beyond understanding the image in terms of how it is represented through its processing. (Bruce & Green, 1990.)

This in turn means that perception is not necessarily limited to a description of the world in terms of visual sensory input alone. It is possible to separate perception in general into two types, local and global. Local perception would be the meaningful interpretation of sensory processing in a particular sensory area. For example visual perception would be a product of visual processing and auditory perception a product of auditory processing, etc. Global perception would be a combination of all the available local perceptions to form an overall description of the world based on all the available sensory information. Global perception is extremely complex and does not simply consist of the result of visual sensory processing, it also incorporates the result of sensory processing from all the other senses, information about the world that we have encountered previously and stored in memory.

Visual processing can thus be explained as low level neuronal processing in the brain whereas visual perception is about the relationship between that

low level processing and higher level cognitive processing. Higher level cognitive processes being so called “higher” functions of the brain, such as language, memory and learning, usually described by psychological theory rather than through neurophysiology. In relation to vision such functions might include object recognition, visual attention and visual memory. Central to this idea is how the links are made between higher level cognitive interpretations of the image and low level neuronal representations. We can interpret and react to visual information extremely quickly, recognising and acting upon visual sensory input almost instantaneously. It is evident that we can make judgements about visual stimuli in as little as 150 ms (Biederman et al, 1982). Further, neurophysiological measurement of cortical cell responses are in a similar time domain (Oram and Perrett, 1994). It is therefore extremely unlikely that high level cognitive processing is separate from the low level neuronal processes that underlie visual perception. The two must be linked in some way. The question is at what point in the processing is the link made, and how? Do we have two very separate and distinct levels of understanding visual information, one encoding and the other interpreting, operating in serial, or is visual perception inextricably interwoven into the underlying neuronal processes? If the latter is the case at what point does low level neuronal activity become high level cognitive perceptions?

#### THE BINDING PROBLEM.

At the level of processing and at the level of perception, local or global, whether separate or simultaneous, a key issue arises, how are all the

separate parts of the visual scene combined to form an overall representation? How are low level neuronal processes linked to high level cognitive interpretations? How are local perceptions linked to global perceptions? Somehow all this information has to be "bound" together, combined at all levels to provide a complete and meaningful understanding of our world. This problem of combining all the relevant information about an object in the visual scene is referred to as the "binding problem" (Hummel & Biedermann, 1990; 1992.) or more simply, the problem of integration of information. At its lowest level, in terms of visual sensory information binding is related to the "segmentation problem" (Marr, 1976) or "feature linking" (Gray & Singer, 1989). The segmentation problem refers to how the various segmented parts of an object in the visual scene are combined to form a single representation. For example how are all the segments that go to make up the outline of an object linked together. The problem is accentuated by the fact that objects have several aspects of form that have to be taken into account when they are encoded, for example whether or not the object is moving or in what orientation it occurs. The binding problem has become a central issue in both developing theories of vision and in developing artificial systems capable of vision. In terms of developing theories of binding, typically top-down psychological approaches have been dominant with little evidence of a multi-disciplinary approach. Recently however the development of bottom-up neurophysiological approaches that has enabled mechanistic theories of binding to emerge. Until recent discoveries in neurophysiology, theories of binding were dominated by "feature detection" theories, which suggested

that visual binding in object recognition was a product of the hierarchical detection of more and more complex features by cells in the visual cortex. Eventually at the top of the hierarchy was a cell that represented that particular object in its entirety (Barlow, 1972). This theory was well suited to information processing models in artificial intelligence (Lindsay & Norman, 1972) and so was dominant in the field for some time. It was also supported by apparent neuronal mechanisms in the primary visual cortex (Hubel & Wiesel, 1962). Recent discoveries in neurophysiology have changed that view, through the advent of so called "population models". Such models suggest that object recognition is a product of the distributed encoding of information across populations of neurons. Previously evidence for mechanisms to support this idea was non existent resulting in little interest being taken in theories based on this idea (Stryker, 1989). With the advent of new neurophysiological evidence and with the failure of artificial systems based on traditional information processing models to either provide adequate support for feature detection theories or produce systems capable of visual binding (Stryker, 1989), multi-disciplinary approaches are now being pursued. Previously general accounts of binding included theories from cognitive science implemented in standard neural networks, e.g. multi-layer perceptrons (Hinton et al, 1986; Chalmers, 1990; Pollack, 1990; Blank et al, 1992). Standard neural networks is a term given to the class of neural networks originating from work in the 1980's popularised by Rumelhart and Mc Clelland's "parallel distributed processing" (PDP) group (Rumelhart & Mc Clelland, 1986). The problem with this approach is that standard neural networks do not really provide



any insight into the mechanisms that underlie binding. Further, standard neural network theory alone has been shown to be insufficient in the development of successful implementations of models of visual binding (Hummel & Biedermann, 1990; 1992). This approach has doubtful relevance in the advancement of understanding binding because to implement a theory of binding in this way requires that you already have both a theoretical and conceptual understanding. An implementation in standard neural networks, which are unrelated to actual brain mechanisms, is simply one possible way of representing a model or theory. It is like translating your theory from a verbal description to a pictorial description. It tells you no more about it, it just provides another form for representing it. Therefore implementing a psychological theory of binding through a network that is unrelated to actual brain mechanisms will tell you no more about binding than you knew before the implementation. Clearly to discover more about the nature of binding, and to build more successful vision systems, developing models that incorporate more realistic brain mechanisms is essential. What is needed, and what is now becoming apparent, is a multi-disciplinary study into the very nature of perception, an iterative procedure based on both top-down theories supported with bottom-up mechanisms that, through implementation, will allow modification and support of each other.

## CONCLUSIONS.

Two levels of understanding vision are proposed: the level of processing, which demands a detailed analysis of structure and function of the parts of the brain involved in vision; and perception, which involves giving meaning

to the product of low level neuronal activity through high level cognitive representations. Perception or higher level cognitive interpretations take place locally, being "sensor specific", and globally, incorporating information from all available senses and from memory. It is possible to study vision in terms of processing alone, but it is not possible to study perception without studying the processing that underlies perception. So to look at visual perception involves understanding visual processing. Through the development of a conceptual model of visual perception based on low level neuronal processing a possible solution to the binding problem can also be proposed.

The development of a truly multi-disciplinary model of vision begins with looking at the neurophysiology of the visual system. An investigation into the architectures and mechanisms that underlie vision provides the next step, bearing in mind that several theoretical assumptions about the nature of perception have already provided an outline for the development of a conceptual model of vision. By taking a truly multi-disciplinary approach it is hoped that both the aims of the thesis will be fulfilled, to provide new algorithms for machine vision and insight into the nature of perception. The intention is to start by introducing what is known about vision from a bottom-up view, and then relate this to theories of visual perception at a higher cognitive level, ultimately providing a deeper understanding of what vision is, and deriving new architectures and algorithms for developing artificial vision systems.

## **SYNOPSIS OF THE FOLLOWING 5 CHAPTERS.**

Chapter 2 aims to provide an overview of the neurophysiology underlying visual perception. This chapter comprises of a simplified account of the low level neuronal processes that constitute visual processing. It is from these descriptions of structure and function that an architecture for a conceptual model of visual perception has been constructed. This chapter forms the basis of the so called bottom-up explanation of visual perception as discussed previously in this chapter. As the visual system, like the rest of the brain is extremely complex, a simplified account of the architecture and function is all that can realistically be achieved.

Chapter 3 aims to provide an overview of both the psychological and neurophysiological theories of visual perception. The neurophysiological theories are accounts of visual processing rather than of visual perception whereas the psychological theories are attempts at explaining perception rather than visual processing. This chapter is more concerned with a top-down approach to investigating visual perception, although the neurophysiological theories are best described as top-down interpretations of bottom-up accounts of visual processing. The psychological theories are “purely” top-down in that they do not have to make reference to bottom-up descriptions. This distinction is not as clear as it seems as there is an enormous amount of overlap between neurophysiological and psychological theories. The aim of this chapter is to provide an evaluation of the available theories of visual perception enabling the development of a theory on which the conceptual model has been based.

In Chapter 4 both the theory and the conceptual model have been made explicit. Initially a theory of perception has been outlined which has then been used to shape the conceptual model. The theory originates from a combination of current explanations of visual perception and the conceptual model provides a structural and functional description that supports the theory. The theory and resulting conceptual model that have been developed in this chapter attempt to both explain visual perception in novel way and provide new and innovative approaches to the development of artificial vision systems.

In chapter 5 additional support for the theory and the model as outlined in chapter 4 has been made. Although chapters 2 and 3 provide the foundations for the component parts of the model this chapter provides more concrete support for the model and theory as a whole.

Finally, chapter 6 comprises of a summary and discussion of the achievements of the thesis and suggestions for further work.

The thesis attempts to provide a “complete” account of visual perception through the development of a theory and conceptual model that supports that theory. It attempts to do this through a multi-disciplinary approach that unites both bottom-up neurophysiological and top-down psychological accounts of vision. Further it is envisaged that the development of such a model will provide new insight into the development of artificial vision systems and new algorithms for perceptual function in such systems.

# **CHAPTER 2**

## **THE VISUAL SYSTEM:**

### **AN OVERVIEW OF THE NEUROPHYSIOLOGY OF VISUAL PERCEPTION.**

*This chapter aims to introduce a simplified account of the neurophysiology of vision, providing an overview from which the outline of a conceptual model might be developed. Part one attempts to explain why we need to understand the neurophysiology of vision, and introduce a general view of visual processing. Part two then gives a more detailed account of the visual processing pathways in the brain and explanations of the various areas associated with visual perception. From the point of view of the thesis as a whole, the description here will be general in the sense that when the conceptual model is introduced in chapter 4 a more detailed account of neurophysiology will be made in order to support the model.*

### **UNDERSTANDING VISION.**

#### **INTRODUCTION TO STRUCTURE AND FUNCTION.**

The aim of developing any model of vision involves understanding both the processing, in the form of neuronal behaviour that takes place, leading to an eventual understanding of the higher level cognitive states that enable us to perceive our world. If we regard low level processing as being the foundations from which higher level interpretations occur, then it seems

reasonable that a good starting point in attempting to explain vision would be to trace its evolution from light falling on the retina to the areas of the brain that deal with that information. It is important to emphasise that most of what we know about the structure and function of the brain originates from extrapolating the results from experimental work with animals to the human brain. The majority of the work presented in this chapter is the result of experimentation in the cat or monkey, and not in the human brain.

Although we have extremely advanced techniques for tracing the flow of information in the brain there is still a great deal to be discovered about the structures and their underlying function. The speed at which understanding advances is hampered by the differing methods of investigating the brain and a lack of consensus in terms of the results such work produces (van Essen, 1985). This means that a lot of neurophysiological evidence is often speculative in terms of the less well understood areas of the brain. As van Essen points out it was only twenty years (now thirty years) ago that the area of the brain processing visual information was thought to be made up of three concentrically organised areas (van Essen, 1985). The next step would be to describe the underlying mechanisms of these structures. To do this the relationship between structure and function in the cortex must be addressed.

The structure of the cortex is not entirely independent of its underlying function. It is not as simple as having a set of pre defined structures that any old function can be applied to. How neurons behave influences how structures develop and evolve. In the same way, how structures evolve can influence how neurons behave, not necessarily individually but as groups.

Many studies have been conducted addressing the issue of cortical plasticity and the effects of visual deprivation on the development of the visual cortex and resultant perception. If the visual apparatus is in some way damaged, for example monocular deprivation through one eye not functioning, then the visual cortex loses its binocularity. Projections from the retina are pronounced from one eye causing a change in the normal structure of the visual pathway where both eyes would share the same distribution of ganglion projections (Le Vay, Wiesel & Hubel, 1981; Hubel, Wiesel & Le Vay, 1977). In the same way, if young animals are raised in artificial environments deprived of specific visual information the brain is unable to readjust to normal function when the animal is no longer visually deprived, suggesting that some of the structures of the brain have not been able to develop fully. (Hirsch & Spinelli, 1971). It is important to note that up until recently it was thought that there are so called critical periods in the development of the visual system, in cats and monkeys this is from birth to 6 or 7 months whereas in humans the critical period is up to 6 years (Hubel & Wiesel, 1970; Thompson, 1985). After this critical period has elapsed the visual system no longer changes in its structure significantly enough to effect visual perception. Recently though new research has suggested that this may not be the case (Gilbert, 1992). It seems that both the structure and function of the visual system are intertwined, changes to either influence visual perception. It is not sufficient to merely model one or the other to understand visual perception.

## A GENERAL UNDERSTANDING OF VISUAL PROCESSING.

The starting point for developing a biologically plausible model of visual perception means understanding and developing a model of visual processing. Initially investigation must follow the path of visual information from it entering the brain to its end product, recognition and identification. The process of investigation demands an understanding of the architectures that form part of that pathway, and how they function in terms of processing that information. To begin to describe vision in simple terms is in itself a complicated task. How the brain deals with light falling on the retina, and describing the areas of the brain that deal with this sensory information, is not as clear as it first seems. Already, confronting the problems discussed in chapter 1 relating to the multi-disciplinary nature of modelling perception is a major issue. Firstly there is a need to have an understanding of physics so that the form, which information is transmitted to the visual system can be understood. There is then a need to understand neurophysiology, the behaviour of neurons and the structures they form as well as psychological theory explaining the eventual behaviours that are a result of interpreting and acting upon the information that has been processed. Because of this diversity of understanding and the enormity of the information available, this chapter aims to introduce a "simplifying brain model" (Sejenowski, Koch & Churchland, 1988), as discussed in chapter 1. Then from this general understanding of global principles a more detailed model can be developed (Wehmeier, Dong, Koch & van Essen, 1989). It cannot be stressed enough that the enormity and complexity of



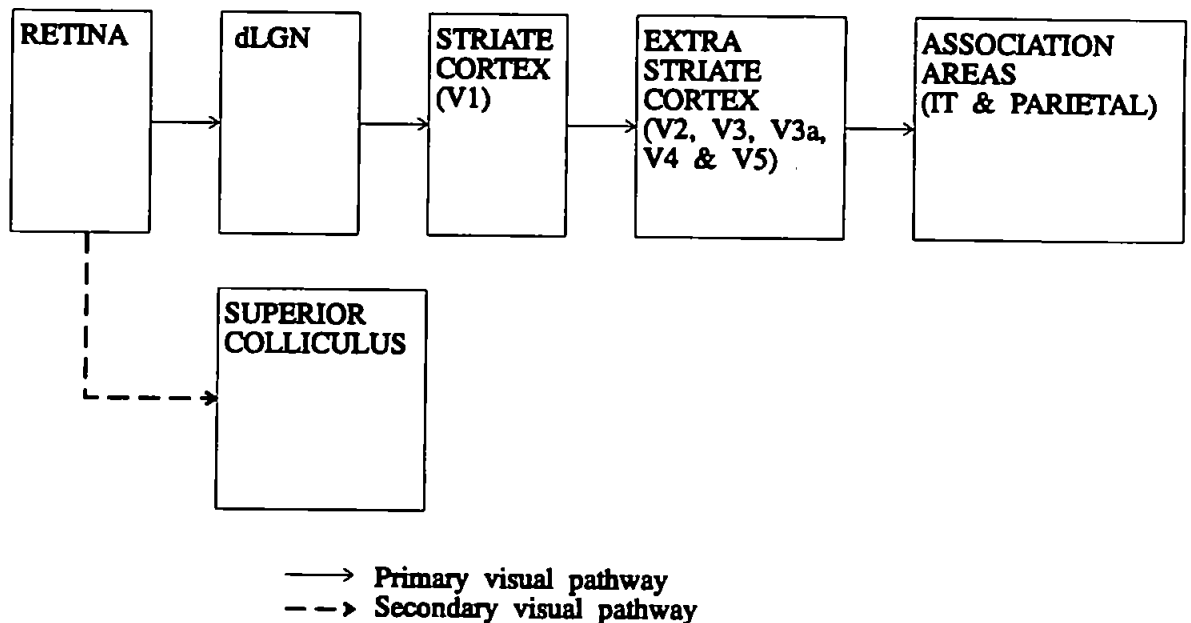
information from neurophysiology alone means that only an extremely simplified description of vision is possible if the issue of perception in its global form is to be addressed.

In its simplest terms the visual process begins with information being transmitted to the perceiver in the form of waves of light. These waves of light are reflected from objects and forms that make up the visual scene and are absorbed by the first part of visual apparatus, the retina. The retina is the starting point for the processing of visual information and is often referred to as a "pre processing stage". The retina is made up of light sensitive receptor cells that absorb the waves of light being reflected from the objects in the visual scene. At this point the first transformation of information takes place as light waves are turned into electrical signals, a medium with which the brain is able to work. (Bruce & Green, 1990; Dowling, 1992.) Although at this stage a certain amount of processing takes place it is generally regarded as preparation, transformation by the retina of the information from the visual scene into a form that can be understood and passed on for in depth processing by dedicated areas in the brain. The image at the retinal level is only two dimensional. As we are able to perceive three dimensions it is assumed that further processing allows us to infer three dimensional interpretations of a two dimensional image. The destination of the outcome of retinal processing of the two dimensional image is therefore all important to understanding perception.

There are two major pathways that retinal information is transmitted on, both via the optic nerve. The first route known as the "primary visual pathway" leads from the retina down the optic nerve to the dorsal part of the

lateral geniculate nucleus (dLGN) in the thalamus and then on to the visual cortex. The path that those signals follow, through the primary visual pathway, can be traced to at least six synapses. Two are in the retina, the third in the dLGN, the other three represent at least three distinguishable levels of processing in the visual cortex (Dowling, 1992).

The visual cortex is divided into two main areas: the striate cortex which consists of the primary visual cortex (V1) where most of the information from the dLGN is projected to and the extra striate areas which are those areas beyond the primary visual cortex (see figure 1).



**Figure 1.** A schematic of the visual processing areas of the brain including the primary and secondary visual pathways.

It is this pathway that is of most interest as it leads to the area in the brain that carries out the processing of visual information that leads to eventual perception. It is well accepted that the cortex is the area of the brain that processes sensory information due to the large number of ganglion axons terminating there, other pathways are considered to be less important in

terms of the processing of the visual scene. (Bruce & Green, 1990; Dowling, 1992; Cowey, 1981.)

The secondary visual pathway leads to the superior colliculus (SC), which is thought to be involved with controlling eye movements and possibly plays a role in visual attention (Churchland, 1986). The SC is also implicated in the integration of the processing of sensory information as it also responds to auditory input enabling the eyes to foveate on objects omitting sound, thereby linking visual and auditory information (Churchland, 1986). However essentially the role of the SC in perception is indirect rather than direct, that is, its role is more with controlling what we see rather than saying what it is we are seeing (see figure 1). There are other visual pathways leading from the retina to other areas in the brain but their function is not as well defined. That is not to say that these pathways are not significant. In relation to both global and local perception as outlined in chapter 1 these pathways may perform key roles. One reason for the lack of interest in these other pathways could be due to the role they play in visual processing. Like the secondary visual pathway their involvement in the visual process is related more to the control of what is perceived rather than performing the action of perception. They are often referred to as dealing with so called reflexive "non cognitive" processing, therefore are not given the attention they perhaps demand (Cowey & Stoerig, 1990). However if perception is seen as an emergent property of the brain as a whole then the importance of these pathways could be said to be being overlooked.

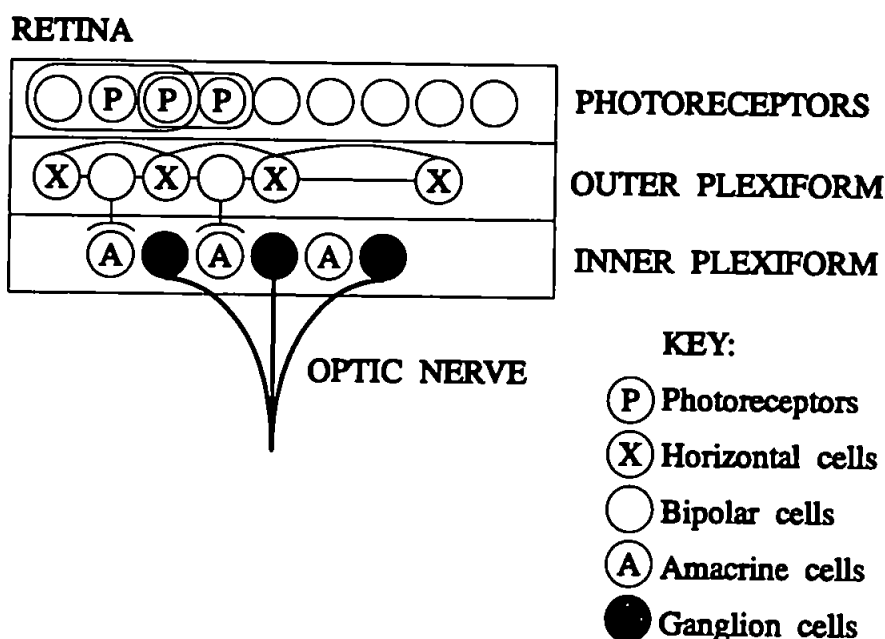
# **INTRODUCTION TO THE NEUROPHYSIOLOGY OF THE VISUAL SYSTEM.**

## **THE RETINA.**

The retina is the point at which visual processing begins. A mapping of the visual scene on to the retina is the beginning of the perceptual process. The retina then encodes that representation and passes the information on for further analysis. A great deal of research has been conducted into the structure and function of the retina, it is perhaps the most well understood part of the visual system. As a result it has been possible to develop extremely sophisticated models of the retina, and it has even been possible to implement such models in hardware (Mead & Mahowald, 1990; Mahowald & Mead, 1991). The retina does more than just sample the visual scene. Although it is beyond the retina that visual processing begins in earnest, it is important to mention key points about the retina's role in the processing of the visual scene as certain aspects of the retinal image are preserved in the processing that takes place in the cortex.

The retina encodes both spatial and temporal information present in the visual scene in a series of light sensitive receptor cells. This information is then transformed into a pattern of electrical activity that translates the spatio-temporal information into a form that can be used by the relevant structures in the brain (Bruce & Green, 1990). The transformations that take place at the retina have been studied in great detail in lower level life forms such as the mudpuppy (Werblin & Dowling, 1969) and the horseshoe crab (Hartline & Graham, 1932; Hartline, Wagner, & Ratliff, 1956). The retina

of vertebrates has also been investigated in depth (Kuffler, 1953; Dowling, 1968). The retina as described by Dowling consists of two layers of synaptic cells (two plexiform layers) each having its own distinct kind of processing. Light hits photoreceptor cells providing input to the outer plexiform layers. In this first layer two main sorts of cells are found, horizontal cells which mediate lateral interactions within the layer, and bipolar cells that send the resultant processing to the inner plexiform layer. Amacrine cells in the inner plexiform layer perform the same lateral mediation tasks that the horizontal cells carry out in the outer plexiform layer. Ganglion cells then act as the output from the inner plexiform layer and the retina as a whole, their axons making up the optic nerve fibres terminating mainly in the dLGN and SC (see figure 2).

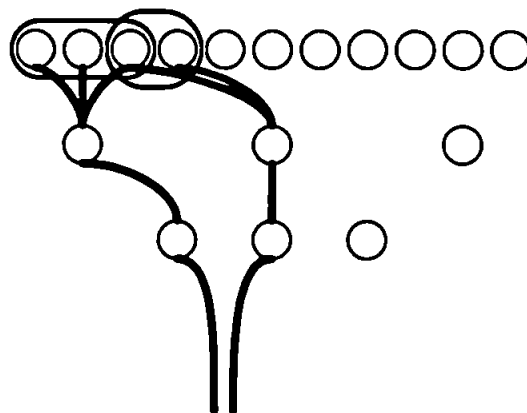


**Figure 2.** The basic structure of the retina.

There are only two action potentials generated in the retina by amacrine and ganglion cells. The photoreceptors, bipolar and horizontal cells respond with graded potentials (Dowling, 1992). Masland notes in a review of

retinal processing that the functionality of the retina was until recently thought to be directly related to the five sorts of cells present in its structure; an example of structure defining function. However he goes on to point out that these five classes of cells comprise of subtypes that multiply the number of functional elements by at least 50. (Masland, 1986).

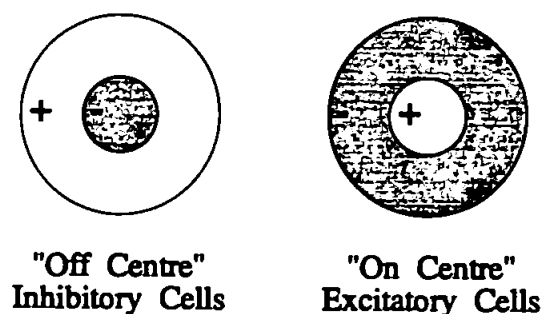
The basic function of the mammalian retina in terms of the processing that takes place can be summarised as follows. There are essentially two paths of information flow, one from the photoreceptors to the brain and the second is lateral mediation pathway that affects the processing within the retina itself. An important feature of the vertebrate retina is the convergence of information, with a ratio of approximately 100 to 1 from photoreceptors to ganglion cells (Dowling, 1992). Each ganglion then has a corresponding group of photoreceptors related to it, this is known as the receptive field of the ganglion. The closer to the fovea the receptive field is the smaller it is, so conversely ganglions with the receptive fields on the edge of the retina have larger receptive fields. Receptive fields are not distinct there is a certain amount of overlap between them (see figure 3).



**Figure 3.** The receptive fields of retinal ganglion cells.

Recent research into the nature of receptive fields suggests that they are able to adapt and change dependent on a number of factors including attentional mechanisms, and contextual information relating to the presentation of given objects (Gilbert, 1992). This has direct effects on the stability of the structure of the cortex. As retinotopy is maintained throughout the primary visual cortex, changes to the receptive field size on the retina directly affect the structure of V1. (Le Vay, Wiesel & Hubel, 1981; Hubel, Wiesel & Le Vay, 1977; Miller & Stryker, 1990) As mentioned previously there are two main sorts of information about the visual scene that are encoded, the spatial information and temporal information. It is possible to distinguish at least two sorts of receptive field properties of ganglion cells, relating to the encoding of either temporal or spatial information.

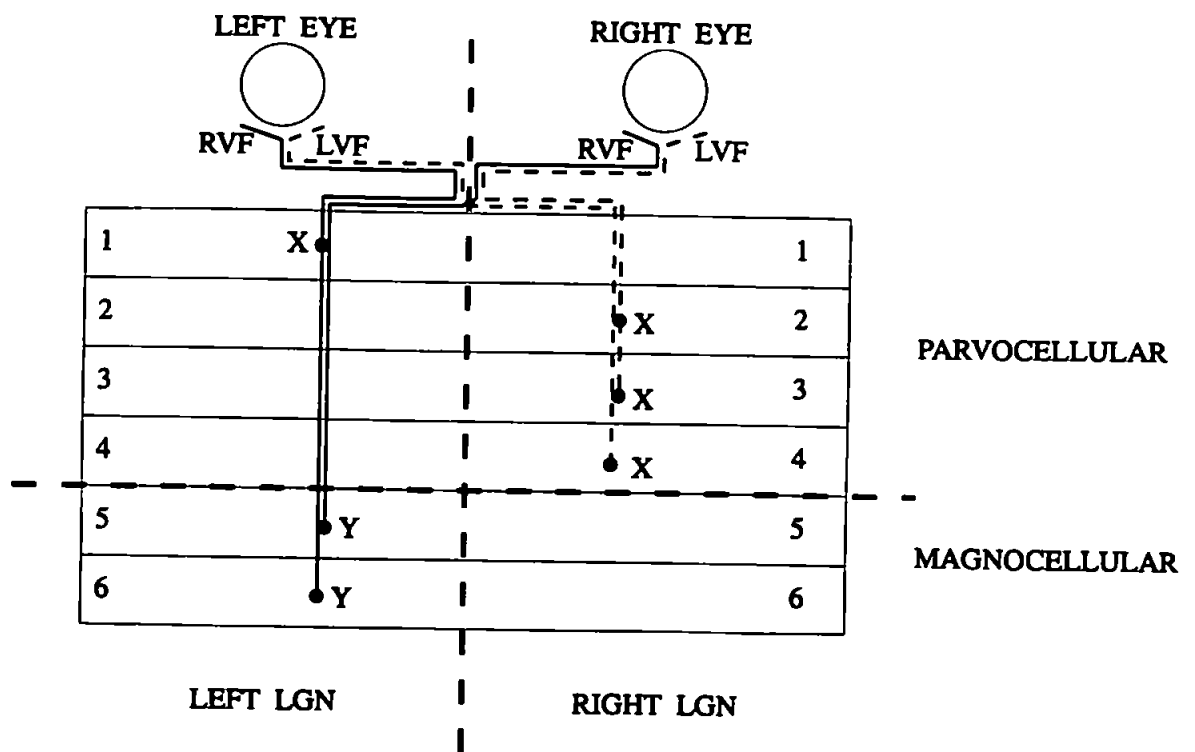
Information is encoded through two processes known as “on-centre” or “off-centre” (Kuffler, 1953). Kuffler discovered that retinal ganglions respond to information from photoreceptors grouped in such a way that they form two concentric circles with either inhibitory cells (off-centre) or excitatory cells (on centre) in the centre circle and the opposite in the outer circle (see figure 4).



**Figure 4.** “On centre” and “off centre” retinal ganglions.

On-centre cells fire with sustained bursts if light falls directly on its centre. If light falls in the surround then the firing is inhibited, finally if it falls across the two a weaker response takes place as the two regions “antagonise” each other. Off-centre cells behave in exactly the reverse way (Dowling, 1992; Bruce and Green 1990). It is possible to discriminate further between two different sorts of concentric field ganglions, X & Y (Enroth-Cugell & Robson, 1966). X ganglions conduct more slowly than Y ganglions and respond in a sustained “tonic” way whereas Y ganglions conduct quickly and respond in a transient or “phasic” way (Carlson, 1986). Hence X cells are thought to encode spatial information and fine detail, whereas Y ganglions code for temporal information and direction. Another difference is in the distribution of X and Y ganglions. X ganglions are concentrated in the central part of the retina whilst the opposite is true of Y ganglions (Bruce & Green, 1990). Both X and Y ganglions project mainly to the dLGN. There is also a third class of ganglion, the W ganglion, that does not correspond to the concentric receptive field and projects unlike the other two types mainly to the SC (Stone & Fukuda, 1974). W ganglions respond weakly and like Y cells prefer moving stimuli. They also have large receptive fields responding in both a sustained and transient way (Carlson, 1986). The result of this processing is that a spatio-temporal encoding of the visual scene takes place and is transmitted down the axons of the ganglion cells to form the third synapse along the primary visual pathway, in the dLGN (see figure 5).





**Figure 5.** The primary visual pathway from the retina to the dLGN showing the X, Y and W ganglion classification.

Obviously not all ganglion cells transmit to the dLGN, as mentioned earlier there are other pathways formed by these ganglion cells. The secondary visual pathway to the SC mainly consists of W ganglions as well as a small amount of Y ganglions. X ganglions along with the majority of Y ganglions terminate in the dLGN. A small number of W ganglions also terminate in the dLGN. Each retina is divided into left and right visual fields, and this structural feature is retained in the passage of information down the optic nerve to the optic chiasma, the point at which the optic nerve crosses. At the optic chiasma left visual field information from both retinas goes to the right hemisphere and vice versa for the right visual field (see figure 5).

## THE LGN.

The lateral geniculate nucleus (LGN) is one of the many nuclei that make up the two parts of the thalamus, which sits on the top of the brain stem, one part being in the left hemisphere the other in the right. The thalamus is considered to be crucial to the integration and transmission of incoming information to the brain, described as “a brain within a brain” (Churchland, 1986). All primary sensory information projects initially to a specific nuclei in the thalamus, where it is then transmitted to the relevant processing area in the cortex (Shepherd, 1979; Carlson, 1986; Churchland & Sejnowski, 1992; Dowling, 1992). Significantly the thalamus is hidden by a mass of both feedforward and feedback axons both providing massive input from sensory organs to their respective processing areas in the cortex and receiving information back from these cortical areas (Mumford, 1991). Not all the thalamic nuclei relay primary sensory information, some correspond to other areas in the brain projecting to related areas in the cortex, and others relay information within the thalamus itself, such as the midline nucleus which receives input from the reticular formation and projects to other thalamic nuclei (Carlson, 1986). Three distinct types of nuclei can be identified: specific central relay nuclei such as the LGN, which send incoming sensory information to the relevant area in the cortex; specific central nuclei, which relay information from the cortex to the cerebellum; and non specific thalamic nuclei, which receive information from the reticular formation and project mainly to the neocortex (Shepherd, 1979). As the majority of the retinal ganglions terminate in the dLGN (dorsal lateral geniculate nucleus) of the thalamus, it is regarded as the gateway to

both the striate and extra striate areas in the visual cortex, these areas being the ones that are assumed to carry out so called cognitive processing, leading to eventual perception. It is important to note that the processing of visual information beyond the retina in higher mammals is not well understood at all (Dowling, 1992). It is estimated that 1,000,000 optic nerve fibres (retinal ganglions) terminate in the dLGN and the same number of neurons are thought to exist in the dLGN (Shepherd, 1979). Neurons in the dLGN have the same receptive field properties as retinal ganglions, responding in the same "on and off centre" way (Dowling, 1992). The existence of an almost 1:1 ratio between retinal ganglions and neurons in the dLGN as well as the retained receptive field organisation or "retinotopy" suggests that little processing is carried out by thalamic nuclei (Shepherd, 1979; Dowling, 1992). Retinotopic representation refers to the fact that the spatial relationship between retinal cells and neurons in the dLGN and further, in the cortex is maintained (Churchland, 1986). However anaesthetics used in the experimentation from which these conclusions about the role of the thalamus are drawn, would significantly effect the results, (Dowling, 1992). There are both neurophysiological and psychological theories of perception that implicate the thalamus as central to sensory processing (Mumford, 1991; 1992).

The structure of the dLGN reflects the intertwined relationship between structure and function in the brain. The complicated coding of sensory information at the retina is maintained in the dLGN As previously noted there are three types of retinal ganglions, W, X, and Y making up the optic nerve, two of which, X and Y, project mainly to the dLGN. This distinction

between different ganglions carrying different sorts of information is maintained in the structure of the dLGN. The dLGN, like the cortex, is constructed of layers of cells (lamina) of which there are six. Each lamina consists of a retinotopic map of one half of the visual field for each eye. Further, the left half of the visual field for each eye projects to the right side of the dLGN and the opposite for the right half of the visual field. (Bruce & Green, 1990). These lamina can be divided into two types, magnocellular and parvocellular, based upon the size of the cells in each layer. The X ganglions form synapses in the parvocellular layers and the Y ganglions form synapses in the magnocellular layers (see figure 5). This classification of cell types coding for different sorts of information lead to the development of a theory of parallel processing in the visual system (Stone, 1983) arising from this structural/functional distinction. This will be discussed in more detail in the next chapter. There is a further functional distinction represented by the architecture of the lamina in the dLGN, that is each eye projects to separate lamina. This is "ocularity" means that layers 1, 4 and 6 correspond to the opposite or "contralateral" eye and layers 2, 3 and 5 to the same side or "ipsilateral". This implies that cells dealing with binocular information do not occur at this stage in the processing (Dowling, 1992) (see figure 5). In simple terms there seems to be separate pathways for different sorts of information about the visual scene, and that is reflected in the encoding of information at the dLGN. An important question has to be how much information processing takes place at any particular point in the perception process? Arguably, the dLGN does not play a significant role in the extraction of information as little of the

retinal encoding is changed. If this is the case why then bother with this stage at all? Surely visual perception would be faster if ganglions projected straight to the next stage of processing. The question then is what is the function of the dLGN?

## THE STRIATE CORTEX.

The next stage in the primary visual pathway is the cortex. Since the early work carried out by Brodmann (Brodmann, 1909) anatomists have described the brain and subsequently the cortex in terms of distinct regions.

Historically the areas of the cortex involved in visual processing have undergone extensive investigation, resulting in detailed descriptions of their structure and related functionality. This has led to the development of what is referred to as “functional anatomy”, the classification of the brain in terms of both structure and the related function for given structures. The cortex is vast and unlike the rest of the brain does not divide so neatly into physically distinct parts. By adopting an approach based on functional segregation, distinctions are more easily made between areas in the cortex resulting in what is commonly referred to as its “functional architecture” (Hubel & Wiesel, 1962). Interestingly most reviews of functional anatomy agree that there are areas of the cortex that have a particular functional anatomy, but the exact nature of the underlying function of such structures is unknown (Churchland, 1986; Churchland & Sejnowski, 1992; Bruce & Green 1990; Dowling, 1992). Understanding the organisation of the cortex through the identification of distinct regions is possible using more than one methodology, for example patterns of connectivity, differences in

neuronal behaviours, lesioning, topographic mapping and chemical and structural differences (Rosenquest, 1985; van Essen, 1985). Despite this the cortex lends itself well to functional segregation and therefore mapping of the cortex continues to be dominated by this approach.

Initially the visual processing areas of the brain can be identified by tracing the flow of information through the visual pathways, notably the primary visual pathway. From the dLGN the ascending primary visual pathway synapses mainly in the area of the cortex known as the "Striate Cortex", also referred to as the "primary visual cortex", "area 17" and V1. For the purpose of clarification it will be referred to hence forth as V1. From here information passes on to other distinct cortical processing areas known collectively as the "extra striate" areas. There are direct projections from the dLGN to the extra striate areas such as V2, V3, V3a V4 and V5 (MT (medial temporal area)), however these are weaker in monkey's and humans than in other animals such as the cat (van Essen, 1985; 1991; Oram & Perrett, 1994; Churchland & Sejnowski, 1992; Bruce & Green 1990; Dowling, 1992). Arguably the primary visual pathway, terminating in V1 must be intact for "conscious" visual perception (Bruce & Green, 1990), supporting the theory that the primary visual pathway is the non-reflexive pathway and the visual cortex constitutes the cognitive processing areas of the brain. There is however, contradictory evidence that suggests that "perception" occurs even if damage to the primary visual pathway is evident as demonstrated by patients suffering from "blindsight" which will be discussed in more detail later in this chapter (Churchland, 1986; Bruce & Green, 1990; Cowey & Stoerig, 1992; Weiskrantz, 1986; Weiskrantz,

1992.). This does not necessarily support a counter argument, that the striate and extra striate areas are not essential for cognition or perception. As previously mentioned information flow does not rely entirely on the primary visual pathway.

Our present understanding of visual processing, and more importantly V1, originates with a collection of work, some of which has already been mentioned, conducted by Hubel and Wiesel (Hubel & Wiesel, 1962; Hubel & Wiesel 1970; Hubel, Wiesel, & Le Vay, 1977; Hubel & Wiesel, 1977.)

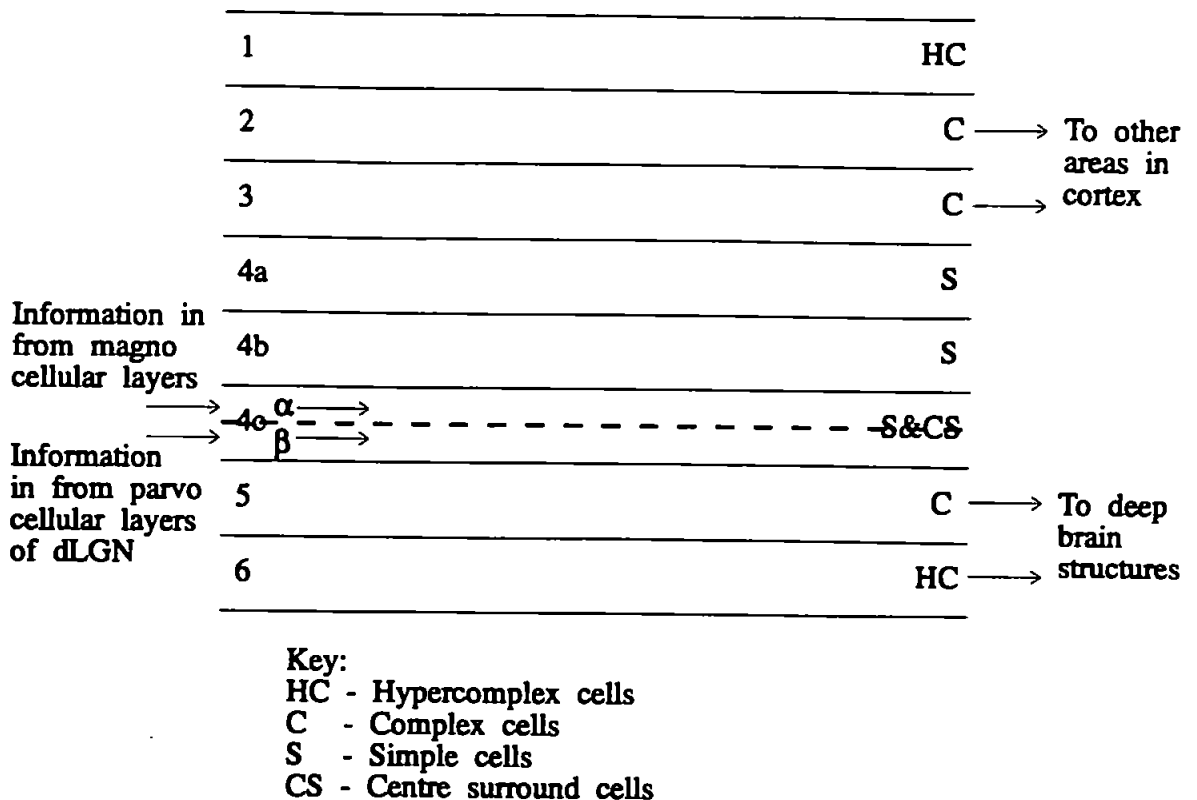
The work of Hubel and Wiesel was primarily experimental neurophysiology based on extra-cellular recordings, predominantly single cell, but also multi unit and staining studies, conducted mainly in V1. This initial work resulted in the formulation of a neurophysiological theory of perception which has been extensively developed and expanded by many researchers (Barlow, 1972; Perrett et al, 1986; Perrett et al 1987; Oram & Perrett, 1994).

Theories of visual perception will be discussed in more detail in the following chapter. Since the publication of the original work on the visual cortex, technology has allowed the development of more sophisticated methods of investigation. Research no longer relies completely on single or multiple cell recordings in animals. With the development of the positron emission topography (PET) scanners and magnetic resonance imaging (MRI) it is also possible to use human subjects. Technology has provided us with the ability to construct a complex and detailed understanding of the structure of the cortex (Oram & Perrett, 1994).

The visual cortex, like the dLGN is constructed of six main layers formed by a mass of cell bodies. Axons of the cells in the cortex mainly project

vertically but there are also a limited amount of “local” horizontal connections (Bruce & Green 1990; Churchland & Sejnowski, 1992). Information from the dLGN typically terminates in layer 4 in V1. As with the dLGN further functional distinctions can be made between the layers that make up V1. This distinction is not directly related to the input received by V1, rather it is a distinction relating to the function of the neurons within each layer. Layer 4 is where information enters the cortex, and can be divided into three sub layers, 4a 4b, and 4c. Magnocellular layers in the dLGN project mainly to the upper half of 4c referred to as  $4c\alpha$ , and parvocellular layers terminate mainly in the lower areas of 4c referred to as  $4c\beta$ . (Dowling, 1992). magnocellular projections also terminate in layer 4b and parvocellular projections also terminate in layers 2 and 3 which is where the clusters of colour sensitive cells are found (these clusters will be discussed later). This has important implications for parallel theories of visual processing that will be discussed in chapter 3. Layers 5 and 6 project mainly to “deep brain” structures, for example back to the dLGN and layers 2 and 3 as well as 4 project to the extra striate cortex and other cortical areas. As mentioned before both feedforward and feedback connections exist and are present in layers 2, 3, 4 and 6 (Oram & Perrett, 1994; Churchland & Sejnowski, 1992; Dowling, 1992) (see figure 6).





**Figure 6.** A schematic diagram showing the basic structure of the striate cortex (V1).

It is possible to distinguish between feedforward and feedback pathways due to the fact that feedforward, or ascending pathways arise in the superficial layers and terminate predominantly in layer 4. Feedback, or descending pathways arise in the deep layers and terminate in the both the superficial and deep layers (van Essen, 1985; Maunsell & Newsome, 1987). An even more general functional description of the visual cortex can be seen in the relationship between input to the cortex and the processing that takes place within it. It has been demonstrated that retinotopicy is maintained in at least 25 visual areas (Rosenquist, 1985). Further, that it is possible to distinguish between “first order transformations”, where the visual field is represented isomorphically with neurons in V1, and “second order transformations” where the relationship is not as strong, apparent in

the extra striate areas (Allman & Kaas, 1971; 1974).

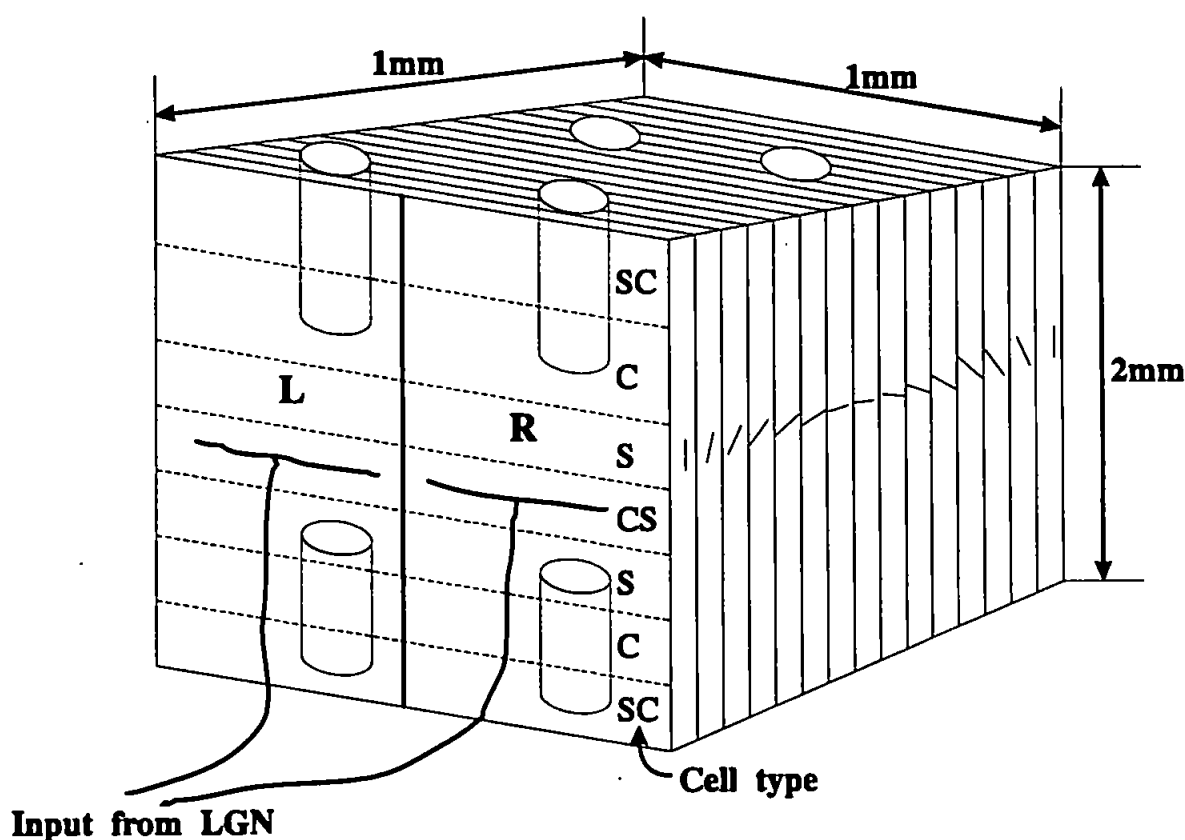
There are two main classes of cells in the visual cortex, pyramidal and stellate. Pyramidal cells project over distance to other areas of the cortex whereas stellate cells are project within the visual cortex. Stellate cells can be classified further into two sorts, “spiny” and “smooth”. Spiny stellates have lots of spines on their dendrites in comparison to smooth stellates which have very few or even none. A functional distinction can also be made, spiny stellates are excitatory and smooth stellates are inhibitory. All pyramidal cells are excitatory, and like spiny stellates have complex dendritic trees. (Dowling, 1992). Interestingly it is the dendritic “structural” feature that seems to determine the functional difference in cells, illustrating the inseparable nature of structure and function in the cortex. Layers 4 (a, b & c) of V1 mainly consists of stellate cells whereas pyramidal cells are found predominantly in the other layers. Layer 4c cells have the same receptive field “centre on” and “centre off” properties as retinal ganglions whereas cells in other layers do not demonstrate this in such a definite way. These cells are also monocular, responding to either the right or left eye but not both (Bruce & Green 1990). Pyramidal cells are generally binocular, responding to input from both eyes, and the further away from layer 4 they are the greater their binocularity (Dowling, 1992). These layer 4 cells have also been classified in terms of function as “simple cells” (Hubel & Wiesel, 1962; Hubel & Wiesel, 1977). These simple cells are thought to process simple information such as the detection of edges, bars and blobs, responding in an excitatory way. They are also particularly sensitive to orientation of lines in their receptive field. This information is

then passed on to the next layer which consists of “complex cells” (Hubel & Wiesel, 1962; Hubel & Wiesel, 1977) which like simple cells respond to edges, bars and blobs present in their receptive field with excitatory responses. They also extract more complex information such as geometric form in the shape of angles etc. and are sensitive to motion. Complex cells have a typically non-linear response pattern like Y ganglions, whereas simple cells respond in a linear fashion to light intensities like X ganglions. Finally there are “hypercomplex” cells” (Hubel & Wiesel, 1962; Hubel & Wiesel, 1977) which are again sensitive to bars edges and blobs, but they also perform something called “end stopping” where the inhibition occurs in the cell if its edge, bar or blob goes beyond the boundaries of its excitatory region.

The existence of hypercomplex cells is much debated, it has been suggested that complex and hypercomplex cells are the same, as complex cells are also sensitive to the ends of bars edges, and blobs. (Bruce & Green, 1990). It has been suggested that simple cells are in fact stellate cells whereas complex cells are pyramidal, the evidence however for such a distinction is not conclusive (Dowling, 1992). However if hypercomplex and complex cells are considered to be the same it is tempting to adopt this cell classification parallel. There is also evidence to suggest that cells in the visual cortex are sensitive to spatial frequency as well as geometric shapes (Campbell & Robson, 1968; De Valois et al 1982).

The relationship between structure and function in V1 is extremely complicated. As well as differences between single cells there are also structural divisions based on the function of groups of cells, “ocular

dominance” is an example of such a distinction. Ocular dominance concerns the cortical organisation of binocular information. As mentioned previously, until the cortex sensory input from each eye is kept separate. The segregation of information from right and left eye is maintained in V1 notably layer 4c where dLGN projections terminate cells prefer input from one or the other eye. These cells are grouped together in terms of eye preference in vertical columns roughly 0.5mm thick, alternating left eye, right eye (see figure 7).



**Figure 7.** A schematic diagram of the functional architecture of the striate cortex (V1).

However in other layers (4a, 4b, 1-3 & 5-6) this is not as easy to demonstrate as both simple and complex cells respond to both eyes. Each cell has a “preferred eye” but the structural organisation is less clear (Oram & Perrett, 1994; Churchland & Sejnowski, 1992; Dowling, 1992 Hubel &

Wiesel, 1962; Hubel & Wiesel 1970; Hubel, Wiesel, & Le Vay, 1977; Hubel & Wiesel, 1977).

A second functional anatomical distinction is evident in cell grouping which is most prevalent in all the layers except 4c, that of orientation selectivity. It seems that V1 can be organised into vertical columns of cells running through layers 1-4b and 5-6 that demonstrate a preference for sensory input lying in a given orientation. Changes of around  $10^\circ$  in the preferred orientation for a given column occurs roughly every 0.05mm continually eventually repeating themselves, across the length of the cortex (see figure 7). The combination of both ocular dominance column and orientation columns was identified as a single functional structure referred to as "hypercolumns" (Hubel & Wiesel, 1962; Hubel & Wiesel 1970; Hubel, Wiesel, & Le Vay, 1977; Hubel & Wiesel, 1977). Each hypercolumn consists of a block of cells running vertically through V1 about 1mm by 1mm by 2mm thick. A single hyper column incorporates 2 ocular dominance columns, one for each eye, roughly 20 orientations columns, one for each  $10^\circ$  orientation preference and both simple and complex cells (see figure 7). The borders of these hypercolumns overlap to a certain degree meaning that the changes across the cortex are graded rather than definite (Dowling, 1992; Bruce & Green, 1990).

There is a further functional structure apparent in this layer and column organisation related to colour vision. Although the conceptual model to be developed does not include a detailed description of colour processing in early visual processing, namely in the retina and V1, the processing of colour information plays a significant role in the relationship between

structure and function and our understanding of parallel processing within the visual system. It is therefore important to explain very simply some aspects of colour processing in V1 which clarify the nature of visual information flow in the so called “higher” visual processing areas of the cortex. It was discovered through staining studies in V1 that within the hypercolumns there were groups of colour sensitive cells prominent in layers 2 and 3 but also in layer 1, 5 and 6. This implies that layer 4 has no real colour processing abilities, and has important implications in parallel theories of the neurophysiology of vision to be discussed in the next chapter. They were referred to as “pegs” as they looked like small round pegs that had been driven vertically into the cortex through the previously mentioned layers (Dowling, 1992) (see figure 7). They were independent of ocular dominance or orientations columns, and on closer inspection were found to be constructed of colour sensitive cells (Hubel & Livingstone, 1983; Livingstone & Hubel, 1984). These “colour pegs” seem to be independent of the rest of processing in V1 and receive independent input from the dLGN (Livingstone & Hubel, 1982).

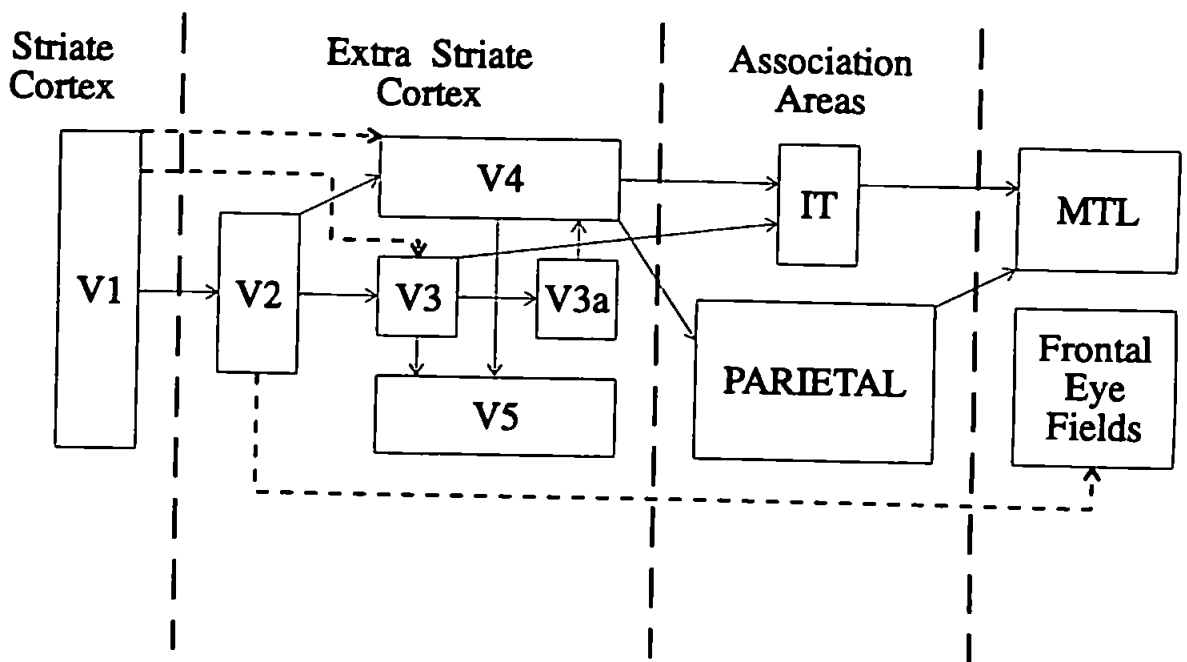
The columnar structure of the cortex particularly the hypercolumns indicate that specific regions of the cortex are dedicated to the processing of information from restricted inputs (Carlson, 1986). This further suggests that because synaptic connectivity in the cortex is mainly vertical, giving rise to this columnar structure, visual processing in V1 is mainly localised. A complete understanding of the visual scene is not well supported by neurophysiology within V1

It seems then that a highly complex pattern of information processing

emerges from the functional organisation of V1, which through simplification indicates specific information is extracted about the visual image at this stage in the processing. Arguably this “simplified picture” of V1 is not only incorrect, it detracts from the real capabilities of V1 in the processing of visual information (Gawne et al, 1994).

#### THE EXTRA-STRIATE AREAS.

At least 25 cortical areas associated with the processing of visual information have been identified in the primate, the striate cortex being one such area (van Essen et al, 1991). In the Macaque monkey it is possible to identify 11 areas that are dedicated primarily to the processing of visual information. There are 4 areas that have a marked preference for visual processing but also process other information. Finally there are 5 further areas that could be dedicated to either of the previous groups, and it is suggested many more remain to be discovered (van Essen, 1985). The areas involved with the processing of visual information beyond the striate cortex (V1) are commonly referred to as the “extra striate” or “pre striate cortex”. Beyond the extra striate cortex are other areas relating to vision namely the “association areas”, for example the parietal and temporal cortices (see figure 8), where it is thought that the resultant visual processing is integrated into a coherent whole.



**Figure 8.** A schematic of the extrastriate and association areas of the visual cortex.

Arguably the extra striate cortex is part of the association cortex, however the extra striate areas tend to be purely concerned with visual information processing whereas the association areas are not necessarily dedicated purely to visual processing (Carlson, 1986). The extra striate areas and the association areas are assumed to perform so called higher level functions or cognitive processing resulting in “conscious” perception. It is these areas and this high level processing that we know very little about (Churchland, 1986; Maunsell & Newsome, 1987). It could be argued that due to the significant amount of feedback from the extra striate areas to the striate cortex and dLGN the role of V1 in high level processing is underestimated. Unlike V1 the extra striate cortex consists not just of one cortical region, but of many, for example V2 and V3 are both visual cortical areas within the extra striate cortex. The identification of these extra striate areas



generally involved classic lesioning studies trying to trace the flow of information in the cortex. By destroying known visual areas, initially V1, and seeing which other areas were affected it is possible to "parcelate" the cortex. For example in lesioning V1, degeneration of another visual cortical area was apparent, V2. The next step was to lesion V2 and so on and so on (Rosenquist, 1985). Although identification of further distinct visual processing areas has been possible, their function is not as clearly defined. It is possible to make some kind of interpretation as to the underlying functionality of the extra striate areas, and some functional and structural themes are continued from V1.

As with the previous accounts of the visual system the outline given here of the extra striate cortex is primarily concerned with charting the information flow through the cortex identifying structures and attempting to define their function (see figure 8 for a simplified overview of the extra striate areas and their pattern of connectivity). The projections making up both the feedforward and feedback pathways from and to V1 are extremely complex. At least 100 pairs of major visual pathways have been identified along with many more suspected ones (Van Essen, 1985). It is only possible to discuss a limited number of the pathways that are significant in terms of a simplified model of visual perception (for a more comprehensive review of the patterns of connectivity see van Essen, 1985). Detailed accounts of structure and function will be discussed when necessary in further chapters. In simple terms then, from V1 the feedforward pathways project mainly to V2, although V1 projects directly to two other areas in the extra striate cortex, V3 and V4. V3 sends projections on to V3A, but V1 is not directly

connected to V3A. From V2 projections go to V3, V4 and also directly to the frontal eye fields. The frontal eye fields are located in the frontal lobes and involve the control of eye movement. V3 then projects to V3a, V5 (MT) and IT. V4 projects to IT and V5 to the posterior parietal cortex. In terms of functionality, V3 and V3a are sensitive to orientation but not colour. These are only the major pathways, there are many more. As mentioned previously retinotopy to some degree is maintained in all these areas but not to the same degree as in V1 (van Essen, 1985; Allman & Kaas, 1971; 1974).

The visual cortex, especially the extra striate areas have been shown to be organised in a hierarchical fashion (van Essen, 1985; Maunsell & Newsome, 1987). Further, it is suggested that at least two anatomically separate streams of information flow exist. The first attempt at classification of two stream of information suggested a dorsal and ventral pathway (Ungerleider & Mishkin, 1982; van Essen, 1985; Maunsell & Newsome, 1987; Bruce & Green, 1992; Oram & Perrett, 1994). The dorsal or “where” pathway proceeded from V1-V2-V5 (MT), terminating predominantly in the posterior parietal region and the ventral or “what” pathway, from V1-V2-V4, terminating in the inferior temporal (IT) regions of the cortex (Ungerleider & Mishkin, 1982). The parietal area of the cortex is involved in the processing of visio-spatial information and language and IT is the “highest” level in the association cortex of the monkey, thought to be the area of the cortex where object recognition takes place (Churchland & Sejnowski, 1992; Carlson, 1986). Through classic lesioning studies it is possible to support this theory of visual processing (Mishkin et al, 1983), but this will be discussed in the next chapter. It has been suggested that this distinction

between two separate pathways originates in the separation of retinal information at the dLGN into magno and parvo cellular layers which is maintained in V1 in the 4c and 4c $\beta$  input layers (Livingstone & Hubel, 1988). However the neurophysiological evidence for this is not conclusive (Goodale & Milner, 1992).

Beyond the parietal areas and IT, projections for both ventral and dorsal pathways go on to the medial temporal lobe. In the medial temporal lobe it seems that the pathways maintain their segregation. The medial temporal lobe includes the tail of the caudate, the claustrum, the amygdala and the hippocampus. These areas are thought to be involved in learning and memory processes as well as the generation of emotional states. It seems that visual processing does not have to reach IT or the parietal regions before it passes to the medial temporal lobe, there are direct projections from earlier visual processing areas.(see O' Shaughnessey, 1994; Oram & Perrett, 1994 for reviews). The implications of this are that object recognition and consequently visual perception could be taking place before visual processing is complete. This could mean that vision is not completely bottom-up, and explain why it takes longer to recognise things we have no experience of compared to things we have previously encountered. Oram and Perrett suggest that how much visual processing is necessary before information goes to the medial temporal areas is dependent on the task in hand. If cues are given as to the nature of the information that needs to be extracted then it takes less time than if no cue is given.

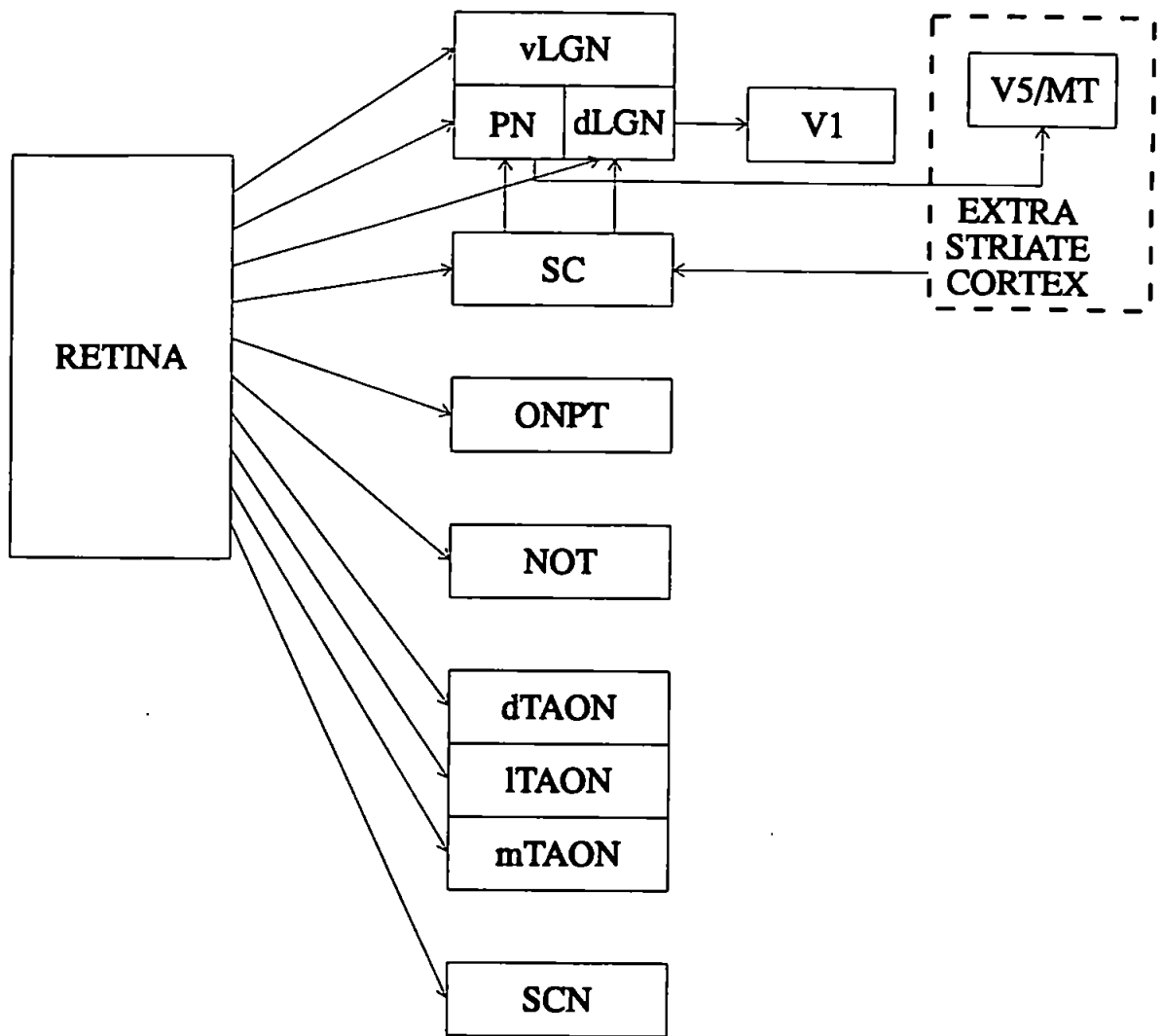
In terms of functionality the extra striate areas have both similarity and dissimilarity with V1. V2 is very much like V1, having the same structure

and functional distinctions, retinotopy is also maintained in V2. Like V1, V2 is sensitive to orientation as well as binocular disparity suggesting V2 is sensitive to depth information (Carlson, 1986). The architecture of V2 consists of thin and thick vertical stripes with faint interstripes separating them. It has been shown that the thin stripes are sensitive to wavelength and the thick stripes to motion. Form sensitivity is distributed across all of the stripes. (De Yoe & van Essen, 1985; Hubel & Livingstone, 1985; Shipp & Zeki, 1985). It is also clear that beyond V2 extra striate areas are tuned to respond to particular types of information (Zeki, 1978a; 1978b; 1980; 1983; 1988; 1992) Lesion studies have shown that by destroying certain areas in the extra striate cortex it is possible to remove the ability to process certain types of information. For example V3, which is anatomically separate from V3a but functionally similar, is sensitive to orientation but not at all to colour. Neurons in V3 have larger receptive fields than both neurons in V1 and V2 suggesting that a more global form of processing is taking place. Lesions to V4 demonstrate inability to perceive colour, and lesions to V5 (MT) show inability to detect motion. Since the development of an anatomical and functional basis for two separate streams of information processing in the cortex there has been extensive revisions. It is now clear that the early distinctions are not fully supported by neurophysiology and it has been suggested that there are in fact 4 separate pathways in the extra striate cortex processing information. one for colour, one for motion and two for form (Zeki, 1978a; 1978b; 1980; 1983; 1988; 1992; van Essen, 1985; Maunsell & Newsome, 1987; Bruce & Green, 1992; Oram & Perrett, 1994). Again this will be discussed further in the following chapter.

Therefore it is possible to create a very simplistic schematic idea of visual processing from V1. V2 can be thought of as a parcelling point where, as in V1, a retinotopic representation of the image is directed to various areas in the extra striate cortex that process certain information. There are at least two main pathways through the extra striate cortex operating in parallel with a possible third and fourth, processing colour, form and motion. Two of the pathways can be seen to originate in V1 where colour and motion information are processed separately, colour in the “pegs” and motion on direction selective cells. These direction sensitive cells are found mainly in layer 4b which projects to V5 (MT) and V3 and in layer 6 which is known to project directly to V5 (MT) where motion information is thought to be processed (Livingstone & Hubel, 1988; Zeki, 1978a; 1978b; 1980; 1983; 1988; 1992; van Essen, 1985; Maunsell & Newsome, 1987; Bruce & Green, 1992; Oram & Perrett, 1994). It is possible to make a further anatomical distinction between the extra striate areas and the association areas of the cortex. Initially there are the areas that are still concerned with the processing of the retinal image, extracting information. Beyond this are further areas that are concerned with the integration of information and possibly learning and recognition of the result of such integration. Areas considered to involve information extraction include; V2, V3, V3a & V4 and V5. The areas thought to be significant in the integration of all this visual information include the inferior temporal cortex (IT), and the posterior parietal cortex (see figure 8).

## **OTHER PATHWAYS: REFLEXIVE “NON COGNITIVE” PATHWAYS.**

As mentioned previously there are a number of other pathways from the retina that do not pass through the dLGN. These pathways do not all project to the visual cortex, some project to the brain stem, and play a significant role in the regularity system, controlling for example eye movement and focus. These secondary or reflexive, non cognitive pathways include retinal projections to, the superior colliculus (SC), pulvinar nucleus (PN), ventral lateral geniculate nucleus (vLGN), olivary nucleus of pre-tectum (ONPT), nucleus optic tract (NOT), dorsal terminal accessory optic nucleus (dTAON), lateral terminal accessory optic nucleus (lTAON), medial terminal accessory optic nucleus (mTAON), and the suprachiasmatic nucleus (SCN) (see figure 9). Pathways that project to the cortex include the SC and the PN whereas pathways to other areas such as the brain stem include the vLGN, ONPT, NOT, d, l, and mTAON and the SCN (Cowey & Stoerig, 1992; Carlson, 1986).



**Figure 9.** The secondary, reflexive visual pathways.

The superior colliculus (SC) as discussed previously, receives mainly W ganglion projections from the retina and is thought to play a major role in the control of eye movements. The SC then projects to the extra striate areas via the inter-laminar areas of the dLGN, and the pulvinar nucleus (PN). The SC receives information from nearly all of the areas in the visual cortex. The PN is one of the many nuclei that make up the thalamus and it both projects to and receives projections from the extra striate areas, namely MT. MT being concerned with the processing of motion information. The pathways projecting to the SC and the PN are inextricably

linked, it is therefore hard to determine exactly what information is transmitted by the PN. The vLGN projects retinal information not to the cortex, but to sub-cortical structures including structures that already form secondary visual pathways of their own such as the pretectum, SC, and the SCN. Its role in the processing of visual information is not at all clear. The ONPT is thought to be involved with the control of pupil size and the accessory optic nuclei, d, l, and mTAON, in the co-ordination of eye movements and the detection of self movement. Finally the SCN which is part of the hypothalamus involves the regulation of the day/night cycle and is sometimes thought of as a biological clock. (Carlson, 1986; Cowey & Stoerig, 1992).

The role of these secondary visual pathways in relation to the processing of visual information resulting in perception not clear. As already suggested one distinction that has been made between these secondary visual pathways and the primary visual pathway is that of conscious/non conscious, reflexive/non-reflexive processing. However as it is clear some of these so called secondary reflexive pathways project to and receive projections from the extra striate cortex, the very areas of the visual cortex with which we associate with the processing of higher cognitive "conscious" processes. Experimentation with patients suffering with the disorder "blindsight" has suggested that the information enabling conscious perception could be carried along these pathways. Blindsight is a condition where part of the visual cortex, usually V1 has been destroyed either through illness in human subjects, or through lesions in animal studies, making it impossible for them to recognise or respond to visual stimuli. It has been shown that if



subjects suffering from blindsight are asked to “guess” where objects are they can correctly locate them even though they are clinically “blind” (Weiskrantz, 1986; 1992). However the evidence is not conclusive and there is much work to do before the role of these secondary pathways is more clearly defined (Cowey & Stoerig, 1992).

## CONCLUSIONS.

There cannot be enough emphasis on the point made earlier (in chapter 1) that a simplifying brain model is the best that can be hoped to be achieved. The outline of the neurophysiology of the visual system presented here is a gross simplification. The information available in the neurophysiological literature is far more detailed and complex. Even though it is possible to pursue more and more detailed accounts of the neurophysiology of the visual system, information is still incomplete. The problem of incomplete information is directly related to the techniques available in experimental neuroscience. Intracellular recordings are extremely hard to conduct successfully, and extracellular techniques are less reliable.

Neurophysiological research is continually providing new information about the structure and function of the brain, meaning that any model developed is by no means complete. As research continues finer levels of detail are achieved, going beyond neurophysiological descriptions, to include biochemical descriptions and even to the level of molecular physics. At this point the problem of multi-disciplinary research becomes most apparent. To provide a global picture of what is actually happening causes an overload of information of which a great deal is meaningless without in depth

knowledge in given disciplines.

It seems from the amount of literature published that more is known about the structure and function of the visual cortex than any other cortical area. It seems that even with this wealth of knowledge we may know a great deal about visual processing, but we know very little about how visual processing becomes visual perception (Oram & Perrett, 1994). This makes the task of developing artificial vision systems capable of perception extremely difficult.

# **CHAPTER 3**

## **THEORIES OF VISUAL PERCEPTION:**

### **NEUROPHYSIOLOGICAL AND PSYCHOLOGICAL THEORIES OF VISION.**

*This chapter introduces theories of visual processing and visual perception that originate in both neurophysiology and psychology. The first part of the chapter discusses neurophysiological, bottom-up theories of both visual processing and visual perception. The emphasis being on theories that explain the processing of visual information. The second part of the chapter discusses psychological theories some of which originate in neurophysiology others are pure top-down approaches. The difference from neurophysiological theories being that psychological theories aim to explain global perception. However overlap occurs between psychological and neurophysiological theories especially in regard to explaining local perception.*

#### **INTRODUCTION.**

#### **PSYCHOLOGICAL AND NEUROPHYSIOLOGICAL THEORIES.**

It is not always a simple task to distinguish between neurophysiological and psychological theories of visual perception. This is because some psychological theories have grown from accounts of neurophysiology, taking ideas suggested in the structure and function of the visual system and

using them to explain behaviour. Conversely some neurophysiological theories originate within the field of neuropsychology, using behavioural accounts to develop theories of brain structure and function and supporting them through classical neurophysiological techniques. With neurophysiological theories it is often hard to decide whether what is being proposed by any piece of research is actually an account of either a localised or global aspect of the neurophysiology of vision or an actual theory supported by such descriptions. With psychological approaches it is often hard to see how behavioural accounts in the form of psychophysical experimentation can be related to the structure and function of the underlying neurophysiology.

One possible distinction that can be made is that neurophysiological theories of visual perception tend to be based on accounts of visual processing, whereas psychological theories are based more on visual perception. This equates to psychology taking a top-down approach to vision, having a theory of what vision is then testing that theory against behavioural experiments. Neurophysiology on the other hand works from a bottom-up approach. This means constructing a theory of visual perception by starting at its lowest level, and building on this until a theory emerges. The aim of this chapter is to illustrate this point and also to combine the best of the two approaches in the development of a theory of perception. Most attempts at theorising about or modelling perception are directed at a specific visual task, especially top-down psychological approaches. For example psychological theories of vision include theories of object or pattern recognition, depth perception, motion perception, object constancy

and theories of selective attention. There are very few “global” theories of vision, most seem to be task orientated, possibly with the exception of Marr’s and Gibson’s theories respectively (Marr, 1982; Gibson, 1950). Arguably such theories are really frameworks within which visual perception is investigated (Bruce & Green, 1990). Neurophysiological theories of vision tend to be directed more at accounts of how the visual system might work. Rather than having different theories argued in conjunction with each other theories tend to be updated or replaced as new information comes to light. These like psychological theories of perception tend to be task orientated, for example face detection, or colour processing. It does seem that neurophysiological theories attempt to be global but within that global theory of visual perception a specific aspect of visual processing is focused on.

## **NEUROPHYSIOLOGICAL THEORIES OF VISION.**

Two main types of theory emerge from the structural and functional descriptions of the visual processing areas of the brain. Firstly there are those theories which are based on the notion of vision as a hierarchical, sequential process. Secondly there are those theories that suggest visual processing is to some extent hierarchical but there are separate and distinct streams of visual processing that occur in parallel. Therefore it possible to class neurophysiological theories of vision as being either of one sort or the other. Underlying either class of theory is a further functional distinction related to the coding of information. In hierarchical sequential theories of

visual processing each neuron in V1 is thought to encode a specific piece of information about the visual image. As processing continues through the visual system neurons encode for more and more abstract pieces of information, until at the highest level there are neurons that respond only to specific combinations of information in the form of whole objects. With parallel theories of visual processing encoding is thought to be based on populations of neurons all contributing to the processing of the visual image, with no one neuron in any visual area having complete information on any aspect of the visual image. Neurons in V1 for example encode form information, whereas neurons in V4 encode colour information. The image is encoded as a pattern of activity across populations of cells.(Churchland & Sejnowski, 1992). Exactly how this “distributed” representation becomes a coherent whole is the subject of much debate and is discussed in more detail later in this chapter.

As the visual system can be divided into a number of separate stages of processing involving anatomically distinct areas of the brain there are also localised theories of vision. These theories may relate only to a particular area of visual processing rather than providing a global model or theory of perception itself. However such theories can provide insight into the way in which global theories of visual processing can be constructed.

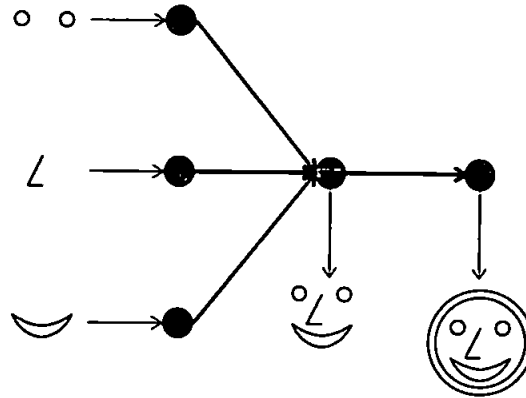
#### **BARLOW'S SINGLE CELL FEATURE DETECTION THEORY.**

The single cell feature detection (SCFD) theory of visual processing and eventual perception evolved from the work carried out by Hubel and Wiesel on the structure and function of the visual cortex as discussed in chapter 2

(Hubel & Wiesel, 1962; Hubel & Wiesel 1970; Hubel, Wiesel, & Le Vay, 1977; Hubel & Wiesel, 1977). The theory is based on Hubel and Wiesel's description of visual processing from the retinal ganglions to neurons in V1. The first level of detection being light intensities in the retinal ganglions terminating in the dLGN. Feature detection becomes more abstracted as simple cells in V1 code for edges and orientation in a given retinal location. The third level of abstraction is with complex cells that are not bound by retinotopic representation (Bruce & Green, 1990; Churchland, 1986). This is supported by the laminar columnar structure of V1, and the strict organisation of information flow within that structure, as well as the existence of the primary visual pathway from the retina to V1. Although Hubel and Wiesel themselves did not propose that visual processing throughout the cortex exists in this way, their work provided neurophysiology with the evidence it needed for a single cell doctrine to develop.

Quite simply the SCFD theory suggests that information is processed in a series of hierarchical sequential steps resulting in a more and more sophisticated description leading to the "Grandmother Cell" hypotheses proposed by Barlow (Barlow, 1972). Barlow suggests that the brain has single cells acting as feature detectors in the cortex responding to more and more abstract geometric features until at the top of the hierarchy they respond only to particular objects, for example one's grandmother. When a particular feature falls into the receptive field of a neuron in layer 4 of V1 tuned to that feature it fires. This neuron projects to those in a layer directly above it which is tuned to respond to an even more abstract feature

until eventually it projects to an area in the extra striate cortex that actually contains a neuron that fires only when a certain persons face has been detected, maybe ones grandmother (see figure 10). The problem is whether this theory can be upheld by neurophysiological evidence.



**Figure 10.** Single cell feature detection theory.

If the SCFD holds true then it would be expected that the transmission of information from the dLGN to the striate cortex would be solely feedforward, and this would be reflected in the structure of the primary visual pathway (Bruce and Green, 1990). Initially it would be expected that if this was the case then dLGN cells would connect directly and only to simple cells in V1. It has been shown that projections from the dLGN connect to both simple and complex cells in V1. Complex cells also receive projections from other cortical cells. It must be noted however that simple cells do only get information in from the dLGN being predominantly in layer 4 of V1 and complex cells do provide output from V1 (Ferster & Lindstrom, 1983). One major problem with the SCFD theory is that there are more feedback pathways than feedforward ones, so as much goes back to the dLGN as comes in (van Essen, 1985; Oram & Perrett, 1994; Churchland & Sejnowski, 1992; Dowling, 1992). However the organisation of the visual



cortex has been shown to be hierarchical in that information entering the cortex proceeds in a feedforward way and is processed in stages. Although specific connection patterns may not be upheld there is a consensus that the cortex is to some extent organised in a hierarchical fashion (see van Essen, 1985; for a review).

Another problem for the SCFD theory is that it assumes that single neurons in V1 code for geometric features. A pattern of light falls on the retina and is coded as an abstract symbolic representation of the retinal image which consists of simple geometric features present such as line angles etc.

Neurons in the visual cortex then respond in a hierarchy to more and more complex patterns of these geometric features. The centre-on cells for example in the LGN do not just respond to straight lines they can respond to any number of patterns. To be a line detector they would have to respond only to straight lines (Bruce and Green, 1990) Bruce and Green argue that proponents of the SCFD theory would suggest that simple cells in V1 respond to lines of cells in the LGN not just single centre-on cells.

Therefore simple cells in the next stage of the hierarchy are coding for geometric features. However simple cells do respond to variations in line orientation in the retinal image, as coded by lines of centre-on cells, so detecting different lines. This is not the whole picture, the response of simple cells increases with other variables such as contrast and position of the light falling in its corresponding receptive field. Therefore the information it imparts is ambiguous (Bruce & Green, 1990).

The SCFD theory also has been shown to be flawed through classic lesion studies. If the theory holds true it would be expected that if V1 was

removed through lesions, object and complex pattern recognition would not be possible. In a series of experiments it has been consistently shown that this is not the case. However in the same series of experiments it was shown that lesions to the extra striate cortex do impair object and pattern recognition, providing support for the SCFD theories claims that whole objects are represented at the highest level in the hierarchy, namely IT in the extra striate areas. (Sprague et al, 1977; Berlucchi & Sprague, 1981). This work also casts doubt on the significance of the spatial frequency tuning of neurons in V1 in contour processing for pattern recognition (Berlucchi & Sprague, 1981). The results of Berlucchi and Sprague support this in that visual acuity is lost with V1, however pattern recognition is not. A possible explanation for this is that the secondary visual pathways are sufficient to transmit information to the extra striate areas. This is a contentious issue and has been investigated extensively in relation to blindsight patients who when pressed to make perceptual "guesses" about objects do so correctly even though they cannot see them. However it is still not certain as to how much information can be passed on via the secondary pathways (see Cowey & Stoerig, 1992; for a review). One suggestion for blindsight is that it is very rarely the whole of V1 that is destroyed, and what is left is sufficient to process images to a limited degree (Campion et al, 1983). The evidence in support of the SCFD theory is not conclusive though the single cell doctrine dominated both theories of visual processing and experimental research in neurophysiology for a long time.

## THEORIES OF PARALLEL PROCESSING IN THE VISUAL SYSTEM.

More recently and as a result of conflicting evidence for single cell theories and extensive investigation into the relationship between the striate and extra striate areas of the cortex, theories of visual processing have been readdressed. It has become clear that beyond V1 processing in the extra striate cortex is not purely hierarchical and by no means a sequential process. By looking at specific aspects of visual processing, for example the processing of colour, motion, shape, and size information, research has shown that there are separate areas for processing this information beyond V1. Destruction, either artificial or through illness and disease, of certain areas of the cortex supports this. Zeki and his colleagues have conducted a vast number of experiments that show that the extra striate areas of the cortex are specialised in their function (Zeki, 1978a; 1978b; 1980; 1983; 1988; 1992). As mentioned in chapter 2, Zeki showed that V5 (MT) is specialised for motion detection, and cells in V4 are particularly sensitive to wavelengths of light used in colour identification as well as orientation for detection of form. Neurons in V3 and V3a are also tuned to form information but do not respond to colour information. This lead to the development of a functional specialisation hypothesis which has also extended to V1. Zeki suggested that both V1 and V2 are also specialised in their function, and they act as a “parcelling” point where information is grouped and sent to the relevant specialised area in the extra striate cortex for processing (Zeki, 1978a; 1978b; 1980; 1983; 1988; especially 1992 for review). This hypothesis is supported by the structural and functional distinctions in V1 where, as mentioned in chapter 2, colour and motion

information is coded separately.

## UNGERLIEDER & MISHKIN'S TWO STREAM THEORY OF VISUAL PROCESSING.

The SCFD theory was predominantly hierarchical, and there is evidence suggesting this to be the case in the primate visual cortex with several areas of processing making up each level with extensive feedback from higher to lower levels (van Essen, 1985). The existence of a hierarchy does not mean serial processing feature detection theory is right. Especially as SCFD theory claims processing to be feedforward, and evidence to support hierarchical processing suggests that feedback connections are equally as important. The parcelling model suggests different areas code for different things. V1 and V2 "parcel" information out to areas that code for different things in parallel (Zeki, 1978a; 1978b; 1980; 1983; 1988; 1992). So within the bounds of the striate cortex functional and structural evidence lend support to single cell accounts of visual processing. However when looking beyond V1 to the extra striate areas the evidence is not so convincing. The failure of the SCFD to explain why pattern and object recognition was still possible after lesions to V1 (Sprague et al, 1977; Berlucchi & Sprague, 1981) along with the evidence for functional specialisation within the extra striate areas (Zeki, 1978a; 1978b; 1980; 1983; 1988; 1992) led to the call for a parallel theory of visual processing.

As mentioned in chapter 2 Ungerleider and Mishkin proposed a two stream theory of visual processing based on two different visual processing capabilities, recognition and visio-spatial. One pathway determines what an

object is and another determines where it is. There theory was based on a proposal by Schneider (Schneider, 1969) who suggested that there were separate sites dedicated to object recognition and object location.

Ungerlieder and Mishkin conducted a series of experiments that enabled them to support such an idea as well as give it substance through proposing the exact anatomical pathways for information flow, ventral and dorsal, to two separate processing sites, IT and the parietal areas of the cortex. The theory is also supported by neuropsychological studies. Patients suffering from visual agnosia due to damage to the occipitotemporal regions of the cortex are unable to identify objects yet can still manipulate them, conversely patients suffering from optic ataxia due to damage to the posterior parietal cortex are able to identify but not move objects (see Goodale & Milner, 1992 for review).

More recently the two stream theory has been substantially reviewed in the light of neuropsychological studies that suggest the distinction between “what” and “where” pathways is not quite so straight forward (Goodale & Milner, 1992). They suggest that neuropsychological evidence used to support the original two stream theory is conflicting. In some cases of patients suffering from optic ataxia they are unable to use the information relating to grasping rather than being completely incapable of doing so. Goodale and Milner suggest that the “where” pathway should be renamed the “how” pathway. They suggest that two systems exist, one relating to predominantly to perception and the other to action, both processing visual information in different ways. There is a certain amount of interaction between the two systems through connectivity, meaning that it is not as

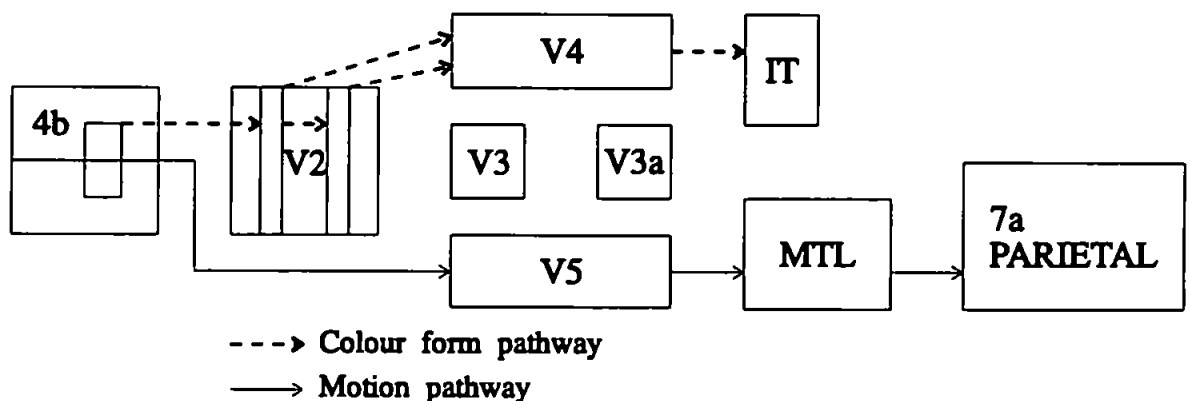
easy to separate perception and action as suggested by previous models (Goodale, & Milner, 1992).

#### MAUNSELL & NEWSOME'S PARALLEL MODEL OF VISUAL PROCESSING.

Ungerleider and Mishkin's two stream theory was primarily concerned with visual capabilities. It also suggested that divergence of the two streams took place early in the extra striate areas (Maunsell & Newsome, 1987). Revision of the two stream theory taking account of early vision was proposed by Maunsell and Newsome. (Maunsell & Newsome, 1987). The two stream theory they proposed was based on the formulation of possible pathways through the extra striate cortex in view of Zeki's work on functional specialisation. Maunsell and Newsome propose that there are two main pathways functioning in parallel, rather than lots of different areas all processing different things as the parcelling theory suggests. Both pathways begin in V1, and it is suggested even earlier in the magnocellular/parvocellular separation in the dLGN, but as mentioned in chapter 2 the basis for such a distinction is not clear. One pathway processes motion and spatial information resulting in visio-spatial processing, whilst the other processes colour and form information resulting in object recognition. Each stream is organised hierarchically, with differing levels of processing in each of the relevant areas associated with that particular stream. Maunsell and Newsome refer to the two pathways as the "colour form pathway" and the "motion pathway" and emphasise the point that they are related to but not the same as Ungerleider and Mishkin's

streams of processing (Maunsell and Newsome, 1987).

The motion pathway originates in V1 in layer 4b which projects directly to V5 (MT). From V5 projections go to the medial superior temporal area (MST) and then to area 7a in the parietal cortex. The colour form pathway originates in the colour “pegs” in V1 and the “interpeg” regions that are sensitive to orientation. This segregation is maintained in V2 where the pegs in V1 project to the thin stripes in V2 which are sensitive to colour. The interpegs project to the inter stripes in V2 which prefer a similar orientation preference as in V1. From V2 projections go to V4 then to IT, specifically the posterior inferotemporal area (PIT) and to the anterior inferotemporal area (AIT). In both pathways retinotopy decreases as they move away from V1 and evidence suggests that as the pathways progress deeper into the extra striate areas more complex information is developed. Finally Maunsell and Newsome cite experimental work that suggests extra retinal inputs via the secondary visual pathways to the parietal and IT regions of the cortex play a significant role in the information processing that takes place there and should not be dismissed as merely modulating visual processing (see figure 11).



**Figure 11.** A schematic diagram illustrating Maunsell and Newsome's two stream theory of visual processing.

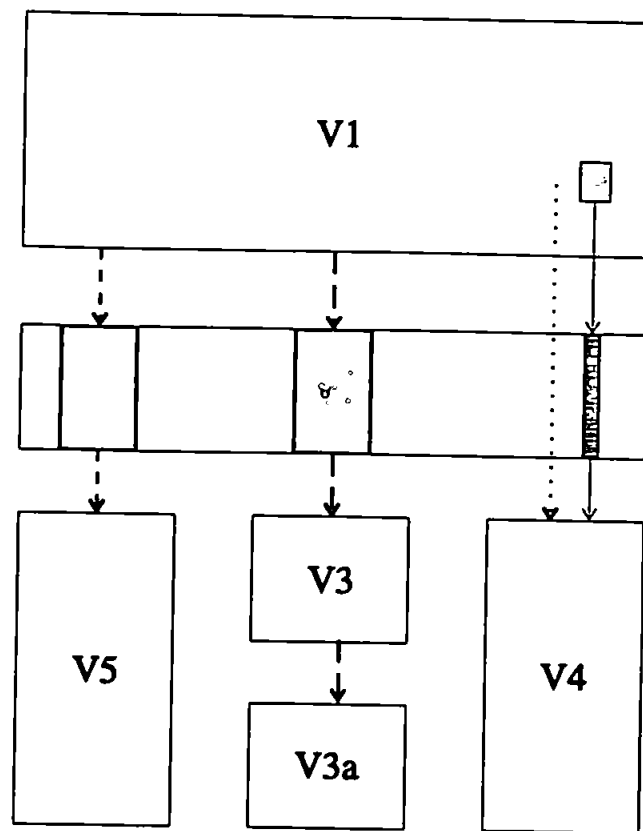
Maunsell and Newsome make it quite clear in their paper that the extrastriate areas are still very much under investigation and that some areas have been the subject of intensive study whilst others have not been scrutinised at all well. They go on to suggest that the two streams of processing are not totally independent of each other either functionally or anatomically. The colour form pathway does contain direction sensitive cells and orientation specificity is present in the motion pathway. Also there is evidence for connections between V4 and V5 and evidence from psychophysics that suggests completely independent processing does not exist. They suggest visual processing consists of two streams of serially connected visual processing areas that are in no way completely sequential due to feedback and diverse connection patterns. The complexity of the functional and structural relationship between the two streams is such that it is not possible to demonstrate conclusively that they operate completely independently of each other (Maunsell & Newsome, 1987). It seems then that although it is possible to regard visual processing in terms of parallel streams of processing there is a certain amount of overlap.

#### **ZEKI'S FOUR STREAM THEORY OF VISUAL PROCESSING.**

Zeki has revised the two stream theory further to fit with recent neurophysiological evidence from positron emission tomography (PET) scans. The results of PET scans carried out in conjunction with visual experimental tasks support Zeki's parcellation hypothesis (Zeki, 1992). Zeki proposes four parallel pathways operate in the visual cortex, one for colour,



one for motion and two for form. Zeki suggests there is an overlap between the pathways but the motion and colour pathways are quite distinct, having origins in the magnocellular parvocellular layers in the dLGN and in V1. One of the two form pathways is linked to the colour pathway originating in the parvocellular layer in the dLGN and passing to V4 via the areas between the colour pegs in V1 (interpegs) and the thin stripes in V2. Then second form pathway is independent of colour, and follows the magnocellular pathway via layer 4b in V1 to the thick stripes in V2 to V3 and directly to V3 from layer 4b in V1. Zeki refers to this as the “dynamic form” pathway as it is thought that V3 processes form and motion information. Zeki argues that the “pigeon hole” nature of V1 and V2’s role in processing and direct connections between extra striate visual areas allows for complete interaction between pathways even though they can be seen to be distinct (Zeki, 1992) (see figure 12).



Pathways:-  
 ———→ Colour  
 ----→ Motion  
 .....→ Colour/Form  
 - - -> Dynamic Form

**Figure 12.** A schematic diagram illustrating Zeki's four stream theory of visual processing.

Support for the four stream theory comes not only from existing structural and functional experimental evidence but also from neuropsychology.

Damage to V4 results in achromatopsia, where the patient still has a fully functioning V1, but is not able to see colour. Interestingly patients are also unable to recall colour information from before the damage occurred. These patients are completely capable of form and motion perception. In contrast patients with damage to V5 resulting in akinetopsia, are unable to see things when they are in motion. The form pathways are less well supported by

neuropsychology, however some achromatopsia patients with damage in the striate cortex cannot perceive form unless either they or the visual world is in motion. Zeki argues that this is because they have to rely on the dynamic form pathway rather than through the damaged colour form pathway.

Strikingly in support of the role of V1 and V2 in the four stream theory is a syndrome that Zeki has termed “chromatopsia of carbon monoxide poisoning”. People who suffer the effects of smoke or fume inhalation tend to have visual problems but their colour vision is not affected at all. In fact colour becomes the dominant medium for object recognition. Zeki suggests this is due to the fact that the colour pegs in V1 and the thin stripes in V2 have a high density of oxygen supplying blood vessels meaning that oxygen is less likely to be deprived to these areas than in others in the event of fume inhalation (Zeki, 1992).

Zeki also proposes that V1 is needed for long term visual experience, suggesting that V1 and V2 are as important to perception as the extra striate areas that are assumed to underlie cognition. Zeki suggests that V1 is used for short term perception in that it initialises the processing of visual information which is then processed in its specialist area and then passed back to V1 and V2 for further specialist analysis. He uses blindsight patients as support for this idea. Blindsight patients cannot see as such but Zeki argues secondary pathways allow enough information to get to the specialised areas for them to “guess” correctly. They cannot add detailed specific information but can guess at motion and colour.

There are other neuropsychological disorders providing support for parallel stream theories such as Zeki's. However it must be noted that

neuropsychological deficits and savings are generally partial. This means that although there may be a distinct deficit in certain processing abilities these are not necessarily completely abolished. Obviously individual cases differ due to the extent and nature of damage. Further, deficits in the processing of one particular type of information processing generally means a deficit is present in another. The nature of the cortex is such that damage occurring to certain areas will affect the processing capabilities of others. Other disorders providing support include akinetopsic patients who cannot tell direction or coherence of motion but can still determine the existence of motion (Baker et al, 1987; Baker et al, 1991). However as mentioned above, patients suffering with akinetopsia generally have impaired form processing as well. Similarly achromatopsics can interpret waveform information but not as colour. Patients in both disorders still had an intact V1, therefore Zeki implies V1 is not just important in terms of providing detailed form information but for visual perception as a whole process. Conversely serious damage to extra striate regions but not to V1 causes problems in long term perceptual abilities. Zeki cites the case of a man whose V1 was intact but had severe stroke damage to the extra striate areas. He could copy drawings in incredible detail but had no idea what he had drawn. This suggest that V1 and V2 are not just simply parcelling points, their role in perception is complex. Zeki parallels this segregation in the cortex to memory, initial processing can be used like a "working memory" (Baddeley & Hitch, 1974) whereas later processing is essential for long term storage (Zeki, 1992).

This also supports Zeki's earlier findings that suggested the parallel

pathways operated hierarchically extracting more and more abstract information as they progressed. There is evidence in support of hierarchical processing within the colour form pathway, where it has been demonstrated that cells in V1 are selective for wavelength and respond even to changes in wavelength brought about by changes in lighting conditions. Cells in V4 however do not change their response in this way, they demonstrate colour constancy. This ensures that an object that is red remains red even when lighting conditions change. This suggests that colour processing in V4 is more abstract than that in V1 (Zeki, 1983). The notion of hierarchical processing does not mean that early visual processing is not essential for visual perception rather it suggests that higher visual processing is less dependent on retinal processing.

One important question that has to be addressed in any parallel theories of vision is the integration of information allowing perception to take place. SCFD theory does not call for information to be combined. Zeki says there is no one "master" area where information converges, claiming there is no anatomical evidence to support such an idea. Zeki says integration occurs through the many links right from dLGN upwards that run between areas and forward and back in cortex. He illustrates this by pointing out that V5 detects motion, but its receptive fields are too large to pinpoint the exact location of what is moving. It is by sending information back to V1, that the visual system is able to determine exactly what and where objects are moving. In reverse Zeki argues this is also an explanation of illusions. V1 does not fall for illusions. It is the larger receptive fields and increased processing power of V2 that perceives things to be present that are not. This

led Zeki to formulate his theory of "multistage integration" in which perception takes place across the visual system simultaneously. Perception is mediated by vast re-entrant (feedback) connections to all areas of the visual cortex. Forward connections are "patchy and discrete" but feedback connections are non-specific and diffuse. This means that a feedback system could perform several important functions including the integration of form, colour and motion information, and provide detailed spatial information for extra striate processing (Zeki, 1988; 1992).

#### EVIDENCE FOR INTEGRATED THEORIES OF PERCEPTION.

Contrary to theories of segregated visual processing there is evidence to suggest that there is a great deal of interaction between processing channels in the visual system. As mentioned previously neuropsychological deficits are not generally distinct. When one form of processing is impaired other perceptual capabilities are usually also affected. Akienotopic patients who have damage to the motion pathways also suffer from form perception impairments. This implies that form processing is not necessarily distinct from motion processing.

The neurophysiological and psychological literature also supports the distribution of processing between processing channels. It has been shown that the chromatic properties of visual stimuli effect the detection of motion (Krauskopf & Farell, 1990). Psychophysical research has also shown that the processing of motion information seems to be dependent on form information. In a series of experiments Adelson and Movshon showed that the coherent motion perception of two spatial gratings was facilitated if the

two gratings had similar spatial frequencies and were of a similar contrast (Adelson & Movshon, 1982: 1984). This has been interpreted as being evidence that segmentation of the object is directly responsible for coherent motion perception (Albright, 1991).

Neurophysiological accounts of perception that suggest that visual processing is carried out in separate channels acknowledge that integration within the visual cortex occurs at all stages. However parallel theories of visual processing are concerned with the overwhelming evidence that supports segregated streams of processing within the visual system. This does not mean that the processing of form and motion information is not an integrated process, rather that is predominantly segregated.

#### PERRETT AND COLLEAGUES -THE SINGLE CELL REVISITED.

Further, and more recent support for the SCFD theory and hierarchical processing in the visual system comes from work conducted by Perrett et al (Perrett et al , 1986; Perrett et al, 1987; Oram and Perrett, 1994), based on earlier findings by Gross et al (Gross et al, 1972). Perrett et al have extensively investigated cells in IT that have been found to be selectively responsive to faces. (Perrett et al , 1986; Perrett et al, 1987; Oram and Perrett, 1994). The implication of this is that while cells in lower levels detect edges, bars, blobs and orientation, at the highest level cells respond to complex abstract combinations of features such as faces. As mentioned previously, IT is thought to be the area of the visual cortex involved in object recognition. Therefore a tenuous conclusion can be drawn that perhaps IT contains cells that respond to individual faces and objects.

In a review article Perrett et al provide examples of evidence from both neurophysiology and neuropsychology to support hierarchical abstraction of complex features in visual processing, where object recognition is dependent on a step by step process culminating in the association cortex possibly in IT. They suggest that Hubel and Wiesel's work along with the clinical condition prosopagnosia, where people are able to identify objects but not faces, provides the foundations for such a theory. The existence of cells that respond to faces provides strong support for such a theory. They stress that the neuropsychological work is controversial. One problem that could be construed is that it is hard to see how IT could be selectively damaged so as to stop face responsive cells but not cells responsive for other objects. On closer examination of their claims it seems that the behavioural responses of such cells are not as clear as it first seems. Cells respond not just to one face but to faces. Cells recorded from monkey cortex responded to human and monkey faces. They did however not respond at all to other complex objects such as an alarm clock. Cells that did respond to faces however did so despite position, viewing distance, orientation and luminance of the face. Perrett et al also stress that this does not mean that visual processing does not occur in parallel, they merely suggest that hierarchical feature extraction is a possibility in the process of object recognition. Interestingly further experimentation showed that cells responded to particular facial features as well as they did to complete faces. However they go on to cite experimental evidence suggesting that certain cells responded best to "normal" arrangement of complete facial features. Finally response latency was affected by view, in that it took longer for



cells to respond to uncommon views. (for review see Perrett et al, 1987). It seems that the evidence is not conclusive. There could be cells responding to faces but at the same time there are cells that are responding to a lot of other things as well. They still maintain that the recognition of faces relies not on features encoded across populations of cells, but on small groups of selective cells and single cells. They also conclude that because cells responsive to faces are clumped together there are areas of the brain to deal specifically with face recognition. This does not mean it is exclusively dedicated to face recognition, but that is one of its particular functions. The implications of these findings are vast. If there are cells responsive to faces then is it possible that there are cells that respond to other objects. If there are then such a theory would fall foul of the problem of combinatorial explosion. It would also suggest that for recognition of new objects and new faces new cells would have to develop. This is not borne out by neurophysiological evidence. Arguably just because a cell responds to a face doesn't mean it is actually a specific feature detector. The numbers of cells that respond in this way are extremely small, only 10% of a sample of cells taken from IT (Perrett et al, 1982). Further criticisms include experimental evidence that shows the responses of cell in IT to be ambiguous, they respond to other things as well as faces. They also respond to faces with differing degrees of intensity (Bayliss et al, 1985). Another problem for purely feedforward hierarchical feature detection is that the response of cells in V4 in the colour form pathway are affected by their attentional state whereas cells in V1 are not (Moran & Desimone, 1985). If visual processing is feedforward hierarchical with the highest

level of processing holding a complete abstract representation like a face then you would not expect one level of processing to be affected by the animals state and not others. Also it has been shown that cells in anterior inferotemporal cortex (AIT) respond very differently in a delayed matching to sample task. Some respond to the sample, different cells respond in the delay period and a further class of cells responded to the colour the animal eventually chooses regardless of the colour it was shown. (Fuster & Jervey, 1982). Fuster and Jervey conclude from this that the extra striate pathways do not function independently of other areas of the brain. Other processes such as attention and memory influenced the strength of response in so called feature detecting cells. The fact that cells respond to the chosen colour, which is different from the colour of the object presented to the retina, suggests that visual information is not completely processed in a feedforward hierarchical way. Information feeds back and forth not only in the visual pathways but to and from other sensory processing areas.

Oram and Perrett in a recent paper outlining how models can be derived from neurobiological constraints revise feature detection theory to fit with a more general view of visual processing (Oram & Perrett, 1994). They discuss in further detail the nature of cells in IT and their role in object recognition. They maintain grandmother cells do exist but now describe them as cells that are responsive to "biologically important visual patterns..." such patterns include faces, limbs, eyes etc. and include groups of features or just single features. They acknowledge work which has shown cells in IT are selective for complex patterns but suggest that biologically important visual patterns are more abstract than these complex patterns.

They cite neurophysiological evidence that suggests that cells in IT prefer specific combinations of features to be present and in a specific order for such cells to respond maximally (Oram & Perrett, 1994). This fits with psychological research that suggests neonates are capable of face recognition at extremely young ages (Fantz, 1961; Bushnell et al, 1989). In one reported case a two day old neonate is capable of distinguishing between its mother and a stranger (Bushnell et al, 1989). Oram and Perrett do emphasise that although visual perception involves the interaction of many different areas of the brain the model they are concerned with is one that involves constructing a visual image from a bottom-up approach. They are trying to model visual perception arising from being asked "what is this picture of?" Exactly how much of visual perception is based on such an approach is questionable, and even if you ask such a question experience surely plays a large part in analysis of individual components of the picture (Oram & Perrett, 1994).

Evidence from neurophysiology and neuropsychology shows strong support for parallel models of visual processing. It is possible to support a SCFD approach within a parallel model. In view of the fact that the mechanisms for integration of information in parallel models is not at all clear SCFD attracts much attention, especially in relation to the development of artificial vision systems. There is still much debate as to whether the SCFD model is correct. Evidence from both SCFD and parallel models suggest that visual processing is hierarchical. However parallel models especially that proposed by Zeki emphasise the significance of feedback from higher to lower levels for visual perception. Whether visual perception can be

explained in terms of bottom-up processing leading to the abstraction of features of increasing complexity is not clear. The strongest support comes from studies of object recognition, or more specifically face recognition. Arguably this is a very specialised area of visual processing, involving a great many other factors. It may not be sufficient to extrapolate the mechanisms of face recognition to general visual processing. One advantage of adopting SCFD as a model of visual processing is that it is easy to see how it may operate at the functional level. Inputs from neurons ultimately converge in a feedforward hierarchical fashion on a single cell in IT. Parallelism in the visual system can be seen even in retinal processing of the image. It seems then that it is an important feature of the visual system right from early vision to late vision. Single cell feature detection might be part of visual processing but it does not explain visual perception. Because a single cell fires in the presence of certain visual stimulus does not mean all our understanding of that object is stored in that one cell. Even if you accept the feature detection theory you must still have an explanation of the mechanisms underlying early stages of processing. Most of the evidence to support single neurons responding to faces concentrates on static images. Bearing in mind that motion information is processed in the parietal areas how does a SCFD approach cope with the recognition of moving faces? Visual perception means combining processing from other visual, and sensory areas as well as incorporating information from past experiences involving the mechanisms of memory and learning. It may be that there are cells that respond to biologically important objects, but this is just a small part of visual processing and it does not explain vision as a whole.

## OSCILLATIONS AS AN UNDERLYING MECHANISM FOR VISUAL PROCESSING.

It seems that processing in the visual cortex can be explained as a combination of hierarchical and parallel processing. This means that within the cortex there is a need for a functional mechanism for visual processing capable of explaining the linking of resultant processing in separate areas of the visual cortex. Since the late 1980's and more significantly the early 1990's a new body of research within the field of neurophysiology has evolved that has changed not only our understanding of how the brain functions but has also had a major implications in the field of neural modelling (Gray & Singer, 1987; 1989; Eckhorn et al, 1988; Gray et al, 1989; Eckhorn et al, 1989; Eckhorn et al, 1990; Engel et al, 1990; Eckhorn & Schanze, 1991; Eckhorn 1991; Engel et al, 1991a; 1991b; Engel et al, 1992.). This work has provided new explanations as to how the brain processes information specifically in the visual cortex. In general terms the findings of this research suggest that information about stimuli can be encoded in the brain as patterns of oscillation. Most of the work has focused on the presence of oscillatory behaviour in the neurons of the cat visual cortex, however this oscillatory behaviour has been found to exist in other areas of the brain as well as the visual cortex, including the olfactory bulb and the thalamus (Skarda & Freeman, 1987; Crick, 1984). Recently oscillatory activity has also been shown to exist in humans (Ribary et al, 1991).

The discovery of oscillatory activity in the cortex occurred simultaneously

by several groups of neurophysiologists.(Gray & Singer, 1987; Eckhorn et al, 1988.) Initially the work was purely experimental, but later research lead to the development of theoretical models based on the results of these experiments ( Eckhorn et al, 1989; Eckhorn et al, 1990; Engel et al, 1990; Eckhorn & Schanze, 1991; Eckhorn 1991.) This theoretical work enable neurophysiologists to experiment with and analyse the complex emergent behaviours of these oscillatory neurons. This was done by replicating the experimental findings through computer simulation and mathematical analysis. The work generated has provided an important link between experimental and theoretical analysis of theories of both brain and mind. It has also provided insight into the underlying functionality of neurons in the visual cortex. This is extremely important, because although we have a reasonable understanding of the structure of the visual cortex we know little about its underlying mechanisms.

The synchronisation of oscillations in separate cells acting as feature detectors was first put forward by von der Malsburg, who proposed cells that discharged with the same frequency and in phase would form strong excitatory links with other cells behaving in the same way. Eventually individual cells would become a group of cells all representing the same stimulus (von der Malsburg, 1981). Since the discovery of oscillatory neurons in the cat primary visual cortex a great deal of research has been conducted into the significance of these oscillatory behaviours. Eckhorn and his fellow researchers also propose phase locking in groups of oscillating neurons as a mechanism for feature linking in the visual cortex. They propose that this temporal coding through phase locking in oscillating

groups of neurons that are spatially separate, but functionally similar, is a mechanism for secondary visual coding. Primary visual coding is represented by the receptive field properties of individual neurons. This secondary processing is still pre-attentive, enabling the appropriate features in the visual scene to be linked together before high level processing takes place. (Eckhorn et al, 1988; 1989; 1990; Eckhorn & Schanze, 1991; Eckhorn, 1991).

In relation to the problems of integration of information in parallel models of vision this provides a possible solution. Zeki proposes in his model outlined previously, that the integration of information from spatially separate sites could be explained by diffuse feedback connections. However he does not discuss in any detail the mechanisms underlying such connections, further this does not account for integration between sensory areas such as the visual cortex and the auditory cortex. In view of this the idea of integration through oscillation sounds very appealing. At least a combination of both feedback connections and oscillations could be considered. In a review of the work so far into temporal codes Engel et al conclude they provide a solution to the integration in distributed neuronal networks. (Engel et al, 1990; 1991a; 1991 b; 1992).

## PSYCHOLOGICAL THEORIES OF VISION.

When discussing neurophysiological theories there is a tendency to talk about visual processing rather than perception. This is generally because it is assumed that processing is an integral part of perception, and it is the underlying mechanisms of perception that are being described at the neurophysiological level. Psychological approaches can be thought of as being on a level above, dealing with perception from a behavioural point of view rather than its underlying mechanisms, trying to establish the nature of higher level cognitive function rather than the so called lower level visual processes associated with early vision. Neuropsychology is perhaps where vision is both thought of as perception and as processing. Psychological theories of vision are not always concerned with attempting to describe visual perception in its entirety. Vision can be thought of as modular, and theories of particular problems of vision are common in psychology. Marr refers to the breaking down of vision into a series of smaller theories as "independent modules of perception" (Marr, 1982). Marr goes on to argue that there are in fact very few general theories that attempt to describe visual perception, because it is so complex it tends to be broken down into its component parts and each being investigated alone:

*"...students of the psychology of perception have made no serious attempts at an overall understanding of what perception is, concentrating instead on the analysis of properties and performance."*

(Marr, 1982).

The problem with this approach is that it becomes difficult if not impossible



to see how all the modules fit together to constitute perception. Further, it makes the unification of top-down and bottom-up descriptions even harder. Reducing many explanations of the nature of visual cognition to underlying neurophysiology becomes too complex to even consider in terms of visual perception. Arguably it becomes easier within the separate modules, but does this equate to perception?

As with neurophysiological theories of vision it is possible to distinguish between two main categories of theories of visual perception, "concept driven" or top-down theories and "data driven" or bottom-up theories (Lindsay & Norman, 1972). These two approaches are related to the terms used to describe the general relationship between neurophysiological and psychological approaches to understanding visual perception as discussed in chapter 1. Data driven theories hold the premise that perception can be explained in terms of interpretation of the visual scene from the data as it streams through the system. No prior knowledge about the world is needed to interpret the world. All that is needed for perception is the information present in the visual world. Concept driven theories on the other hand suggest that the visual scene is interpreted by making comparisons to stored knowledge about the world acquired through experience, or genetically determined. Perception is not about constructing an image from data it is about using data to test hypotheses about the world.

## CONCEPT DRIVEN THEORIES OF PERCEPTION.

Early explanations of vision focused on concept, or hypothesis driven models of vision. The retinal image provided a hypothesis that could then

be tested against what was known about the world. Such approaches to explaining vision were strongly linked to the constructivists view of cognitive psychology, that the retinal image was insufficient on its own. It was through building upon the retinal image through successive stages of processing including referral to information stored about the world that resulted in perception (Best, 1986). This approach to the explanation of perception was popularised in the 1970's as "information processing theory" (Lindsay & Norman, 1972). In turn artificial intelligence models were developed based on this approach (Tenenbaum & Barrow, 1976). Although there was a great deal of neurophysiological research into how the brain processed visual information the findings of this work were seen as the data element in a process that was far more concerned with cognitive states. The emphasis for any model was that it had to be couched in a discipline. That discipline was predominantly psychology, however there were theories within psychology that were firmly rooted in neurophysiology. Such theories though were transformed to fit psychological terms and research methods. Neurophysiology provided descriptions of the mechanisms and psychology provided the models.

The gestalt psychologists proposed that perception involved the grouping of the elements of objects in the visual scene into whole objects or "gestalts". The gestalt psychologists were particularly interested in how we perceive ambiguous images. They proposed that "the whole is greater than the sum of its parts". Ambiguous images provide exactly the same retinal information yet they have more than one possible interpretation. The gestalt psychologists suggested that this is because vision is top-down, you have

two possible interpretations of the data and what you perceive depends on which interpretation you apply. High level cognitive processes are in control of the underlying neuronal mechanisms of lower level vision (Bruce & Green, 1990). They went on to describe what are commonly referred to as the "gestalt laws of organisation" which comprises of a series of principles that underlie the grouping of various component parts of an image to form a gestalt. These include, proximity, similarity, structure from motion, continuity, closure and goodness of fit. They suggested that the visual scene was represented isomorphically in the brain according to these principles. The work of the gestalt psychologists had a significant impact on early approaches to the development of artificial vision systems (Waltz, 1975) and to Marr's computational model of vision (Marr, 1982). The grouping principles aid scene analysis by allowing various segmented parts of the image to be grouped together to enable object recognition. However one of the biggest criticisms of the gestalt approach was that there is no evidence to suggest the mechanisms that underlie the laws of organisation exist (Bruce & Green, 1990). Gregory points out that if the visual scene is represented isomorphically in the brain, you would need a further eye in the brain to see it (Gregory, 1973). It seems that the emphasis should not be on the details of the theory but rather on its more global implications. The suggestion that perception is a top-down process is still very much debated even today. The effect of experience on perception provides a great deal of support for this to be the case. This has been highlighted through the results of psychophysical experiments with visual illusions. Cross cultural research into the effect of the Muller-Lyer illusion suggests that previous visual

experience from our environment determines whether or not we see the illusion (Deregowski, 1972).

Gregory also proposes that visual perception is predominantly top-down. He suggests that visual perception involves testing hypotheses about the world in order to make sense of it:

*"Perception is not determined simply by the stimulus patterns; rather it is a dynamic searching for the best interpretation of the available data...the senses do not give us a picture of the world directly; rather they provide the evidence for checking hypotheses about what lies before us."*

(Gregory, 1972).

Gregory also looks to visual illusions for support for his theory suggesting high level cognitive processes shape the interpretation we make about the visual image. As well as the Muller-Lyer evidence he sites the example of depth perception in visual illusions. Using a hollow mask of a face he demonstrated that it is only when the viewer is close to the face that depth information from the image itself is used to interpret it. Meaning that when the hollow face is held at distances of more than a few feet it is perceived as being convex as opposed to concave or hollow. We are perceiving what we expect to see rather than what we are really seeing.

## MARR'S THEORY OF VISION.

More recently interest has moved away from the idea of vision being a top-down process and the emphasis has been on data driven approaches, an example of which is David Marr's computational approach (Marr, 1982). The approach proposed by Marr is particularly relevant, as it provides an

example of an attempt at both outlining a theoretical framework and then developing a model within that framework. Marr's theory of vision has had a significant impact on attempts at understanding visual perception, in the field of cognitive science and on the development of artificial vision systems in the fields of engineering and computer science. It also poses an important question about the nature of multi-disciplinary approaches to science that have direct bearing on the approach taken in this thesis. It could be argued that these approaches are not truly multi-disciplinary, and that at the level of description of the model they are firmly linked to a particular sort of representation. For example with Marr, his approach is very much a computational approach and the level of description for his model, although based on general ideas from both neurophysiology and psychology, is dominated by the fact that it translates everything into a computational form. This means that Marr's model is subject to the problems associated with the translation from one description to another "universal" description. The other problem that is apparent in this approach is that research is then directed towards trying to find evidence to support the existence of the model of vision rather than looking at what vision really involves. Psychologists for example, interested in Marr's theory of vision dedicate their research to look for behavioural evidence to suggest that Marr's levels exist rather than looking at what vision actually is. Although Marr's intention was that vision could be described at many levels, the major impact of his work has been the implementation of computer vision systems that are in fact far removed from the visual processing mechanisms that takes place in the brain.

Marr's approach to the study of vision took the opposite view to the traditional information processing theories that had preceded it. Marr believed that vision could be explained in terms of bottom-up processing. His model was "data-driven" rather than top-down "hypothesis-driven". The theory was based on the assumption that you did not need prior information about the world to perform the task of visual perception. All the information needed for interpreting the world was present in the incoming data (Marr, 1982). A symbolic representation of the light falling on the retina was derived from the positions, orientations and movements of surfaces and their relations to one another. Marr proposed that visual processing took place on three levels. Firstly the retinal image has to be translated into data that can be processed. Translation involves matching each cell on the retina with the corresponding grey level image associated with it, giving rise to a "grey level image" (GLI), from which a grey level description is computed. Computation of the description is through the detection of certain features in the GLI intensities that help to explain it. These intensities are represented symbolically as signatures for features corresponding to edges, bars, blobs and terminations in the image. The result of this level of processing Marr referred to as the "primal sketch". The second level of processing involves describing the "layout" in the world of the structures in the primal sketch in relation to the viewer. This can be explained as taking the symbolic descriptions of features and their geometry and putting them in a frame of reference, or co-ordinate system, making explicit orientations and depth. This level involves computing the underlying structural description from the features in the primal sketch and

results in the 2 1/2 D sketch. The third level of visual processing attempts to solve the problem of higher level visual processing. It is not a well-defined problem and ultimately not complete in Marr's theory. The emphasis is on the construction of a three dimensional object based on an object centred description as opposed to a retinotopic description that has been the case up until this stage. This is because if the description was viewer centred every time the view changed the object would appear to change and so its description would have to be altered. As we know, objects do not change every time our view of them changes, their form remains constant. Therefore a computation of the "volumetric primitives" of the two dimensional image takes place. Objects are described by taking a principle axis within them and building up a hierarchy of generalised "nested cones". This gives a two dimensional "stick" object three dimensional structure. Marr refers to this level of processing as the 3D sketch. (Marr, 1982).

The early stages of visual processing proposed in Marr's theory are based on strong neurophysiological understanding of the vision system, for example the work of Hubel and Wiesel. They discovered cells in the primary visual cortex were organised in a very specific way, that led them to believe they acted as feature detectors looking for things like orientation, movement and edges in the information transmitted from the retinal image (Hubel & Wiesel, 1962). The later stages of the model are not so well supported by neurophysiological evidence. This might explain why the 3D sketch is not so clearly developed in the model. Whereas the primal sketch has strong foundations in neurophysiology, the 3 D sketch is about explanations of higher level cognitive function, something that has little support in

neurophysiological experiments. Also it is relatively easy to translate the primal sketch into a computational description that is easily implemented (Marr & Hildreth, 1980). The 2 1/2 D sketch also translates relatively well in a modular fashion but has not been implemented as a complete model. However the 3 D sketch does not translate in the same way, since the information it conveys is not easily described computationally. The implementation of the primal sketch has not been without problems. Neurophysiological evidence since the Marr-Hildreth algorithm was devised has contradicted its formulation (Schiller, 1982). So in translation to the computational level information was misinterpreted, providing evidence to support the notion that a single language is not necessarily the solution to multi-disciplinary study. This also supports the idea that implementation is perhaps best thought of as an aid to the development of theories rather than the end product of a theory.

#### VON DER MALSBURGS CONSIDERATIONS.

Contrary to Marr's theory of vision is one relating more to the early approaches to vision and Gregory's views. In brief von der Malsburg has constructed an outline for the development of a model of vision referred to as "considerations for a visual architecture" (von der Malsburg, 1990). Importantly von der Malsburg insists that the theoretical work in modelling vision is central to understanding vision. The depth of understanding and available technology should enable a model of vision to be constructed through the assimilation of information from many disciplines to form a coherent whole:



*"What is needed now is the integration of all the components into one coherent functional whole: the definition of a visual architecture."*

(von der Malsburg, 1990).

Central to the view of von der Malsburg is that vision is in fact an active process of testing and constructing visual information from stored knowledge we have about the world. This is a very different approach from that of Marr, who saw vision as being independent of our understanding of the world. Unlike Marr's theory of vision von der Malsburg suggests that top down information in the form of schema's are crucial for perception. Therefore von der Malsburg's account of perception goes beyond the visual processing areas of the brain to include discussion of the role of memory in visual perception. The account of the nature of perception given by von der Malsburg is theoretical, so unlike Marr's account it is not addressed at a computational level necessary for possible implementation.

There are two key components to von der Malsburg's theory: architecture and data structure. According to von der Malsburg architectures are structures which shape and direct understanding of a system:

*"Architectures are standardised, flexible formulations of data structures and processes."*

(von der Malsburg, 1990).

Data structures are the elementary parts that make up the architectures, for example neurons are the data elements of the brain. It is the organisation of

those elements within the architectures that lead to the understanding of vision. The theory is very much a series of "considerations" and it is not always possible to agree with some of the assumptions made by von der Malsburg. However the approach outlined is sympathetic to the aims of this thesis. Rather than translate information into one language, as with Marr's approach, transfer of information between disciplines is much preferable. By constructing a well-defined theory it is possible to combine information from various disciplines, thus also combining both a top-down and bottom-up approach.

## CONCLUSIONS.

It would seem then that there is no clear answer as to which theory, neurophysiological or psychological is correct. Development in either field seems to be related to trend as much as anything else. Within neurophysiology new evidence tends to lead to the adaptation of existing theories to fit new information. With psychology there seems to be separate evidence to support either one view or another. It seems clear that in terms of understanding real perception no one theory is right. It would seem fair to suggest that vision at the neurophysiological level is a combination of distributed hierarchical processing with the possibility of specialised cells for biologically important information (although this is far from conclusive). In terms of providing a psychological description of perception again it seems that it must be a combination of the two approaches, data driven and concept driven.

In the next chapter the evidence from chapters 2 and 3 will be used to

**develop both a theory and a conceptual model of visual perception.**

# **CHAPTER 4**

## **A MODEL OF VISUAL PERCEPTION:**

### **AN OUTLINE OF A CONCEPTUAL MODEL OF PERCEPTION FROM A TOP-DOWN AND BOTTOM-UP APPROACH.**

*Chapter 4 can be divided into two parts all of which are heavily based on the general descriptions of the neurophysiology and psychology of vision provided in chapter 2 and 3. Part one briefly outlines a very general theory of perception. Part two describes a more detailed conceptual model outlining possible architectures and functions underlying the theory and includes a section on the architecture and function of the model and a section on their underlying mechanisms.*

### **A THEORY OF VISUAL PERCEPTION.**

The aim of this chapter is to create a hypothesis about how perception might be explained both theoretically and as a conceptual model. This means that some aspects of the model are purely speculative, whereas others are more factual and therefore described in more detail. In turn this means that some aspects of the model are more relevant to the development of artificial vision systems and others to the advancement of our understanding of the neurophysiology and psychology of perception. It is envisaged that the conceptual model will provide an “architecture” with which to direct further

work. Therefore the less detailed, more speculative aspects of the model can be addressed in future research.

## INTRODUCTION.

In chapter 3 psychological or top-down theories of perception were classed as being one of two possible types. The first class of theory suggested that vision was essentially concept driven (top-down), where information entering the brain at the retina consisted of hypotheses that were interpreted through testing against experiential or genetically predetermined knowledge. Theories based on this approach include Gregory's theory which describes perception as hypothesis testing (Gregory, 1973). The second class suggested that vision is a data driven process (bottom-up) where stored knowledge is not needed to perceive, all the knowledge that is needed is already present in the world itself. Such theories of vision include Marr's and suggest that no previous experience of the world is needed in order for perception to take place. Marr's theory however does imply that although early visual processing may be structured like this higher level interpretation, for example object recognition relies on stored knowledge in the form of a collection of three dimensional model descriptions (Marr, 1982). Both classes of theory are limited, data driven theories such as Marr's best explain low level vision whereas concept driven theories such as Gregory's provide better explanations of high level visual processes such as object recognition.

The theory of vision I wish to propose here is not new. It merely combines or extends the best aspects of existing theories. Unlike existing combination

theories (Nakayama, 1990) this theory will attempt to explain perception in its fullest sense rather than just provide a functional explanation of the organisation of the visual system. At the same time it aims to provide insight into how architecture, function and possible underlying mechanisms support the theory as well as provide ideas for artificial vision systems.

#### A COMPLETE THEORY OF VISUAL PERCEPTION.

The theory underlying the model suggests that visual perception results from a combination of both data driven (bottom-up) and concept driven (top-down) processes. Representations are constructed within the visual system which form hypotheses about the world. Such representations are not necessarily generated entirely from visual input. Other sensory input is combined with visual input in the form of a poly sensory hypothesis at the highest level of processing. These hypotheses are then tested by comparing them to knowledge acquired through previous interactions with the world, and knowledge that is genetically predetermined. Stored associations made between previously encountered visual information and other sensory information also contributes to the interpretation of hypotheses. The result of this "testing" allows for some form of identification or labelling of the visual image based on this top-down knowledge. Conformation results from the labelled hypothesis being propagated back through the system to be "checked" against bottom-up data.

Data driven processing amounts to the construction of representations of the visual scene using knowledge that is implicit in the world and exists independently of high level knowledge gained through intelligent reasoning

processes. It is essentially an unconscious process. Concept driven processing involves the use of high level cognitive processes both conscious and unconscious that allow us to reason about data driven representations. So visual perception involves building bottom-up or data driven representations that can be compared in some way to existing knowledge about the world resulting in the interpretation of the world which can then be acted upon.

#### SOME PREDICTIONS OF THE THEORY.

- It is possible to describe vision in terms of the three levels outlined in chapter 1. Visual processing equates to the construction of a hypothesis using data driven information. Local perception equates to the identification of a hypothesis in terms of its visual content alone. Finally Global perception equates to the use of top-down information in the form of memory and other sensory input resulting in perception.
- Visual perception is facilitated by experience. Knowledge about the world enables us to perceive it more quickly and accurately than if we have no knowledge.
- A visual representation is not necessarily the same as a visual perception. Information processing using top-down information means that the resultant perception can comprise more information than is made available through data entering the visual system.
- If a hypothesis cannot be constructed through damage to the visual system then visual perception would be greatly impaired. However, it is possible that if the visual image has been experienced before, top-down

information could evoke a visual response. As it would be impossible to check this response against incoming data it would be expected that it may be extremely difficult to consciously report on the visual information being experienced.

## **A CONCEPTUAL MODEL OF VISUAL PERCEPTION.**

The development of a conceptual model is strongly rooted in neurophysiology. The aim is to build a model based on the architecture and function of the visual system that will uphold the top-down theory of perception outlined. The development of the model is strongly influenced by chapters 2 and 3, however there is significant difference between what was described there and what will be described here. Chapters 2 and 3 described in general terms the visual system in a biological brain. What is described here is a model based on the biological vision system, not a replica of that biological visual system. The biological visual system is extremely complex, and it is almost an impossible task to build a conceptual model on such a scale. Even if it is possible it still may not function in the same way that the biological vision system does when it stands alone from the rest of the brain. To develop a conceptual model of perception capable of supporting a theory of perception requires gross simplification. This is even more important if at any time the model is to be implemented in software. Therefore the conceptual model outlined here is by no means definitive, it is highly speculative and at best will be biologically plausible rather than biologically real.



The approach taken to developing such a model involves starting with the simplest possible description and continually refining it until a detailed model emerges. This is done firstly for architecture, outlining the structure of the model, and then gradually functional descriptions of structure can be added resulting in a complete conceptual model of visual perception. The first stage of the development of the model is to begin building a structure that gives a global interpretation of the system. This process equates to the general philosophy of the thesis that you start with a "grand theory" or "general model". Once this has been achieved satisfactorily the global model can be broken down into its component parts and structure can be outlined at the local level for the various modules. This section of the chapter aims to illustrate how the conceptual model is developed.

#### ARCHITECTURE & FUNCTION.

The conceptual model is based on the neurophysiology described in chapter 2 and 3. In its simplest form the model can be outlined in terms of a functional architecture. Basically it will comprise a series of units corresponding to the major areas that make up the visual system in the brain. Each of the units outlined can be looked at in two ways, firstly through connections between the units as a whole system, and secondly as a series of independent sub systems (see figure 13).

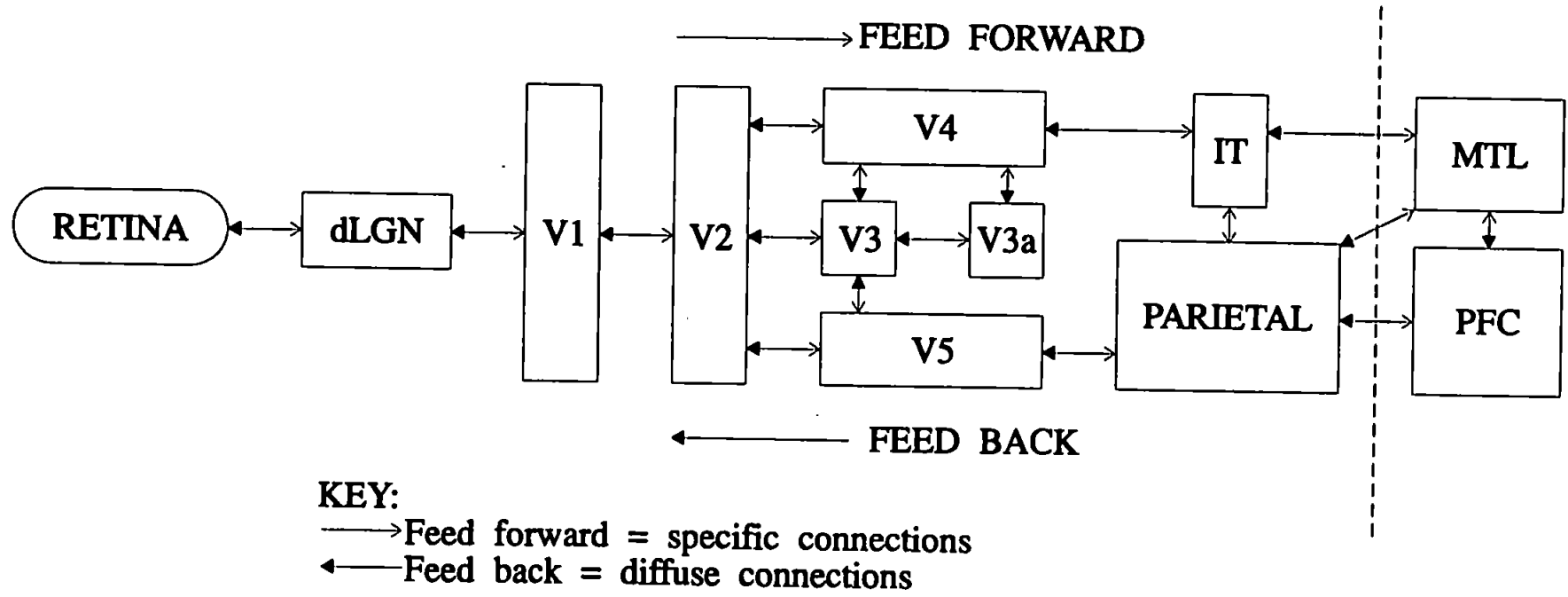
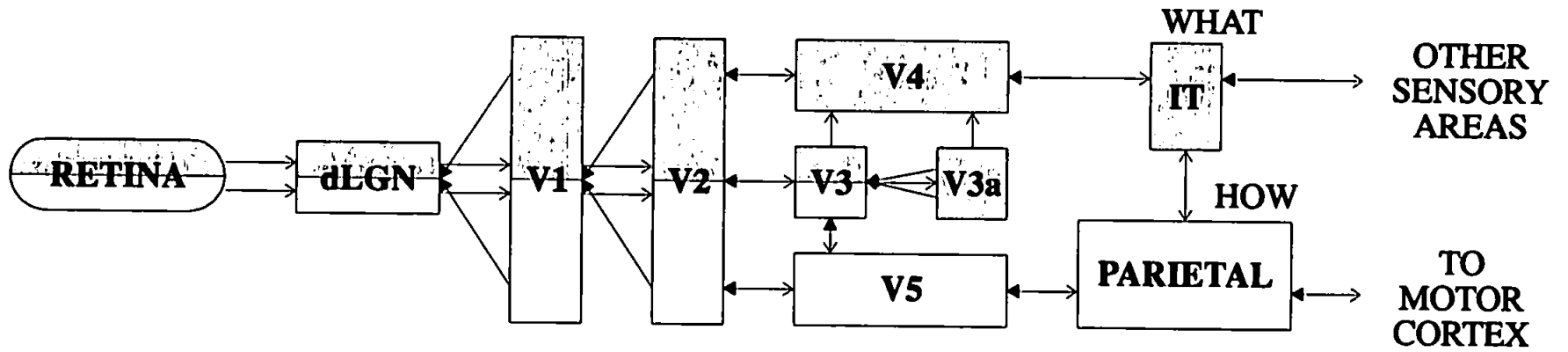


Figure 13. A simple outline of the structure of the model.

The model comprises of a series of hierarchically connected visual areas with both feedforward and feedback connections throughout the whole system. As with the biological visual system the feedforward connections are specific, (specific information is transmitted along certain pathways) and the feedback connections diffuse (information is projected back across the whole system regardless of its nature). The retina and dLGN to V1 represent the equivalent of the primary visual pathway. V1 represents the striate cortex and V2, V3, V3a, V4 and V5 the extra striate areas. IT and the PARIETAL regions represent the association areas and MTL (medial temporal lobe) and the PFC (pre frontal cortex) represent “memory” in its broadest sense. It is suggested that MTL is needed for the consolidation of memories, especially declarative memory, and the PFC is strongly implicated in working memory. In terms of the model detailed discussion of these areas is not necessary and it is sufficient to refer to them as “memory” (See O’ Shaughnessy, 1995 for a detailed review of the structure and function of these two areas).

In general terms the overall functionality of the model can be classified into two streams of information flow operating in parallel as suggested in the biological visual system. One stream is the “WHAT” pathway and the other the “HOW” pathway (see figure 14).



**KEY:**

- The What pathway
- The How pathway
- Specific feed forward
- Diffuse feed back

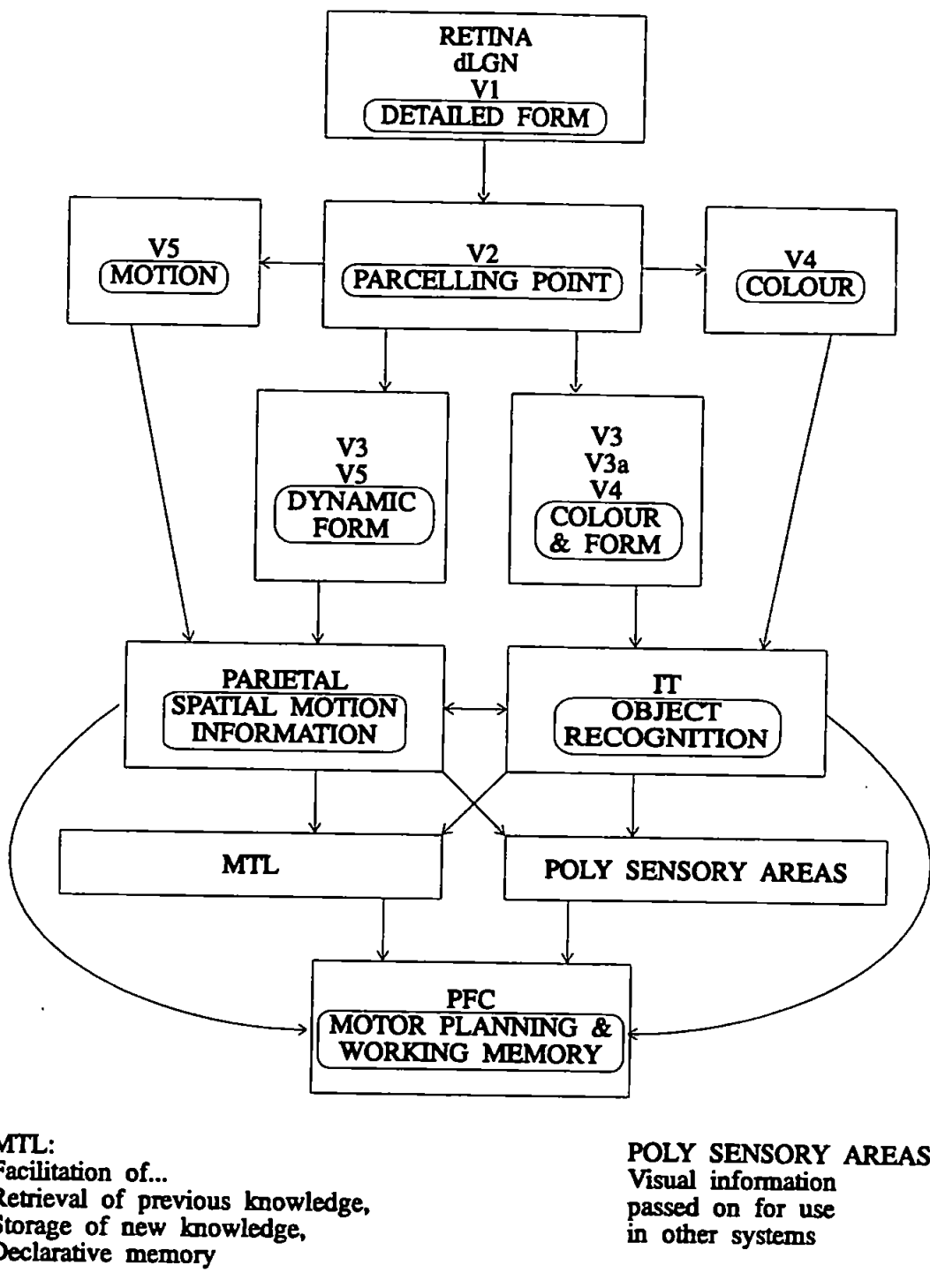
Figure 14. The HOW and WHAT pathways.

The WHAT pathway is concerned with the identification of visual stimuli whereas the HOW pathway is concerned with how to use or respond to the visual stimulus. The separation of information into two pathways begins at the retina and continues through V1 and V2. In the biological vision system the division of two major pathways in the dLGN, V1 and V2 is extremely complicated as both pathways pass through the same structures. In the development of the conceptual model this has been simplified.

Simplification is necessary at all levels to enable the development of a model capable of explaining biological perception in its global sense and to be able to use the model in the development of artificial vision systems.

Beyond V2 the segregation of information flow becomes less complicated, the WHAT pathway being a general description of the colour and colour form pathway terminating in IT where objects are identified, and the HOW pathway the same for the motion and dynamic form pathway terminating in the PARIETAL regions where spatial information is processed. The WHAT and HOW distinction is thought to be maintained in the PFC where there is evidence to suggest that “what” and “where” information is processed separately (Fraser et al, 1993). This suggests that working memory, thought to be located in the PFC, like the visual system is modular (see Service, 1993 for a brief review).

Functionally the model will perform as dictated by the biological visual system (see figure 15).



**Figure 15.** A simple outline of feedforward functional processing.

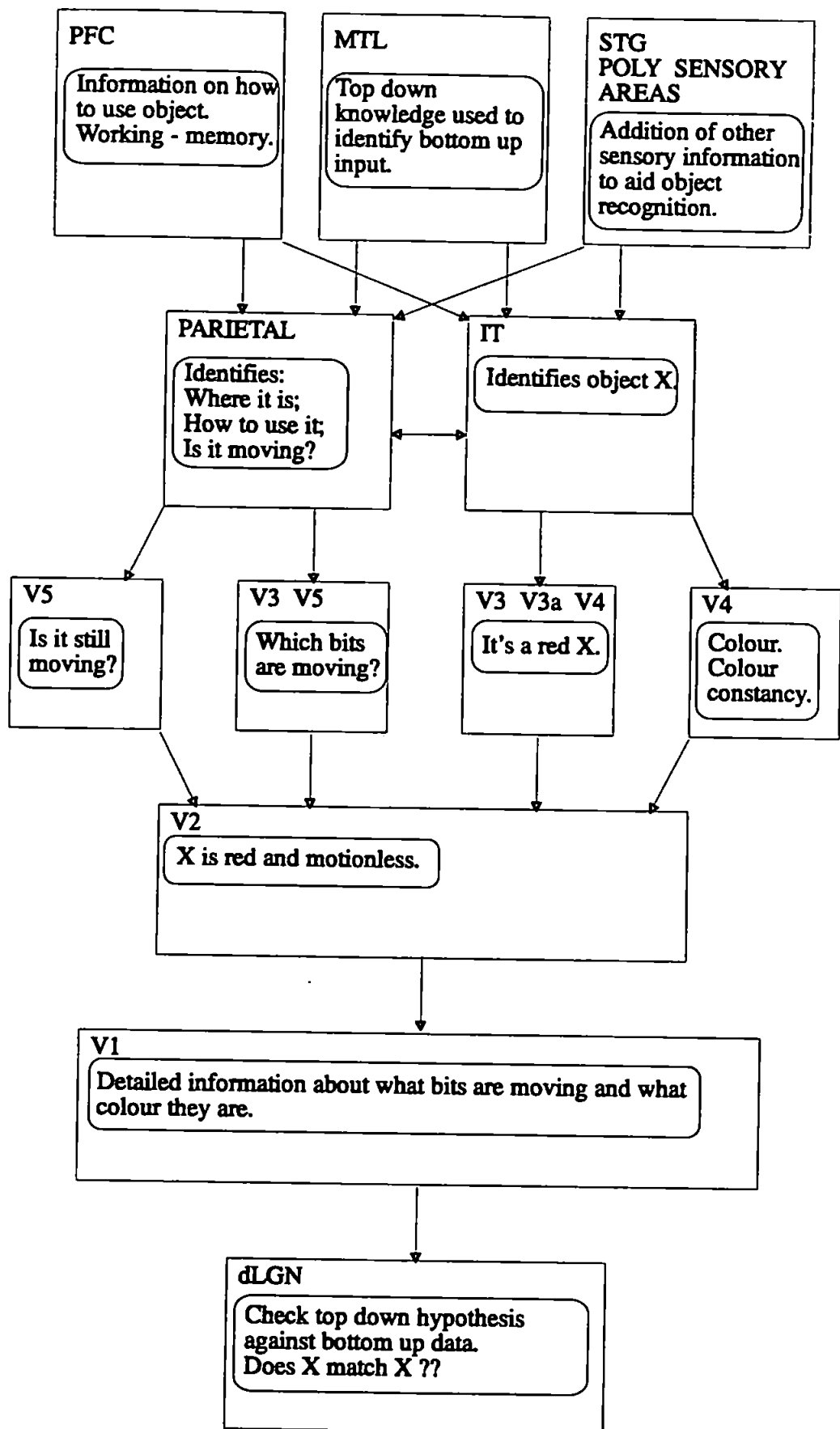
Data enters the system at the retina where it will pass on to the dLGN. From the dLGN data is relayed to V1 which constructs an early representation of the image in terms of form. V2 then acts as a parcelling point where specific information is sent for further processing, directly to V4 for colour,

to V4 via V3 and V3a for colour and form information combined, to V5 for motion detection and to V5 via V3 for dynamic form information. At the highest levels in the visual system IT is the area where object recognition takes place, and the PARIETAL region where spatial layout information is interpreted. From here information is processed for two main purposes firstly for object recognition (what), and secondly for action (how). Therefore IT connects both to the MTL and directly to the PARIETAL region. The MTL is necessary for the retrieval of past experience or "memories" that facilitate recognition in IT, and for the storage of new experiences or objects for future use. IT also relies on HOW information from the PARIETAL region to identify objects. Where objects are in relation to each other can be used to aid interpretation as can information on what to do with objects. It is possible to identify things by their purpose as well as by form information especially when form information is incomplete in some way. The PARIETAL region is predominantly concerned with spatial and motion information and so has direct connections to the motor cortex and PFC, the motor cortex for action generation and the PFC for planning and working memory. The PARIETAL regions also need direct connections to IT so that a label for the visio spatial information is available immediately. As with IT, connections to the MTL also allow for previous experiences to be incorporated in HOW decisions and new information can be stored for future use.

As with biological vision there are extensive feedback connections throughout the system. These connections are non specific, meaning that all of V2 and V1 and the dLGN receive feedback from both high level streams.

One suggestion for a possible role for diffuse feedback connections in the biological visual system is to provide a possible mechanism for complete integration of information processed in the four separate streams at all levels (Zeki, 1992). Functionally separate information is distributed to areas that did not receive it through feedforward pathways, via diffuse connections. Therefore abstract global information about object identity and visuo spatial relationships is integrated with detailed form information via these diffuse feedback connections (see figure 16).



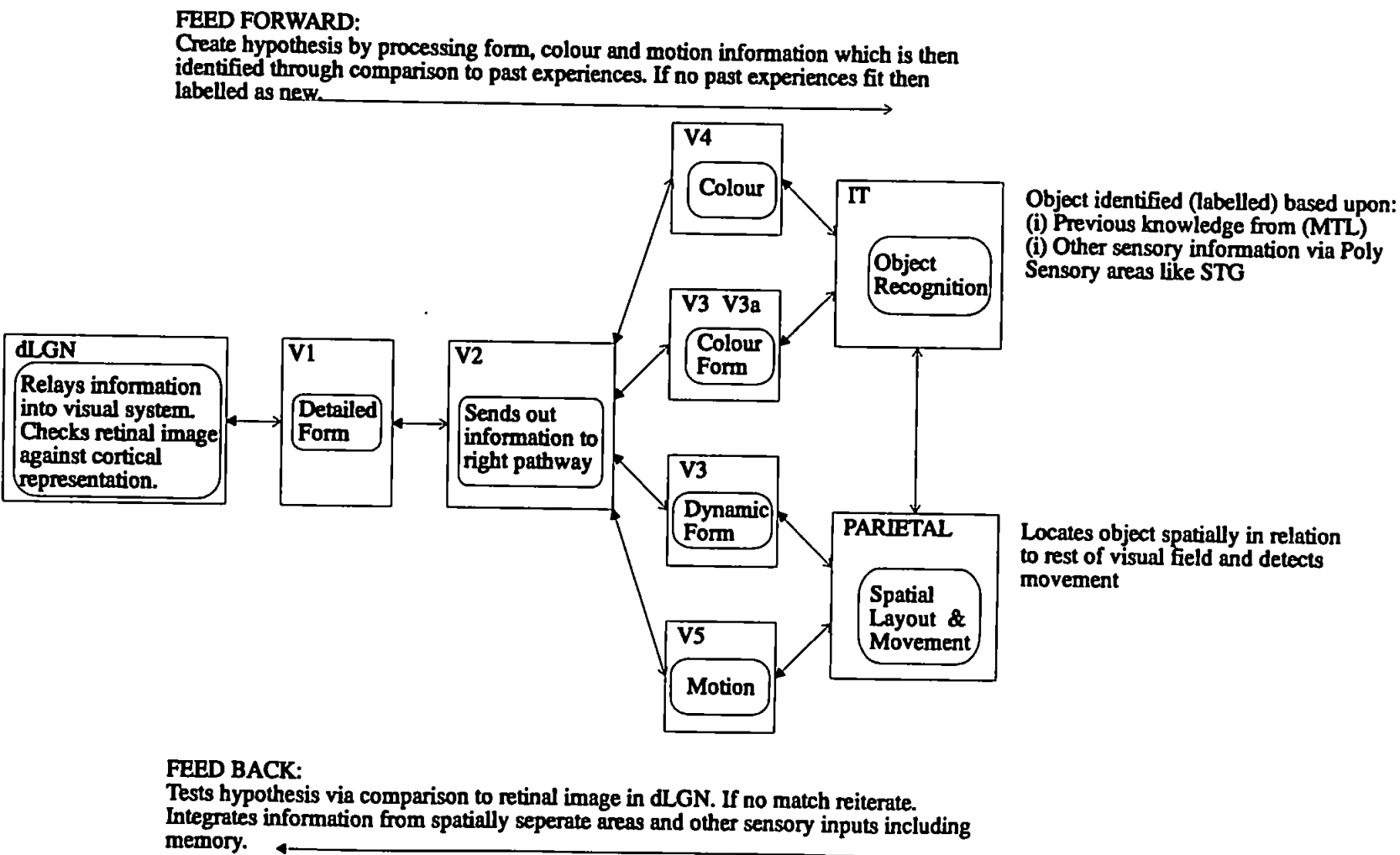


**Figure 16.** A simple outline of feedback functional processing.

Feedback connections in the model perform three crucial roles:

- 1 To enable the integration of information for perception including:
  - Information processed in functionally separate structures such as colour and motion information.
  - Top-down information about the world stored in memory accessed via the MTL and PFC.
  - Information from other sensory processing available via poly sensory areas in the association cortex notably the superior temporal gyrus (STG) and the posterior and anterior superior poly sensory areas (STPp and STPa respectively).
- 2 To maintain the visual representation in all areas for detailed analysis:
  - Allowing for detailed low level form information to be associated with high level abstract information about object identity.
  - To enable an incorrectly confirmed hypothesis to be re evaluated.
- 3 To provide a mechanisms for checking hypotheses constructed about the visual scene:
  - Once a hypothesis has been constructed and compared to stored knowledge feedback enables the possible interpretation to be checked against the incoming data.

At this point the model based on both bottom-up neurophysiology and top-down psychology can be summarised by figure 17.



**Figure 17.** A summary of the model so far.

The aim of the model is to construct and test hypotheses about the world

based on incoming data from that world. Top-down information in the form of both stored and genetically determined knowledge is used to test hypotheses, as is bottom-up data entering the system in the dLGN. The model functions as follows:

The visual image is basically processed in two ways, one to identify WHAT things are in the image and two HOW to use and react to the visual image. Although integration of WHAT and HOW information is important it is not always necessary. You can identify things without actually having to know their exact spatial location in relation to other objects, and how to manipulate and use them. By the same token you do not necessarily need to know what something is in order to react to it or know how to manipulate it. The first step towards perception involves the construction of a hypothesis. Initially the hypothesis consists of detailed form information (V1), then more abstract properties such as colour and motion information (V3, V3a, V4 and V5 via V2) are incorporated. The result causes particular cells in IT to fire. These cells or groups of cells represent particular features of objects and possibly in some cases whole objects themselves. The representation formed in IT is very much abstracted with no detailed form colour or motion information. Detailed information is associated with the representation in IT via diffuse feedback connections. At the same time in the PARIETAL regions spatial information is processed and passed on to IT which enables the features to be associated by where they are in relation to each other to form the collective hypothesis in IT.

From IT connections go directly to the MTL where two things happen. Top-down information in the form of stored knowledge is matched against the

hypothesis in IT and if a match is made identification of the hypothesis occurs. If however there is no previous recollection of the representation in IT it is stored in memory for future use. This aspect of the model is speculative and in relation to biological vision there is no concrete neurophysiological evidence to provide possible mechanisms for how that matching occurs. However in relation to artificial vision systems there are possible pattern matching algorithms that could support this aspect of the model (Carpenter & Grossberg, 1987; Carpenter & Grossberg, 1987). At the same time top-down information in the form of other sensory processing is incorporated via input from poly sensory areas namely the STG, STPp and STPa. Sensory information from other areas can be crucial in the identification of a hypothesis if visual information alone fails. It is also necessary to incorporate information from other senses for complete perception. This information is then fed back through the visual system for confirmation, detailed analysis and integration. This confirmation occurs where feedback connections converge on the dLGN. The dLGN acts as a "detail assessor" checking that the identified representation in IT correlates with the incoming data from the retina.

In the PARIETAL region the hypothesis is interpreted in terms where things are in relation to each other in the visual image. This is continually updated so that objects that are moving can be monitored at all times. This visuo spatial information can then be used in the planing and execution of actions in relation to the visual image. The PARIETAL region also connects to MTL so that top-down stored knowledge about spatial relationships can also be incorporated in the spatial analysis and prediction of movement. Direct

links to the PFC facilitate motor actions and working memory. Connections with IT are also used for the facilitation of action through object identification in working memory. In sum then, a hypothesis has been created and tested both in relation to top-down and bottom-up information. Figure 18 gives a simplified example of how the model might function.

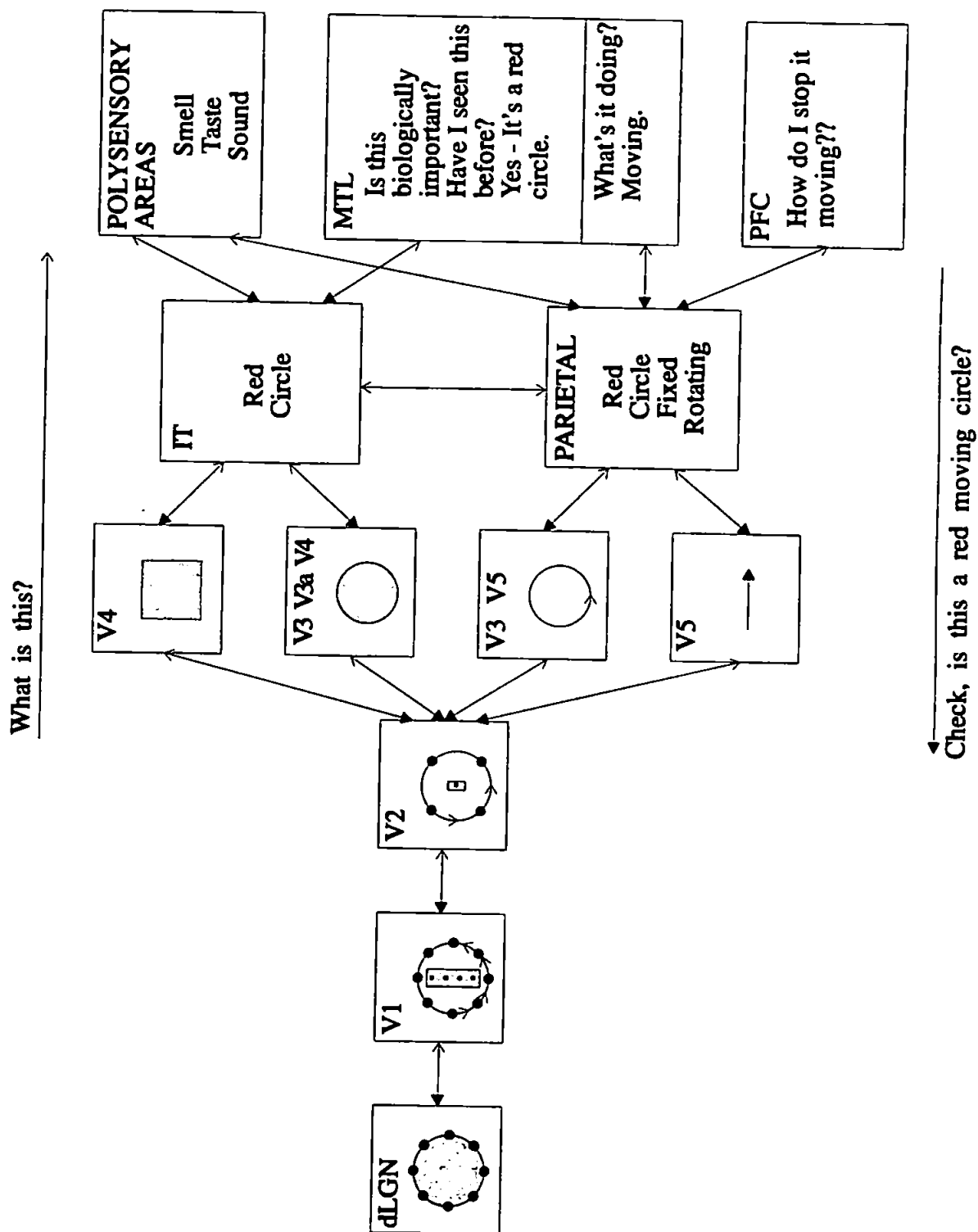


Figure 18. A simple representation of visual processing in the model.

The dLGN acts as a blackboard providing the most up to date representation of the visual image from the retina (Mumford, 1991). V1 carries out

detailed form processing recognising edges blobs and bars and is also responsive to motion and colour information. The representation in V1 is not of for example a red circle but of a series of edges, bars, blobs and termination in given orientations, with wavelength sensitive cells responding to colour. V2 then holds a more abstract version of V1 where the edges, bars, blobs and termination relate to larger receptive fields. It still does not represent a red circle. V4 responds to colour information alone, but V3 and V3a combine input from V2 about form with V4 colour information to hold a representation of what forms are what colour. V5 responds to motion information which is combined in V3 to represent form and motion together. Within the model IT is the point at which a representation is formed that could be classified as “red circle”, but has no detailed information about form colour or motion. This representation takes the form of groups of cells that fire in response to all the attributes associated with red circle, but not to any specific details. The response is to the high level representation “red circle” rather than to detailed form, colour or motion information. MTL provides the mechanism which enables access to stored knowledge against which the representation in IT can be compared to and identified with, as well as enabling new information to be stored for future identification purposes. It is important to emphasise at this point that in relation to the model the role of the MTL is to facilitate memory. Detailed descriptions of the structures and mechanisms that underlie memory are beyond the scope of this thesis.

MTL also provides the same service for the PARIETAL region. The PFC can generate and co-ordinate the execution of actions relating to the visual



stimulus through input from the PARIETAL region. The connections between the IT and the PARIETAL region, direct and indirect via the MTL, perform a number of functions including:

- The integration of spatial information with object identification. IT relies on the PARIETAL regions for information about spatial location of a given object in relation to other objects. The PARIETAL region knows that motion is occurring in a particular form in a particular direction, but relies on IT for information as to the identification of form.
- An aid for object recognition when incomplete information is available to IT. Some things are not easily recognisable from their visual representation. Procedural information from the PFC such as how something is used may facilitate object interpretation in IT via the MTL.

The poly sensory areas enable the incorporation of other sensory information detected in relation to the object, for example smell, taste, and sound. Two way connections from the poly sensory areas to IT and the PARIETAL regions also perform a number of possible tasks including:

- Facilitates recognition in that if there is insufficient top-down information to identify an object other sensory information may provide invaluable association cues.
- Reconciliation of ambiguous bottom-up information, for example determining a real object from a replica object such as a real apple from a wax apple using olfactory information as well as visual information.
- Following on from the previous example, facilitating the correct response to an object when there is a choice. For example a hot baked potato requires careful handling, although visually the cues do not distinguish

between hot and cold, smell can elicit information that would enable care to be taken in its handling.

Meanwhile feedback which is continually occurring at all levels ensures that the representation is maintained in the visual system. It also enables adjustments to be made to the processing at each stage based on top-down information. Such adjustments might occur as a result of extra information from poly sensory processing and from memory. This may mean that the representation fed back through the visual system has more detail than the original representation resulting from feedforward retinal input to the system. These adjustments can then be checked against the representation in the dLGN for confirmation against incoming bottom-up data. This is particularly necessary for new visual examples that have not been experienced before. It is essential that a correct representation is committed to memory.

#### MECHANISMS UNDERLYING ARCHITECTURE AND FUNCTION.

Within the model it is proposed synchronous oscillatory activity will underlie the maintenance of a distributed visual representation for integration, detailed analysis and hypothesis checking. This will mean that the visual representation is distributed across the whole system as opposed to the single cell approach discussed in chapter 3, which suggests that at the highest level of representation there are cells that respond to given objects. In the model in IT there are cells that respond to extremely abstract properties of the visual scene but the representation of the visual image is distributed across the whole system. It is then integrated or "bound"

together by neurons all oscillating in synchrony. This oscillatory activity is a function of feedback connections throughout the system. Oscillatory activity has recently become the focus of much attention in the role it has to play in visual processing. It has been suggested as a mechanism for feature linking or binding in early visual processing and there is a wealth of literature, both neurophysiological and simulation based, providing support for its role in visual processing (see Basar & Bullock, 1992; Denham & Troup, 1992; Troup, 1991 for comprehensive reviews and detailed discussion of the nature of oscillations).

As with most models of visual processing feedforward connections enable neurons to act as "feature detectors". The role therefore of feedforward connections in the model is to enable neurons in feedforward processing to respond in an excitatory fashion to certain features if they are present in the visual image. Therefore neurons in V1 are firing in response to detailed aspects of form such as edges bars and blobs in given orientations. Neurons in other modules will respond to other particular aspects of the retinal image. Within the model diffuse feedback connections produce synchronous oscillations in the relevant neurons namely those that have received retinal input allowing them to be associated with each other in each module as well as across modules. As feedback occurs at all levels neurons in V1 may produce synchronous oscillations as a result of feedback from V2 before recognition in IT occurs. This means it would be possible to identify the colour of something before it is actually perceived. The hypothesis about colour could be tested against thalamic input before recognition in IT has occurred.

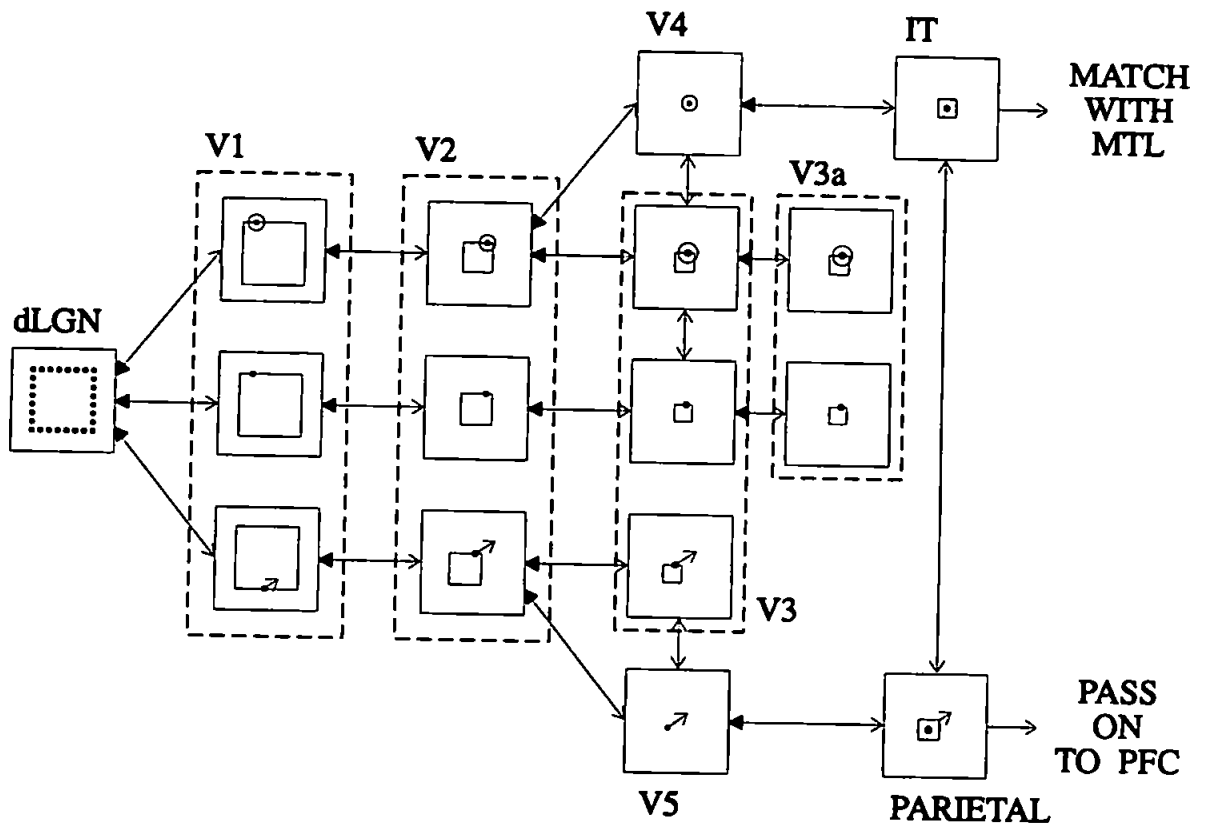
At the very highest level neurons in IT respond to complex features and in some instances actual objects. The pattern of response generated in IT is then matched against or compared to memories for instances of previous patterns via the MTL. If a match is made then synchronous oscillations are induced in the neurons representing the features that fit with the stored pattern. Top-down information in the form of poly sensory input that represents other sensory information associated with the visual image also contributes to the generation of oscillatory activity. It is also feasible that poly sensory activity associated with the visual stimulus could also oscillate in synchrony with neurons in IT representing that stimulus. It is worth noting that the extent of oscillatory behaviour in neurons in high level processing is speculative in regard to the neurophysiology of visual processing. In terms of the aims of developing the model oscillatory activity has to be speculative in relation to understanding human visual perception. However in the development of an artificial model of visual perception it could be extremely relevant. Direct connections with the PARIETAL region allows for spatial information to be incorporated into the representation in IT. Neurons in IT oscillate in synchrony with neurons in the PARIETAL region relating to their location.

Synchronous oscillations also provide a mechanism for integration or feature linking within modules as well as across modules. Lateral connections allow information with similar receptive field properties to be associated or bound together. This would mean that the various edges, blobs and bars that might make up a particular line in a given orientation in V1 can be associated or bound together through synchronous oscillations.

However this kind of processing is localised so it does not allow for the integration or binding of information in edges, blobs and bars having different receptive field properties, yet belonging to the same object.

Integration of this nature is actuated by diffuse feedback connections and not localised lateral connections. (for a detailed discussion based on neurophysiology and simulation work of laterally induced oscillations see Basar & Bullock, 1992; Troup, 1991; Denham et al, 1991; Denham & Troup, 1992).

To implement the model poses some extremely complex problems even at the theoretical level. Figure 19 illustrates how this might possibly be done.



**Figure 19.** A schematic diagram of a possible implementation of the model.

By comparing the diagram above to that in chapter two reflecting the extremely complex functional structure of the dLGN shows how little

information can actually be extracted from the biological visual system in the development of artificial vision systems. For the purposes of possible implementation the retina has been omitted and retinal input reduced to a pattern of activity in a module likened to the dLGN. As mentioned previously in chapter 2 there are already detailed models of the retina available. Input in this theoretical implementation is a one dimensional red square moving slowly across the visual field to the left. It is represented in the dLGN as three grids of excitatory neurons. Visual field information as well as ocular dominance information has been omitted and the dLGN has been reduced to a simple retinotopic representation of form, colour and motion information. As with the biological visual system this is processed separately reflecting X, Y and W information pathways. This representation is then mapped directly onto a module representing V1 in the same way reflecting the 1:1 relationship between the dLGN and V1. Again the main theme carried across from the biological vision system is that of functional separation rather than anatomical detail. The form information is processed separately from motion and colour information in readiness for separate streams of processing later on. Again the complex way in which the biological visual system expresses its functional architecture has to be greatly simplified. If a model of processing purely in V1 was being developed it would be possible to reflect this detail far more accurately. This however is supposed to be a model of visual perception rather than a model of specific visual processing. This is then mapped again directly to the V2 module where the receptive field size has been reduced significantly. The square now consists of a grid

for form, colour and motion information. At this point the information is split into three pathways. One to V4 where a cell fires in response to a given colour, in this case red. Another to V5 where the same occurs for motion in a given direction. Lastly one to V3 where form information is incorporated with colour or motion information with even larger receptive fields. V3 then projects to V3a but only for colour and form information. V3 grids for colour and form connect to V4 whereas the grid for motion connects only to V5. V3a projects directly back to V3 enhancing colour form projections to V4.

V4 then sends information to IT where a pattern of activity occurs which represents the object as red square. It does not reflect specific information although in the biological system there is some evidence for retinotopy to exist in IT. The same occurs in the PARIETAL region, but here cells represent spatial information in relation to each other using a co-ordinate system. Finally the pattern in IT is matched with patterns stored in memory, facilitated by the MTL, for identification. A positive match results in neurons in IT oscillating, which in turn, through diffuse feedback connections, causes oscillatory activity in all the neurons at the various other stages in the visual system. This oscillatory activity is then correlated against the retinal input in the dLGN. A positive correlation results in a further strengthening of synchronous activity in V1 which in turn influences the strength of oscillatory activity across the whole system. Synchrony across the system enables binding and integration, and a positive correlation indicates a confirmed hypothesis.

## PREDICTIONS OF THE MODEL.

- Top-down information from other sensory areas can cause a representation in IT to be active, i.e. a group of cells oscillating in synchrony, before feedforward visual information has reached IT. Information still has to feed back through the system for detailed analysis and for confirmation with the retinal input in the dLGN.
- Top-down information in the form of either a memory or other sensory information is sufficient to evoke a pattern of synchronous oscillatory activity in the model without retinal input. Such a pattern of response would be weak compared to a retinally induced one, and it would not be possible to check the visual hypothesis against a retinal input correlated in the dLGN. This would mean that if blindfolded you were asked to identify a lemon using smell and touch as sensory inputs a visual representation would be generated. You would not however be able to confirm that the lemon was in fact a lemon, as you would not be able to confirm the hypothesis that other sensory information had provided. If for example the lemon had been coloured green it would be impossible to check. The pattern of activity in V4 would say yellow, as the associated colour for lemon is exactly that, however without retinal input it would be impossible to test this hypothesis.
- It is also possible that if the response pattern in IT is incomplete and that the stored pattern consists of features not present in the retinal image then neurons not responding could be induced to do so due to MTL input. This means that incomplete visual information can be completed by comparison of a partial pattern to a stored complete pattern.



# **CHAPTER 5**

## **SUPPORT FOR THEORY AND MODEL:**

### **ADDITIONAL SUPPORT FOR BOTH THE THEORY AND MODEL.**

*This chapter provides additional support for the theory and conceptual model explaining perception outlined in the previous chapter. Firstly it addresses support from neurophysiology and psychology for the claim that visual perception has a strong top-down component. Secondly it looks at how neuropsychology also provides support for various aspects of the theory and model notably the top-down component of visual perception. Thirdly support for the role of oscillations as a possible mechanism of visual processing in the model is discussed.*

### **INTRODUCTION.**

Support for both the theory and conceptual model originates initially in chapters 2 and 3. This aim of this chapter is to provide additional support for certain aspects of both the theory and model that have not previously been made explicit.

## **VISION AS BOTH A TOP-DOWN AND BOTTOM-UP PROCESS.**

### **SUPPORT FROM NEUROPHYSIOLOGY AND PSYCHOLOGY.**

The conceptual model of visual perception presented in the thesis has been based upon a general theory of perception that claims vision to be both a top-down and bottom-up process. Based upon this, visual perception is explained in general terms as the construction of hypotheses which are identified through comparison to previous knowledge, and then verified against incoming data.

As mentioned in chapter 3 it is not a new idea that visual perception is reliant on top-down processes in the form of memory being used to make sense of bottom-up incoming sensory data. Neurophysiological descriptions of the visual system indicate both massive feedback as well as feedforward pathways are in operation indicating that deep brain structures are exerting their influence on incoming sensory data. Such structures include those thought to be associated with the storage and retrieval of memories.

Recently research in developmental neurobiology has provided architectural support for the claim of the model that visual perception is strongly reliant on top-down processing in the labelling of bottom-up sensory data. It is widely reported that the human visual system is immature at birth (see Burkhalter, 1993 for a review). Recent research has shown that both feedforward and feedback connections in the visual system only emerge just prior to birth. It is not until 4 months of age that feedforward connections are developed in the same way as the mature adult cortex. Importantly

feedback connections are at this stage immature by comparison (Burkhalter, 1993). Burkhalter suggests that this implies that in the first few months of life the infant is only capable of local processing, for example image segmentation in V1, it is not until later on that the infant is capable of global processing in the form of object identification.

It is possible to interpret these findings more generally to provide support for the theory underlying the conceptual model. If the visual system is reliant on memory for the labelling of hypothesis then it would be expected that the focus of the immature visual system is to develop a repertoire of memories that can be used to aid identification. Therefore the first few months of the infants life are devoted to the development of feedforward pathways to enable a repertoire to be constructed. It is only when enough sensory stimulation has been encoded and stored that the feedback pathways are of any use. Therefore they develop more slowly and do not mature as quickly as the feedforward paths. So it is not until the visual system is capable of processing both local bottom-up information and global top-down information that perception is complete. It is possible to provide further support based on the above interpretation in the developmental psychology literature.

There is a whole host of experimental evidence from developmental psychology to suggest that the functional capabilities of the visual system are extremely limited at birth (for example Banks, 1980; Banks & Salapatek, 1983). Infants will respond to high contrast patterns (Banks and Ginsburg, 1985) and objects that are moving (Slater et al, 1985) but form perception is limited (Banks & Ginsburg, 1985). This fits with the recent

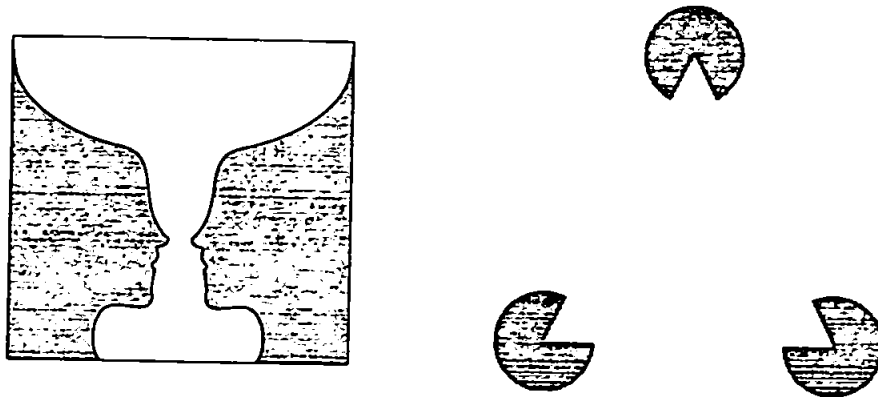
neurophysiological findings outlined in the previous section that suggest the visual pathways are not fully developed until well into the first year of life. Research into infants ability to store patterns also provides experimental results that support the neurophysiological findings in the previous section. It seems that infants under 2 months respond equally well to new patterns as they do to ones which they have experienced before. However infants over two months of age respond better to patterns they have experienced before (Olson & Sherman, 1983; Banks & Salapatek, 1983). There is an exception to this in that genetically important patterns such as the pattern of a mothers face produce responses a few days after birth (Bushnell et al, 1989). This supports the neurophysiological data suggesting that the visual system is immature at birth. Further it supports the theory outlined in chapter 4 in that it suggests that memory or experience is essential for perception, in that we need top-down information to perceive the world. Developmental psychologists have even gone as far to claim that the very young infant is essentially a "stimulus seeker" with a biologically programmed disposition to actively search for novel stimuli keeping the visual neurons active to promote development (Haith, 1980). In relation to the conceptual model this could be interpreted as building a repertoire of experiences for the future labelling of hypotheses.

The role of past experience in perceptual processes is well documented in psychology, and there is a great deal of experimental evidence supporting top-down processes in the interpretation of bottom-up data. Research by Biederman into how context affects ones ability to identify objects demonstrated the importance of top-down information in perceptual

recognition tasks. Subjects were asked to identify both contextually related and unrelated objects in real world scenes. Performance for contextually related objects was significantly faster and more accurate than for non related objects (Biederman, 1972; Biederman et al, 1973). Exposure to stimuli facilitating recognition is also well documented in psychology. In the 1960's Haber and Hershenson demonstrated in a series of experiments that in a word recognition task the more times a string of letters was presented the more likely subjects would perceive them correctly (Haber & Hershenson, 1965). As well as supporting the role of top-down information in visual perception it is also possible to demonstrate hypothesis testing experimentally. In an experiment which involved showing two images one to each eye tachistoscopically Engel was able to show how subjects interpret bottom-up visual data. Subjects were shown two identical stationary figures, one to each eye. One figure was naked and the other clothed. When subjects were asked to report what they had seen they either said a figure getting dressed or a figure doing a "strip tease". Subjects had perceived motion that was in fact not present to make sense of the visual information they had received (Engel, 1956).

Visual illusions provide further evidence for the role of top-down processes in perception. Many visual illusions present the visual system with a representation that has more than one possible interpretation yet only one retinal image. Examples include the goblet/two faces illusion (see figure 20), the mouse/old man illusion, the young woman/old hag illusion etc. As well as ambiguous illusions there are illusions which involve the visual system actively adding information that is not actually present in the retinal

image. Kanizas triangle is one such illusion (see figure 20).



**Figure 20.** Examples of visual illusions: The goblet/two faces illusion and Kanizas triangle.

Here the visual system interprets the retinal image as a triangle even though there is no actual triangle present, only the illusion of one. The interpretation of such illusions is reliant on top-down information, not on bottom-up processing alone. There is also experimental results showing that contextual effects influence the visual systems interpretation of ambiguous illusions. If subjects are shown pictures of young women before being shown the young woman/old hag illusion almost 100% report seeing a young woman as opposed to 65% when not shown the pictures first (Leeper, 1935). The same effect has been demonstrated using the mouse/old man illusion but instead of specific pictures of a rat, various animal pictures were presented showing generalisations in interpretation (Bugelski & Alampay, 1961).

Other top-down information which facilitates perception includes genetically predetermined information and information from other sensory processing areas. Genetically predetermined input to visual perception is much debated, as with most psychological research there is evidence both in

support and against the claim that visual perception is innate. Animal studies have shown that certain biologically important visual behaviours are present from birth. Herring gull chicks will peck at the red spot on the adult birds beak as soon as it is born, before it has made the association between food and the red spot (Tinbergen, 1952). There is evidence to suggest that certain visual information in humans is genetically determined. Experiments in developmental psychology have shown that infants are capable of demonstrating certain innate visual behaviours such as depth perception (Gibson & Walk, 1960) and object constancy (Bower, 1966). It has also been suggested that infants are able to recognise their mothers face at 3 months of age (Barrera & Maurer, 1981), even though they are unable to easily discriminate between carers and strangers faces until they are 5-7 months old (Cohen et al, 1979). According to the research discussed in an earlier section the visual system is neurophysiologically incapable of processing the kind of global information needed for face recognition. This could be interpreted as being support for the suggestion that some biologically important visual information such as face identity could be facilitated by pre determined information about faces, enabling a certain amount of perception to occur without repeated experience. This may also explain the existence of cells in IT responsive to faces.

More recently it has been suggested that "backward masking" is controlled by the top-down process of voluntary visual attention (Ramachandran & Cobb, 1995). Backward masking is where a target stimulus is presented followed by a non-target stimulus, causing subjects to be completely unaware of the existence of the original target stimulus. Gregory goes

further to suggest that bottom-up data is "topped up" by top-down information (Gregory, 1995). Gregory's interpretations can be seen as further support for the role of top-down information in both the theory and conceptual model.

#### SUPPORT FROM NEUROPSYCHOLOGY -BLINDSIGHT.

The existence of certain neuropsychological disorders affecting visual perception also provide support for the theory and the model, blindsight mentioned previously in chapters 2 and 3 is one such disorder. The destruction through illness or accident of the striate cortex leaves a person clinically blind. It must be reiterated that vision in patients suffering from blindsight can be partially, rather than fully impaired. Research has shown that some of these supposedly clinically blind patients are in fact able to make judgements about stimuli which they reportedly claim they are unable to "see". This is what is referred to as blindsight, and has been extensively investigated (for reviews see Weiskrantz, 1986; 1992). As the amount of information they can report about such stimuli is restricted to things like whether or not it was moving, or what colour an object was rather than detailed form information, it seems that information about the stimulus is a product of processing in the extra striate cortex. Arguably the existence of blindsight could be interpreted as demonstrating that perception is not purely a bottom-up process. If this was the case destruction of the striate cortex should result in complete blindness. The ability of people suffering from blindsight to make judgements about visual stimuli however limited suggest that visual perception is still possible even if the bottom-up



pathways are destroyed.

Attempts to explain the existence of blindsight have given rise to two main theories, briefly mentioned in chapter 3. The first suggests that the secondary visual pathways mentioned in chapter 2 allow sensory processing to occur. The emphasis originally has been on the pathway leading from the superior colliculus (SC) via the pulvinar to V5. However more recently, research has shown that the destruction of other pathways, specifically the lateral pretectum and the accessory optic system, have a greater effect on the ability to see with existing damage to V1 than destruction of the SC. Further to this some blindsight patients are able to detect colour information in a stimulus, since the SC is unable to transmit wavelength information this suggests that other pathways are responsible for the transmission of colour information (see Cowey & Stoerig, 1992 for a review on the pathways mediating blindsight). The second theory suggests that the secondary visual pathways do not contribute to the ability of blindsight patients to make perceptual judgements. Perception in cortically blind patients is due to scattered light falling on areas in V1 where cortical tissue has been spared, and are able to function as normal (Campion et al, 1983). Neither theory is conclusive, there are lots of possible pathways that could account for the resultant perceptual judgements demonstrated in blindsight patients (see Cowey & Stoerig, 1992).

In relation to the theory blindsight provides support for its claim that visual perception results from the construction and subsequent testing of hypotheses about the world by the visual system, a combination of both top-down and bottom-up processing. In terms of the model a blindsight patient

would at best only be able to construct limited hypothesis. Detailed form information would not be available due to the damage to V1. Although there are at present no secondary pathways defined in the model it is possible to speculate their ability to carry limited information to the extrastriate areas. For example it is possible to envisage a pathway from the SC to V5 via the pulvinar. The processing that is carried out in V5 would only be able to provide a hypothesis of "motion" and limited spatial information but nothing more. It is possible that the information from the secondary visual pathways is not sufficient to engage the memory processes, meaning that object labelling would not occur. Further to this it would be impossible due to damage to V1 to "test" any hypothesis against incoming data in the dLGN. This would mean that conscious perception would not be possible. However it would be possible to send the resultant motion processing information on to other cortical areas such as the motor cortex, or the polysensory areas for integration with other processing. This would mean it may be possible to point to a visual stimulus without being visually aware of it. The importance of top-down information in the labelling of hypotheses is emphasised by other neuropsychological disorders such as prosopagnosia. Sufferers from prosopagnosia are impaired in certain aspects of perceptual recognition in that they are generally unable to identify the faces of familiar people, they can however differentiate between faces (for a general review of the literature see Young & De Haan, 1992). In relation to the model this demonstrates an ability to construct a hypothesis but there is a problem in labelling that hypothesis. Feedback would still be possible meaning that the hypothesis could be "checked" but it would not have a label hence would

not be identified. This indicates that visual perception of some sort is occurring, but complete visual perception is not. This will be discussed further in chapter 6 in relation to levels of processing as outlined in chapter 1 and how they relate to the model outlined in chapter 4.

Another form of agnosia, associative visual agnosia, can be seen to provide support for the claim made by the model outlined in chapter 4 that top-down information in the form of poly sensory information is also important in visual perception. Some patients suffering from associative visual agnosia are able to replicate drawings of objects with ease, but when asked verbally to draw objects they are unable to do so. They are unable to associate the verbal description with the visual description. When asked to describe verbally the object they have been unable to draw they can do so without problem so it seems they are aware of what it is they have been asked to draw even though they are incapable of doing so (Ratcliff & Newcombe, 1982). Another case study of a person suffering from some form of visual agnosia also provides evidence to support the role of poly sensory processing in visual perception. This particular patient was unable to recognise objects visually. However when studying pictures of objects he made unintentional gestures with his hands that in some way related to the picture he was studying. The hand gestures enabled him to identify the objects in question (Carlson, 1986). The model suggests that if a hypothesis is unlabeled through incomplete visual input or through failure to match with a stored memory then the result of processing from other sensory areas can facilitate recognition. It could be that the motor response pattern generated by the agnosic facilitates the retrieval of a stored description of

the object. This can then be instantiated at all the possible levels of sensory description previously associated with the object. Damage to the visual processing areas of the brain would mean that it could not be fed back through the visual system and "confirmed" at the dLGN. However it could be fed back through other sensory processing pathways and identified at a different level of description.

The evidence presented by neuropsychology is in no way conclusive. As mentioned in chapter 3 it is extremely difficult to determine the exact nature of damage to the visual cortex resulting in neuropsychological disorders. Further, most disorders are not clearly defined and tend to be partial rather than complete. However it does seem that there is a great deal of psychological as well as neurophysiological evidence that can be used to support the model and especially the theory outlined in chapter 4.

## **SUPPORT FOR MECHANISMS UNDERLYING PERCEPTION.**

### **INTEGRATION OF INFORMATION THROUGH SYNCHRONOUS OSCILLATORY ACTIVITY.**

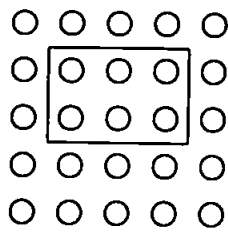
Within the conceptual model it is proposed that the problem of integration of information can be solved by synchronous oscillatory activity. Such activity serves to integrate information both within modules via lateral connections and across modules via diffuse feedback connections. Further to this synchrony as a function of feedback also provides a mechanism for confirming hypotheses about the world by correlating them against incoming data. The key to such a mechanism is that the visual image can be represented by distributed temporal activity across populations of neurons rather than by dedicated processing in specific neurons. This means that a given neuron can become part of the representation for more than one stimulus.

The role of oscillatory behaviour in the integration of distributed information was introduced in chapter 3. This work had a major impact on both the development of our understanding of visual processing and the development of artificial models of visual perception. Since the early work investigating oscillatory behaviour a great number of models have been developed that exploit synchrony as mechanism of feature linking. This early work, both experimental involving the recording of cells, and simulation based involving the development of models, suggested that synchrony in oscillations occurred in two ways. Firstly through lateral

connections between neurons or groups of neurons with similar receptive field properties, and secondly by diffuse feedback connections linking neurons with unrelated receptive fields (see Werner et al, 1993 for a review).

In a series of neurophysiological experiments Gray and Singer demonstrated that neurons in the cat visual cortex oscillated at a frequency of around 40 Hz. They also demonstrated that these oscillations correlated with a rhythmic firing pattern, suggesting that groups of neurons were using temporal information to synchronise their oscillations in spatially separate areas of the visual cortex, (Gray et al, 1987; 1989). Their experiments showed that groups of neurons within a functional column with the same receptive fields responded to a moving stimulus by discharging rhythmically. Gray and Singer concluded from the experimental results that when responding to a given stimulus, adjacent neurons in the stimulated receptive field are more likely to fire at the same time and in synchrony. Further more this phenomenon is not a product from interference from other areas of the brain such as the thalamus, which is known to produce oscillation. In the discussion of their results Gray and Singer suggest that although the underlying mechanisms that are producing this behaviour are unknown, it seems that simple inhibitory and excitatory links between neurons is a sufficient explanation as to how oscillations are produced, (Gray & Singer 1989) (see figure 21).

## Light Bar Stimulus



## Rhythmic firing



## Correlation



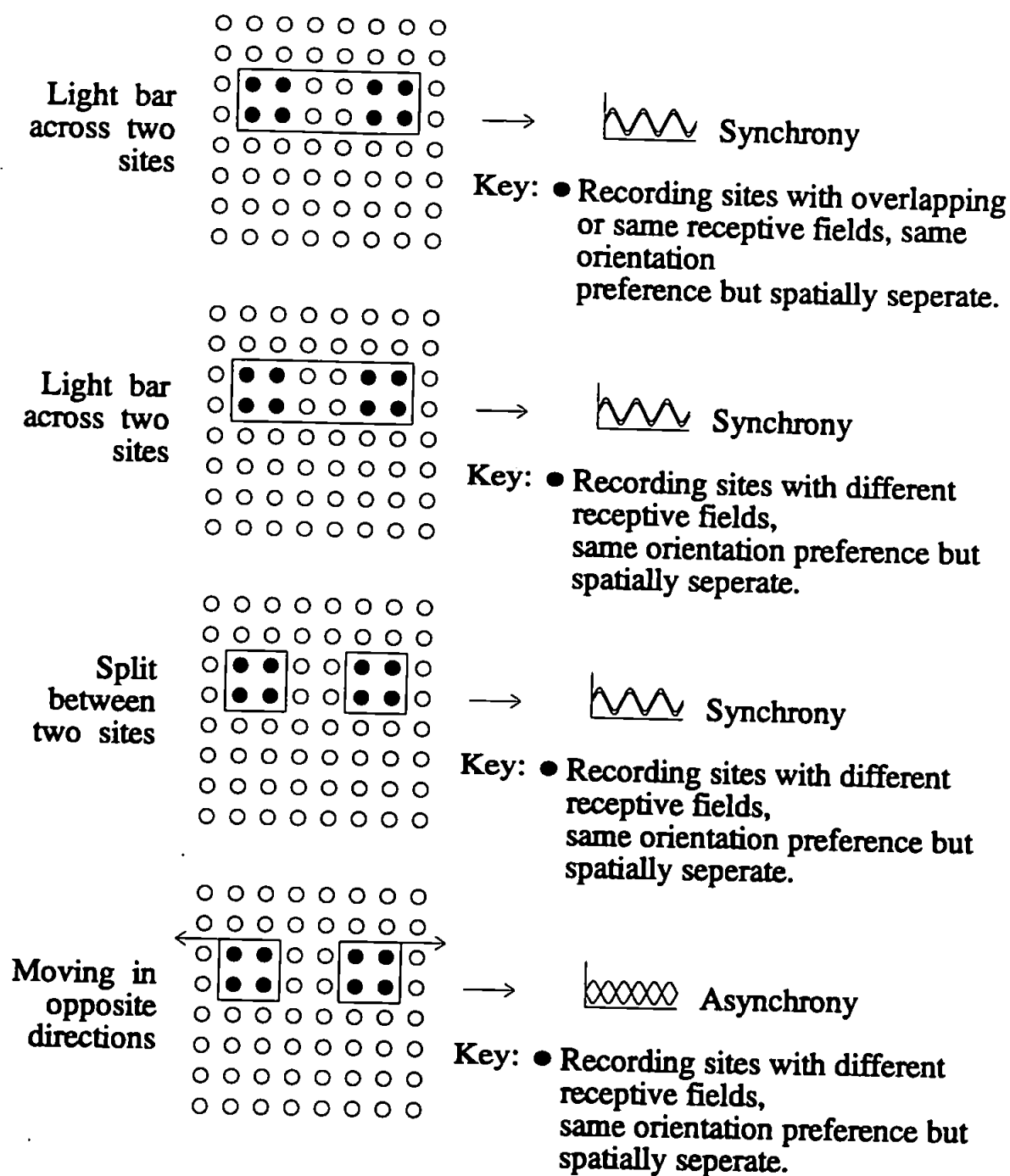
## 40 Hz Oscillations

**Figure 21.** A summary of Gray and Singers findings.

In more detailed experiments Gray et al went on to show that these synchronous oscillatory responses explained how spatially separate features of a given stimulus could be bound together, (Gray et al, 1989). They recorded at between 5 and 7 spatially separate sites in the visual cortex, and used a cross correlation function to determine whether the oscillatory responses were present at any of the sites when stimulated. They then selected from 132 recordings, 99 pairs of sites producing oscillations that had the same orientation preference, also having overlapping receptive fields. To determine whether synchrony (or phase locking) was present, a cross correlation function was applied to the results of the oscillations of these pairs when they were stimulated by a single bar of light falling across both sites. It was demonstrated that synchrony was present when the two sites were stimulated by the single bar. The conclusion is that synchronisation of oscillation was being used to identify the stimulus as a long bar in a given orientation (see figure 22). In a second experiment recordings were made in pairs of sites as far apart as 7 mm. These sites did not have the same receptive field, but still shared the same orientation

preference. The rationale behind this was to see if synchrony was just a result of the receptive fields being the same. If stimulation was achieved through the use of a stationary light bar that fell across two sites, the oscillatory responses were in synchrony. If the light bar was split into two separate stationary bars of the same orientation, each falling across one of the sites synchrony was also evident. If the two bars of light were moved across the recorded areas in the same direction again synchrony is evident. However if the two light bars were moved across the two separate areas in different directions there was no synchrony in the oscillations (see figure 22).





**Figure 22.** Gray and Singer's 4 experiments.

Gray et al conclude:

*"We propose that the synchronisation of oscillatory responses in spatially separate regions of the cortex may be used to establish transient relationships between common but spatially distributed features of a pattern."*

(Gray et al 1989).

Through the mechanism of synchrony it is possible to establish overall or "global" features including continuity, orientation preferences, and consistency of motion, (Gray et al 1989).

Further research based on these findings explored the results more extensively with a larger data set, (Engel et al 1990). As in the previously mentioned work the study aimed to investigate activity in functional columns. Measurement of the local field potential (LFP) indicated that functional columns of neurons, each column having a different orientation preference, oscillated in synchrony. By recording multi-unit activity (MUA) it was demonstrated that groups of neurons in these functional columns also demonstrated oscillatory behaviour. By taking the local field potential it could be seen that neighbouring columns did not contribute to the oscillations produced by a given column. The cross correlation of MUA activity in the pairs of recording sites produced the evidence for feature detection through synchrony. In this more detailed analysis previous results were confirmed and a more detailed understanding was created. Firstly, the strength of synchrony depends on the distance between the spatially separate sites. The further the recordings were apart, the weaker the synchrony. With groups of cells having non-overlapping receptive fields synchrony is a product of orientation preference, as well as being sensitive to stimulus features such as motion. It was found that if the stimulus was changed, for example a stationary bar of light moved, the reaction of the groups of neurons producing oscillations changed, indicating that the oscillations encoded feature specific information and did not just respond to

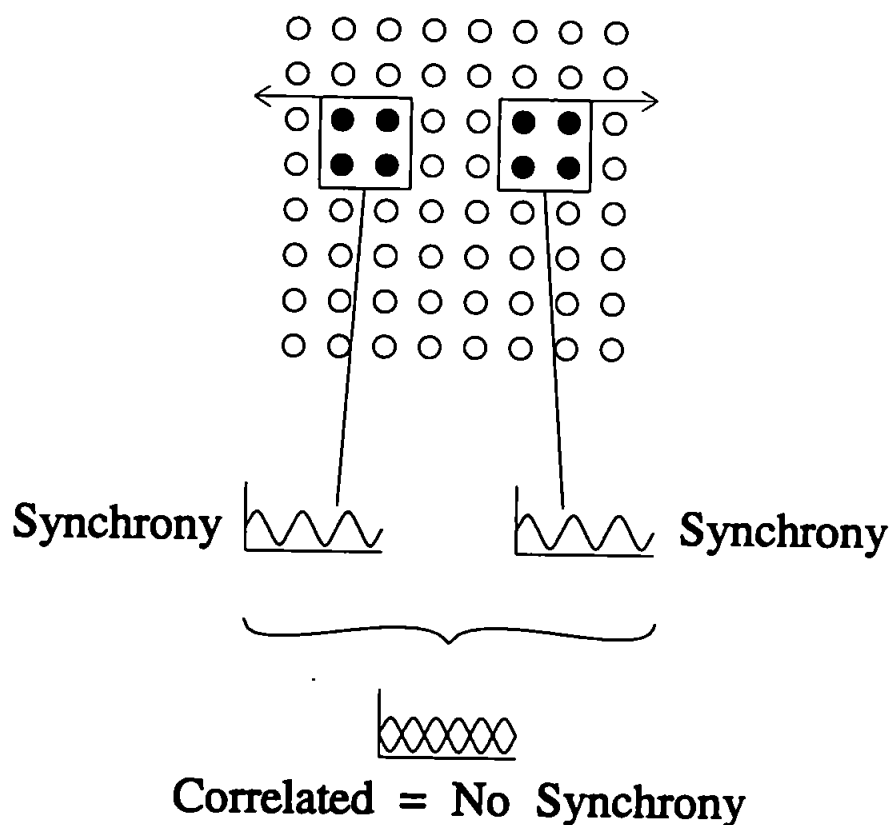
any kind of stimulation with the same pattern of activity. Also a long single bar produced stronger oscillations than two separate bars, (Engel et al, 1990).

There is a great deal of controversy as to the relevance of this kind of neurophysiological research. Phase information has been largely ignored on the grounds that it is unreliable. Most of the problems can be traced to the methodology used to extract the data. The wrong kind of anaesthetic can effect the recordings taken, and incorrect analysis can also be a problem, (Engel et al, 1990). However the results obtained by this recent body of research have overcome these problems.

Parallel work carried out over a similar period of time by Eckhorn et al made further distinctions as to the nature of this oscillatory activity (Eckhorn et al, 1988; 1989; 1990; Eckhorn & Schanze, 1991; Eckhorn, 1991). They distinguish between two types of oscillatory activity, stimulus forced, and stimulus induced. Stimulus forced synchronisations are driven by the image itself and are generally coarse representations of the visual scene, referred to by Eckhorn et al "crude instantiations" (Eckhorn et al, 1990). Stimulus induced synchronisations are internally driven and are thought to be the mechanisms by which more detailed and complex associations are made. The mechanism for synchrony in both cases is thought to be by feedback and in the stimulus-induced synchrony through mutual coupling of local oscillators, (Eckhorn et al 1990; Eckhorn & Schanze, 1991; Eckhorn, 1991).

The results of these experiments have led to suggestions that synchronous oscillatory activity as a result of neurons responding to particular objects

can be likened to Gestalt criteria such as proximity, continuity, similarity and common fate (Engel et al, 1992; Sompolinsky et al, 1990). In Gray and Singers experiments synchrony was evident in non overlapping receptive fields only if the stimulus was a continuous bar. If two bars moving in opposite directions was observed both groups produced synchronous oscillations but there was no correlation in the oscillatory activity between the groups (see figure 23).



**Figure 23.** Oscillatory activity between groups showing no correlation.

Engel et al suggest this is evidence to support the Gestalt principles for grouping such as continuity and similarity. If the object falls in the same receptive field but with differing orientation preferences synchrony can still be achieved through lateral connections. This is interpreted as representing criteria such as common fate and proximity.

Thus there is a great deal of evidence from these recent findings to support the use of synchronous oscillatory activity as a "binding mechanism" for early visual processing. Employing such a mechanism allows segments of the visual representation to be bound together to form coherent parts of that representation. The hundreds of edges bars blob and terminations that contribute to a particular line can be linked via synchronous oscillatory activity. In relation to the model stimulus forced synchronisation enables initial crude representations to be formed. Subsequent stimulus induced synchronisation through lateral connections enables more detailed associations to be formed. Within V1 and V2 in the model, segments that share the same receptive fields can be linked by coupled oscillators that form strong feedback connections, and do not necessarily rely on sharing the same orientation preferences (Sompolinsky et al, 1990). Segments that do not share the same receptive fields, but have the same orientation preferences, synchronise their oscillations via weak specific connections (Sompolinsky et al, 1990). In terms of viability in the conceptual model, there is a great deal of simulation work based on the experimental findings described so far, that demonstrate the ability to build artificial models capable of such behaviours (for example Cotterill & Nielsen, 1991; Denham et al, 1991; Denham & Troup, 1992; Dorizzi, & Grammaticos, 1991; Eckhorn et al, 1990; Eckhorn & Schanze, 1991; Eckhorn, 1991; Grossberg, & Somers, 1991a; 1991b; Kurrer et al, 1991; Kurrer et al, 1992; Sompolinsky et al, 1990; Troup, 1992).

Later experimental and simulation work has suggested that synchronous oscillations are also an interhemispheric phenomena suggesting that inter

area association or "binding" via synchronous oscillator activity within the model is feasible (Engel et al, 1991a). Engel et al also suggest that synchrony can be mediated by long range cortico-cortico connections. Synchronisation of oscillation between cortical areas has also been reported notably between V1 and V2 (Eckhorn et al, 1988). Further to this synchronous oscillatory responses have also been recorded between V1 and the posteromedial lateral suprasylvian area (PMLS) mediated by long range cortico-cortico connections (Engel et al, 1991b). This provides significant evidence to support association between striate and extrastriate areas via synchronous oscillatory responses. The PMLS is an association area thought to be primarily involved with the analysis of global motion information, responding to movement in large scale patterns. The PMLS also has significantly large receptive fields compared to V1. As already discussed in chapters 2 and 3, V1 is responsible for the analysis of localised detail of objects rather than global interpretations (Engel et al, 1991b). The existence of synchrony provides evidence to suggest that it could be implicated as a possible mechanism for binding between cortical areas coding for different properties of the visual image. This would suggest that temporal binding through synchronous oscillations is not dependent on the receptive field properties of neurons and therefore can easily bind information that is supposedly unrelated.

In sum the use of synchronous oscillations as a mechanism for binding or associating information both within and between modules in the model presented in chapter 4 is well supported in neurophysiology. Whether or not it is the sole mechanism for binding information in the biological brain is

still not conclusive. It does however lend itself well to simulation and therefore provides a strong candidate for binding both in the conceptual model and in the future development of artificial systems based on that model.

#### TOP-DOWN INPUT CAUSES OSCILLATORY ACTIVITY.

Within the model it is proposed that both bottom-up and top-down inputs generate synchronous oscillatory activity. Feedforward retinal inputs drive bottom-up oscillations allowing the retinal image to be represented across the visual cortex in the manner discussed in the previous section. Feedback connections provide a mechanism for both the integration of information across modules and also for top-down inputs to contribute to perception. As outlined in chapter 4, the visual image is interpreted through comparing it to previously stored examples in memory. A match induces oscillations in the relevant neurons in IT confirming the identification is correct. This in turn, through feedback connections, enhances synchronous oscillatory activity throughout the system. Synchrony is also induced in neurons that may not have received retinal input. Such neurons are driven by top-down input from memory and are known to belong to the visual image from past experience but are not explicit in the retinal input due to things like occlusion.

Recent simulation work has shown that it is possible to generate synchrony from top-down inputs fed back into the system (Bugmann & Taylor, 1994). They claim that in their model synchronisation is a product of visual recognition in higher levels of processing rather than causing visual

recognition itself. One of the predictions of their model is that synchronisation occurs at all 4 levels supporting the notion of inter area and interhemispheric integration through synchronous oscillation. The results of simulations partially uphold this prediction. An important point that they raise in the paper is that models of synchronisation to date all include artificial neurons that are capable of self sustainable oscillatory behaviours which enables synchronisation to occur. External input alone does not induce synchronisation in such models. If neuronal firing was driven entirely by retinal input then synchronisation would not occur. However this could be looked upon as a simulation problem in that this may be the case in relation to artificial models but does not mean that the biological vision system cannot produce oscillatory behaviour if it is driven by feedforward retinal input alone. The binding of information in V1 for example relies on stimulus induced oscillations according to the neurophysiology outlined in the previous section.

It is extremely difficult to assess the relevance of these claims in relation to understanding human visual perception as their model only attempts to explain a small aspect of visual processing. Also the model is only loosely based on functional and structural neurophysiology and is more concerned with the intricacies of simulation rather than the biological mechanisms of vision. Finally it addresses important issues, such as the binding problem without substantiating them in any way. However their model does demonstrate that for the purpose of developing artificial vision systems at least it is possible to induce synchronous oscillations as a result of top-down input. Therefore in relation to the model outlined in chapter 4 it is



possible to combine the results of this paper with neurophysiological knowledge to make certain predictions about visual perception which can then be both incorporated into conceptual models as in chapter 4 and investigated further both neurophysiologically and psychologically. An example being the claim that feedback could be considered as a possible mechanism for the integration of visual information processing in parallel architectures.

Further evidence to support the claim that integration in the conceptual model is driven by top-down feedback comes from neurophysiology.

Correlated firing in thalamic relay cells has been found to be driven by feedback inputs from V1 to the dLGN (Sillito et al, 1994). This they claim provides additional support for the role of synchronous oscillatory activity in feature integration. They go on to suggest that the role of this synchronisation via feedback is to maximise the capability of relay cells in driving cortical processing. They expand this by suggesting that the correlation of a cortical pattern of activity in V1 with thalamic relay cells is a mechanism for "testing" that the original pattern in V1 is still correct, and if it is, to enhance the relations between its component parts. The dLGN contains the most recent representation of the visual image which is continually updated reflecting the transient nature of the visual scene. The suggestion that the role of the thalamus extends beyond merely relaying information on to the cortex is not new. It has been widely implicated in models and explanations of visual attention (including: Desimone et al, 1990; Crick, 1984; La Berge et al, 1992; Taylor & Alavi, 1993). It is also central to many models and theories of sensory processing (Harth et al,

1987; Mumford, 1991; 1992). In relation to the conceptual model outlined in chapter 4 of the thesis Sillito et al's findings provide neurophysiological support for oscillatory activity underlying hypothesis testing. The transient thalamic representation serves to provide a pattern against which the hypothesis that has been generated through cortical processing can be tested. If it is correct then a correlation in the form of synchronous oscillatory activity occurs enhancing the pattern of activity representing the visual image in V1. It is also possible that new information about the visual image has become available, and can now be incorporated into the representation, thus the visual image is continually updated with new information. This also serves to aid the interpretation of ambiguous hypotheses which may have been unlabelled on their first iteration through the system.

#### INTEGRATION OF POLY SENSORY INPUT VIA TOP-DOWN SYNCHRONOUS ACTIVITY.

There is also evidence to support the proposal of oscillatory activity being the underlying mechanisms for integration from other sensory areas. Top-down synchronisation could be occurring as a result of input from poly sensory areas or directly from other sensory areas. Recent evidence suggests that in several different species synchronous oscillations occur between cortical processing areas notably between the somatosensory cortex and the motor cortex (Murthy & Fetz, 1991). Further synchronous oscillatory activity has also been found in the association areas which include poly sensory areas (Llinas et al, 1991). Llinas et al reported the presence of

oscillatory activity in layer 4 of the guinea pig frontal cortex which oscillated at the same frequency as previously mentioned cortical oscillations.

The implications of this in relation to the model are that input from other sensory processing areas could be associated with resultant visual processing through synchronous oscillatory activity. It is also possible that inputs from the poly sensory areas could induce oscillatory activity in the visual cortex as suggested by Bugmann and Taylor's top-down model of synchronisation. However as with much of this work this is purely speculative and bears more relation to the development of artificial vision systems than it does to explaining human visual perception.

#### INTEGRATION OF PAST EXPERIENCES VIA SYNCHRONOUS ACTIVITY.

It is proposed in the model that past experiences or "memory" is used to identify and label hypotheses that are constructed via feedforward pathways projecting to IT. The model suggests that a given hypothesis is matched or rather checked against stored exemplars for recognition purposes. As mentioned in chapter 4 it is not the aim of the thesis to provide detailed explanations of the module referred to as memory. (A comprehensive related account of a possible model of this module can be found in O' Shaughnessey, 1995). In relation to developing artificial vision systems it is acceptable to check patterns of oscillation in IT with stored exemplars in the module of the model referred to as memory. If a match or close match is made then identification can be facilitated. This then leads to induced

synchronisation in the related neurons across modules via the feedback pathways.

Relating the conceptual model to the understanding of human visual perception it is a little tenuous. There is however evidence to suggest that synchronous oscillatory activity is present in brain structures that are implicated in memory related activity such as the entorhinal cortex (Freeman, 1975). Later research suggests that oscillations in the entorhinal cortex occur at around 5-7 Hz corresponding to the theta rhythm associated with the hippocampus which is also strongly implicated in memory (Alonso & Llinas, 1989; Llinas, 1991). How the theta rhythm associates with the cortical rhythms which underlie the oscillatory activity implicated in visual binding both in neurophysiology and the conceptual model is not clear. The only concrete common factor is that recordings of responses with an oscillatory nature have been made. It is possible to hypothesise that as there are cortical cells that oscillate in a wide frequency band between 10-45 Hz as well as those oscillating in the narrow 35-50 Hz band (Llinas et al, 1991) there are some cortical cells that correspond to the theta rhythm. This could enable the permeation of memory via wide band oscillations into cortical processing, this however is purely speculative. Engel et al emphasise the danger of implying too much from oscillatory activity saying that the presence of such activity does not prove the temporal coding hypothesis in anyway (Engel et al, 1992). This also applies in attempting to link cortical rhythms with theta rhythm in explanations of cortical processing in the human brain.

In relation to the conceptual model then, it is possible to posit these kinds

of mechanisms in relation to artificial visions systems. Concerning the use of the model to understand human visual perception the evidence is by no means conclusive. The role of synchronous oscillatory activity in binding memory in the form of past experience with sensory input for identification purposes is speculative. However by implicating such a mechanism in human visual perception should stimulate interest in pursuing its existence and relevance further.

## CONCLUSIONS.

This chapter has provided additional support for both the model and the theory outlined in chapter 4. The evidence presented is in no way conclusive but aims to substantiate some of the claims that were made in relation to the theory and model but were not presented previously in the thesis. It seems that it is reasonable to suggest that visual perception is not a single process but a combination of both bottom-up and top-down processing where the retinal image provides a hypothesis that is labelled according to information from memory and poly sensory processing. Further, oscillations as an underlying mechanism provide a solution to the integration of information across parallel processing modules.

# **CHAPTER 6**

## **SUMMARY:**

### **A DISCUSSION OF THE ACHIEVEMENTS OF THE CONCEPTUAL MODEL AND SUGGESTIONS FOR FURTHER WORK.**

*A brief summary of the work is outlined followed by a discussion in three sections. The first section discusses what has been achieved in the thesis in relation to its original aims. The second section discusses the conceptual model further in relation to perception. The third section attempts to discuss the implications of the thesis. Finally suggestions for future work are outlined including suggestions for simulation work, psychophysical and neurophysiological experimentation and how the conceptual model may be used as a neuropsychological tool.*

## **INTRODUCTION.**

As stated in chapter 1 the aim of the thesis is to develop both a theory and a conceptual model of visual perception, the purpose of such being twofold. Firstly to provide a better understanding of visual perception and secondly to provide insight into new architectures and mechanisms for artificial intelligence. Ultimately the applications of a conceptual model would be its eventual implementation as an artificial vision system which could be used to perform visual tasks as well as provide insight into biological vision. The

development of such a model was to be based on a bottom-up approach, through investigating what we understand about the neurophysiology of vision, and from a top-down approach, through looking at general theories of vision. Through combining both neurophysiological and psychological accounts of vision a strong multi-disciplinary framework would be established with which to investigate the nature of visual perception.

## **HAS THE THESIS ACHIEVED WHAT IT SET OUT TO ACHIEVE?**

### **OVERVIEW.**

In comparison to the complex accounts of both architecture and function in the visual systems of primates and humans the model evolved is extremely simple. This reflects the extremely complicated nature of visual perception as a whole. There are both models of artificial vision systems and of human visual processing that are both complicated and detailed, significantly more so than the model in the thesis. This is true of psychological top-down models and neurophysiological bottom-up models. Such models are only really detailed in relation to certain specific aspects of visual processing or visual perception. The seemingly impossible task that is created by trying to move away from localised models to global models means that simplicity is the only option. This section aims to address the claims made in chapter 1 and discuss the model in relation to these claims.

## A CONCEPTUAL MODEL THAT PROVIDES BETTER UNDERSTANDING OF HUMAN PERCEPTION.

The claim that the development of a conceptual model of human visual perception will lead to a greater understanding of that process can be seen as an emergent aspect of the thesis. The early chapters enabled an assimilation of a broad range of knowledge about the nature of vision that would not necessarily have been a result of a research path restricted to a more "narrow" domain. For example if the aim of the thesis had been to develop a model of feature integration in early visual processing the research would have been restricted to investigations into the structure and function of V1 alone. This may have resulted in the development of an extremely detailed model of feature integration, it may even have led to the implementation of a conceptual model in software, but it would not have shown any real understanding of visual perception as a whole process. The question is what has the model achieved in helping us understand human visual perception. In general terms the conceptual model suggests that it is possible to regard visual perception as being a combination of both data driven and concept driven processing. The model incorporates a top-down, concept driven component with a bottom-up data driven component to explain how visual perception occurs. The implication of a top-down component in visual perception also suggests that vision as a sensory modality does not necessarily occur in isolation. Information from other sensory areas, as well as from memory indicates that the traditional view of sensory processing as being completely modular is not necessarily the case, although at a local



level processing is in fact restricted to purely visual information. more specifically the model proposes a possible mechanism for the integration of information globally via oscillations driven by top-down diffuse feedback connections and also locally via weak connections between populations of cells representing stimuli.

In relation to understanding human perception one question that emerges from the development of the conceptual model is how much can be explained by developing models of perception? Arguably understanding visual perception is best explained by psychological or neurophysiological experimentation, through the testing of hypotheses or the gathering of data rather than through the construction of conceptual models. This may be true to some extent, but what the conceptual model does is provide the foundations for conducting experimental work. So many theories of perception have been based on a subset of experimental evidence that bears no relation to the visual process as whole. Experimentation that is guided by a general idea of how perception works is far more productive than experimentation that is carried out in isolation of the global interpretation of the system. It is envisaged that it will be possible to extend understanding of human perception with the eventual implementation of the model in software when some of the predictions of the model can be tested against experimental evidence.

## A CONCEPTUAL MODEL THAT CAN EASILY BE IMPLEMENTED IN AN ARTIFICIAL SYSTEM.

The second main aim of the thesis was to develop a conceptual model that could be easily implemented in software and would provide new and innovative algorithms for use in artificial vision systems. To a certain degree the development of the model has achieved this status. It would be possible to build the model in software in the structural and function form outlined. As mentioned in chapter 5 there are a number of simulation studies to support various structural and functional aspects of the model. The role of such a model once it has been successfully implemented in software is not so clear.

Arguably an implementation of the model as it stands would not provide a great deal of benefit to the artificial intelligence community. So far, in artificial intelligence research, it has been acknowledged that to build systems capable of intelligent behaviour we need to look more closely at how the brain functions. The failure of expert systems to realistically extend beyond a single domain, caused attention to focus on neural network theory providing artificial intelligence (AI) with systems that could "learn". However both expert systems and neural networks have been extremely successful in the development of artificial vision systems, especially the application of standard neural network techniques to pattern recognition problems. Both expert system and neural networks are based upon very general and very simple characteristics of brain function rather than detailed explanations. This does not mean that performance can be greatly enhanced by developing an AI model based on detailed structural and

functional understanding of the brain. However the progression from expert systems which are far removed from the brain to neural networks that are more closely linked to the brain has demonstrated an improvement in the capabilities of artificial vision systems in that they can now "learn".

One of the important emergent properties of the model in relation to artificial vision systems is that like the brain the model does not separate structure from function. Standard neural network approaches to artificial vision not only separate structure from function but also apply different functions to the same structures (Churchland, 1986; Churchland & Sejnowski, 1992). The model reflects the relationship between structure and function through maintaining separate pathways for object identification and object manipulation. This relationship is further maintained in the modular processing of information within the model which is based upon the brain. As Churchland points out the structural and function relationship in the brain is the key to its ability to make sure that information requiring different amounts of time to encode, process and act upon information emerges as a coherent whole (Churchland, 1986; Churchland & Sejnowski, 1992). For real time artificial vision systems this could provide the key to enabling realistic emergent behaviours.

What the development of the model has done for AI vision systems to some extent mirrors its achievements in relation to understanding human visual perception. Essentially it provides a general understanding of perception in its entirety providing insight into the sorts of structures and functions that might be useful to incorporate in an artificial vision system rather than providing a rigid detailed account of architecture and function that must be

adhered to. This point can be illustrated in the suggested mechanism for the integration of information in the model. The model suggests that the integration of information can be explained by synchronisation of oscillatory activity across and between modules. This is based upon biological understanding, but can be extrapolated to artificial systems quite successfully. This does not mean that the artificial system has to be architecturally isomorphic with the biological system, or that the oscillatory neurons have to be exact replicas of the best biological interpretation on single neuron behaviour. In fact most artificial single neuron models are based on approximations of biological models of the single neuron such as the leaky integrator models (Bressloff & Taylor, 1990) based on the Hodgkin Huxley equations (Hodgkin & Huxley, 1952).

It would seem possible that simple object recognition and pattern discrimination tasks could be carried out by a software implementation of the model. Shape, colour and motion information could be computed and possible simple discrimination tasks could also be demonstrated. However to produce a successful artificial vision system would most likely mean that a large amount of the biological constraints of the system would have to be omitted to fulfil the software requirements. Arguably the system that would emerge would be no more biologically realistic than those neural network models already functioning. It is also highly probable that the implementation of a biological system would not be as successful as existing systems. What the model has provided artificial intelligence with is some pathways which could be explored in more detail.

## **A MULTI-DISCIPLINARY APPROACH.**

### **A UNIFICATION OF TOP-DOWN AND BOTTOM-UP APPROACHES.**

In chapter 1 it was suggested that one of the most distinguishing features of the model was that it would adopt a multi-disciplinary approach to its development. This would include investigations into the neurophysiology of vision, predominantly looking at architectures and function and investigations into the psychology of vision looking at theories of perception, and how perception manifests itself as higher level cognitive processing and ultimately in behaviour. Finally by proposing the eventual implementation of the model understanding of the computational aspects of the model would be necessary. As the implementation of the model was not included in the investigation of the thesis the multi-disciplinary nature of the conceptual model was restricted to neurophysiology and psychology alone. Computational constraints were addressed in terms of the possibility of being able to actually build the model in software, especially in relation to the proposed underlying mechanisms of the model, but in depth understanding was not necessary at this stage in the research. Bearing in mind the problems with taking a multi-disciplinary approach as outlined in chapter 1, has the thesis been able to provide a truly multi-disciplinary model of visual perception?

Essentially the unification of disciplines was achieved in the sense that the conceptual model incorporates a top-down psychological account of perception which was applied to a bottom-up neurophysiological account of the architectures and function of the visual system. So the top-down

component explained what the bottom-up component would do globally, but the functionality of the architecture itself was actually based on bottom-up neurophysiology. This can also be interpreted as maintaining the segregation of disciplines as well as unifying them. The unifying element is the model itself and the segregated interpretation comes from the two disciplines explaining different aspects of the model, psychology the general theory of perception and neurophysiology the foundations for the architectures and their underlying functionality.

#### HOW UNIFICATION MIGHT HELP UNDERSTAND HUMAN VISUAL PROCESSING.

This unification and segregation within the model can be expressed in relation to the proposal that biological vision operates at three levels. In chapter 1 it was proposed that visual perception could be explained as three levels, local processing, local perception and global perception. Local processing correlates with the processing that takes place within each module in the model. The processing of form information in V1, V2 V3, and V3a, the processing of colour information in V4 and motion information in V5. Local processing has its foundations in the structural and functional descriptions offered by neurophysiology. Local perception correlates with the level in the model where a representation has been formed that relates to a more complete description, the level at which information is combined to form a hypothesis. The distributed pattern of activity across the modules could then be thought of as being a local perception. At this level the visual image is restricted to a purely visual description with no top-down input

from poly sensory information or from memory at all. It is the result of labelling and the checking of the hypothesis that equates to the third level of global perception. The combination of sensory processing from other areas and experiences from memory mean that visual perception is not restricted to the traditional approach in psychology and neurophysiology to compartmentalise behaviour. At the level of local processing and local perception it is possible to modularise, not only limiting vision to the result of processing in a particular cortical area, but also to distinct modules within that area.

This existence of levels could also provide a solution to how higher level cognitive states as defined by psychology might be reduced to low level neuronal behaviours as defined by neurophysiology. It would be possible to infer that high level cognitive states equate to local and global perception whereas low level neuronal behaviour equates to local processing. Having three levels allows for a gradual transition from low level processing to high level processing, with the middle level of local perception comprising both a level of description which is based on low level neuronal behaviours and yet one which aspires to a higher level of description. The construction of a hypothesis begins with low level neuronal behaviours. Within the model artificial neurons extract information about the visual image in parallel within two hierarchical streams. The actual hypothesis is a result of this processing and is represented in IT by cells that correspond to particular patterns of activity. It is at this level that the distributed representation becomes unified as a hypothesis. However it is only through the comparison to memory and the incorporation of poly sensory processing

that local perception becomes a global perception. Therefore the level of local perception provides a level at which higher level processing emerges from low level processing. It must be noted that throughout the whole model, at all levels, low level neuronal behaviour drives high level representations.

## **HOW THE CONCEPTUAL MODEL IS DIFFERENT FROM EXISTING MODELS.**

In very general terms the conceptual model developed is different to existing models in that it attempts to add more detail both to broad psychological theories of perception and to particular aspects of more detailed neurophysiological accounts of perception. This has significant implications in relation to furthering our understanding of human perception and the development of artificial vision systems. By making explicit possible functional structures underlying broad psychological accounts of vision it is possible to see how behavioural explanations of perception can be linked to neurophysiological descriptions of structure and function. Although the addition of a detailed functional description to existing neurophysiological accounts of vision is speculative in terms of its foundations in experimental neurophysiology, it provides a basis for future investigation. Finally by making explicit the architectures and mechanisms underlying both neurophysiological and psychological accounts of perception it is possible to facilitate the development of artificial computer models of visual processing.



## NEUROPHYSIOLOGICAL MODELS.

The functional structure of the conceptual model was based on parallel models of processing within the biological visual system outlined in chapter 3. In particular the model was developed from Zeki's four stream theory of visual processing (Zeki, 1978a; 1980; 1983; 1988; 1992). However the model differs from that proposed by Zeki in several ways.

Primarily the conceptual model extends Zeki's model to incorporate psychological accounts of visual processing. Zeki's model is purely an anatomical account of vision, he does not attempt to link his model to general psychological theories as discussed in chapter 3 such as Gregory's (Gregory, 1972, 1973). This means that Zeki's work can tell us a lot about structure and function, and can be used to question the nature of specific neurophysiological dysfunction, however it does not tell us a great deal about perception as a whole process. It relates more to understanding visual processing within a limited set of cortical structures. In reverse general psychological theories of perception do not provide detailed insight into the architectures and mechanisms that underlie vision. They attempt to explain and understand the so called higher level thought processes both conscious and unconscious that are a consequence of visual processing.

In particular the conceptual model incorporates the theoretical assumption that vision can be explained as both a top down and bottom up process. This is embodied in the structural and functional account of perception by the addition of a description of how memory and polysensory processing might effect visual input, as well as the role of the dLGN in the visual process.

This aspect of visual processing is not addressed by other neurophysiological accounts of vision such as Zeki's.

The importance of the role of the dLGN in the conceptual model is one of the aspects that enable it to be distinguished from other models. Within the model the dLGN is described as playing an active role in the visual process, rather than just relaying information between the retina and the cortex.

Most of the parallel models of visual processing discussed in chapter 3 refer to the dLGN mainly to add support to the segregation of function within the visual system. The classification of X and Y ganglions upheld in the dLGN, as discussed in chapter 2, is interpreted as being evidence for separate streams of processing in the cortex. There are other theories which suggest that the thalamus, and therefore dLGN, plays an active role in sensory processing.

As discussed briefly in chapters 2, 4, and 5 Mumford proposes a model that suggests that the thalamus acts as a blackboard which holds the most recent results of sensory processing which can then act as a reference point for the integration of information throughout the cortex (Mumford, 1991; 1992).

The conceptual model that has been outlined in chapter four makes the role of the dLGN more explicit in that rather than just holding a representation of the most recent results of visual processing it actively engages in association with both the retina and with V1 to enable the confirmation of labelled hypotheses propagated back throughout the visual system. Sillito et al showed experimentally that thalamic relay cells associated with a given stimulus will fire in synchrony as a result of feedback from V1 only if they are also depolarised by their retinal input. They interpret the role of this

cortical feedback as being a mechanism for testing whether or not the cortical representation matches the retinal image. If a match is made through synchronous firing then connectivity is strengthened. They go on to say that these findings are consistent with the proposal that binding in the visual cortex occurs as a result of synchronous oscillatory activity (Sillito et al, 1994).

Within the model the dLGN behaves in the same way as described above. Synchronous oscillatory discharges within the dLGN will result from cells being stimulated simultaneously via feedforward input from the retina and via feedback input from V1. As a consequence of this the connections involved in the relay of this information between V1 and the dLGN will then be strengthened. The synchronous activity generated in the dLGN implies that the LGN is no longer merely relaying information, but carrying an evaluation of both top down and bottom up information.

As mentioned in chapter 3, Zeki's model of visual processing suggests that there are four streams of information flowing in a feedforward way from the retina to the cortex. Each stream deals with the processing of a particular aspect of the visual image such as form, colour, and motion. Zeki suggests that information processed in distinct streams is then integrated by the diffuse feedback connections between cortical areas. These diffuse feedback connections are "non specific", they are not segregated into pathways in the same way that the feedforward connections are. Although Zeki proposes a functional structure for the integration of information, neuronal mechanisms are not made explicit. The conceptual model suggests a possible neuronal mechanism for the integration of information that is processed in separate

sites, that mechanism being synchronous oscillations. This mechanism is speculative, however as discussed in chapters 3 and 5 there is both neurophysiological evidence coupled with simulation studies that have shown that synchronous oscillations are capable of such behaviours. The conceptual model brings together recent neurophysiological evidence which enables existing models such as Zeki's to be extended to fit with a general theory of visual perception. Existing neurophysiological accounts of vision such as Zeki's are incomplete in relation to the proposed theory of perception outlined in chapter 4. In turn existing models of vision based upon more detailed analysis of both structure and function are still only "simplifying brain models". The more detail that is incorporated the smaller the scope of the model. Wehmeier et al have developed a detailed model of vision, but to enable this to be achieved they are only able to model the retina, dLGN and V1 (Wehmeier et al, 1989). This means at best they are modelling a subset of visual processing rather than attempting to develop a model of visual processing. They do not include explanations of the role of memory and poly sensory processing, and they do not include recent advances in our understanding of the possible role of the dLGN as the conceptual model developed here does.

By incorporating recent neurophysiological research into particular structures forming the primary visual pathway it is possible to provide structural and functional support for the general theory proposed in chapter 4. Further to this it has been possible to develop a conceptual model that goes beyond the scope of existing models.

## PSYCHOLOGICAL MODELS.

As already mentioned in the previous section the main distinction between the model proposed here and existing psychological theories and models of perception as a complete process, is that the present account provides detailed insight into the possible structural and functional mechanisms of vision. Through linking a theory based in psychology with neurophysiological description, detailed interpretation of perception is possible. The conceptual model is based around a general theory of visual perception drawn from existing mainstream global theories. As discussed in chapter 3 there are essentially two major views in the psychological literature regarding global theories of vision.

Firstly there are those theories that claim that vision is purely data driven, the visual image is interpreted with no reference to stored information and those that claim vision is concept driven, using stored knowledge about the world to interpret the visual image. Neither of these views alone fits with evidence from neurophysiology into the nature of vision. For example as mentioned in chapter 2 there are as many feedback connections in the brain as feedforward. If vision is purely data driven what is the role of these extensive feedback connections from higher level processing areas to lower level processing areas? Surely it implies that vision is concept driven as well as data driven, relying on stored knowledge to help interpret the visual image. Feedback connections allow the results of processing in higher levels of the visual system and other relevant further processing areas of the brain, to be propagated back through the system to aid interpretation of the visual image.

The model differs from existing data driven theories such as Marr's (Marr, 1982) by extending the general view of vision to include a concept driven element to account for the massive feedback connections in the brain, therefore avoiding one of the major criticisms of Marr's account of object recognition that essentially his model is biologically implausible and hence sheds no new light on human perception (Werner et al, 1993). Further the model is different from Marr's in respect of the relationship between architecture and function. Marr's account of vision identifies the computational level as being the most important in describing visual processing. Architecture and function can be treated as being completely separate. Within the conceptual model outlined in chapter 4, architecture and function are inextricably linked.

The theory on which the conceptual model is based attempts to unify existing theories such as Marr's with concept driven theories such as that proposed by Gregory to fit with neurophysiological data. The conceptual model goes beyond merely unifying data driven and concept driven accounts of vision, it attempts to make explicit the nature of the concept driven input. As discussed in chapter 3, Gregory's theory refers to the concept driven element of perception as "top down knowledge". However, exactly what form this knowledge takes in terms of structure and function in the brain is not made clear. The conceptual model attempts to clarify this by defining "knowledge" as memory, genetically predetermined information and polysensory information. Further to this, embodying the unified theory of perception in a conceptual model enables specific structures and functionality underlying the redefined concept "knowledge" to be made

explicit.

## **SPECIFIC IMPLICATIONS OF THE CONCEPTUAL MODEL**

### **LIMITATIONS OF OSCILLATIONS FOR BINDING.**

Within the model it is proposed that synchronous oscillations provide a possible mechanism for "binding" both within modules and between modules. As discussed in chapters 1, 3 and 5 the proposal for binding through oscillations is relatively new in terms of neurophysiological evidence in their support. A great deal of excitement has been generated by their discovery, however very little has been published which discusses the full extent of their possible limitations.

One general criticism of synchronous oscillations as a mechanism for binding is that what the discovery of 40 HZ oscillations demonstrates is not binding but segmentation (Hardcastle, 1994). Hardcastle takes a classic view of vision similar to that adopted in the development of the conceptual model, and identifies visual perception comprising of three stages. Firstly segmentation, secondly binding and finally association. Segmentation involves early visual processing, where the visual image is segmented into "simple cohesive features". From segmentation higher level visual processing "binds" image segments together, incorporating some top down knowledge to form an object hypothesis. Finally memories and associations enable objects to be correctly identified. Hardcastle argues that the experimental work of neurophysiologists such as Gray and Singer relates only to the first stage of this description of visual processing. Therefore

synchronous oscillation does not explain binding, rather it can be seen as a possible mechanism underlying the segmentation of the visual image in early visual processing.

Hardcastle suggests that the work of Gray and Singer demonstrates segmentation of the image in relation to very simple single features such as orientation. Gray and Singers experiments do not show how cells coding for very different information might behave. Cells are segmenting the image in terms of one possible constraint, i.e. orientation, but not binding different sorts of information such as colour information processed in V4 and form information processed in V1. However as mentioned in chapters 3 and 5 and acknowledged by Hardcastle, work by Engel et al has showed that it is possible to record synchrony between areas of the cortex coding for very different information. Engel et al discovered synchrony between area 17 (V1) and the PMLS. Area 17 (V1) has relatively small receptive fields, coding mainly for fine grain spatial resolution and orientation, whereas the PMLS has large receptive fields and is thought to code for motion, but is severely limited in its capabilities to encode orientation information (Engel et al, 1991a; 1991 b; 1992).

Hardcastle goes on to argue that even as an explanation for segmentation the evidence for synchronous oscillations is limited. It is dangerous to imply that a very small number of single cell recordings can be used to infer a general trend in the behaviour of all the neurons in V1. Further, with the current state of neurophysiological research, it is impossible to accurately test the proposal that synchronous oscillations underlie binding.

A second more general criticism of synchronous oscillations as a solution to



the binding problem is that most of the experimental work has been conducted on the cortex of the cat. There is much discussion about the limited evidence for the existence of 40 Hz oscillations in the primate, and therefore the relevance of this work in understanding human visual perception. Arguably it is as tenuous to extrapolate from primates to the human brain as it is any other animal. The foundations for our understanding of visual processing and for that matter the structure and function of the biological brain as a whole is largely based on, or at least originated from animal experimentation. Dismissing the implications of oscillations on these grounds hardly seems justifiable, as it could imply that perhaps we should question the whole of our understanding of the human brain that has originated from such work. Therefore it is important to make clear that the existence of oscillations is limited and requires further investigation.

More recently research has been directed at investigating more extensively oscillation in monkey cortex, the results being contradictory. There is evidence to suggest that there is both oscillatory, and synchronous oscillatory activity present in awake and anaesthetised monkey cortex. Livingstone recorded both single unit and local field potential oscillations in V1 of the monkey, and also demonstrated phase locking behaviour with two light bars (Livingstone, 1991). Krieter and Singer recorded 40-60 Hz oscillations in MT, however the responses were short, less than 300 Ms and not always present (Krieter and Singer, 1992). Work carried out by Young et al was not so supportive and found that there was little 40 Hz oscillatory activity in V1 and MT, and in IT only two of fifty recording sites in a

monkey performing a face discrimination task oscillated in the 40-60 Hz range (Young et al 1992). However as with more positive conclusions drawn from cat experiments, the methodology for measuring this kind of activity is limited, meaning that more research is required and the results treated carefully.

More specifically it has been argued that synchronous oscillations in respect of visual binding are computationally slow. They do not meet the temporal constraints necessary for perceptual integration. Psychophysical research has show that the time taken to perform perceptual grouping, segmentation and figure ground separation requires less that 200 ms. Biederman et al demonstrated that it was possible for subjects to determine the "semantic" relations between objects in a novel visual scene in less that 150ms (Biederman et al, 1982). This would mean that the visual image would have to be segmented, bound and associated within this time. Further, neurophysiological experimentation has shown that response to visual stimuli in recordings from awake monkey is less than 100 ms. In a discrimination task in which different views of faces were presented to the animals the latency between stimulus onset and a response in cells in STPa was as little as 70 ms. (Oram and Perrett, 1994). However it must be noted that a response in a cell does not necessarily mean that recognition has occurred.

Contrary to these criticisms evidence suggests that binding by synchronous oscillatory activity is capable of meeting these time constraints. Tononi et al has achieved figure ground segmentation in simulation in around 100-200 Ms which fits with psychophysical and neurophysiological data.(Tononi et

al, 1992). The original work by Gray and Singer also demonstrated that synchrony was present within a few tens of milliseconds. It could be argued that it might be possible to equate oscillation onset times with cell recording times, but not so successfully with behavioural response to visual stimuli. Certainly simulations have concentrated on replicating responses in populations of artificial neurons rather than in the generation of behaviours. Edelman's group argue that they have successfully used oscillatory networks to drive a behavioural response to visual stimuli producing similar distribution of reaction times over a given set of trials as psychophysical studies (Tononi et al, 1992).

Another more specific criticism of the use of oscillations as a mechanism for perceptual integration is that they are computationally more expensive. This criticism related more to the development of artificial vision systems than the explanation of biological vision. Although oscillations are computationally expensive and the physical simulation time is longer than in standard neural networks (SNN's) the advantages outweigh this disadvantage. The main advantage is the number of representations one can sustain on a single network, facilitated by the fact that you can encode more than one parameter (Skarda & Freeman, 1987; Grossberg, 1988), phase amplitude and frequency, rather than just amplitude in the form of excitation level as in SNN's. Further, Hummel and Biedermann have developed a model that is capable of the temporal binding of visual relations using 44 cells compared to the  $1.985 \times 10^{11}$  cells needed for a standard enumeration solution (Hummel & Biederman, 1990).

The fact that oscillatory artificial neurons are computationally expensive is

not a realistic problem at this stage in the development of their use in visual perception. At this point they should be seen as the basis for proving a methodology rather than for building systems for the real world. Once the methodology has been established it is envisaged that technology in the form of computing power will have caught up. To a limited degree, they are already feasible with today's computational power, and have been used to replicate binding demonstrated by neurophysiological experimentation (for example: Eckhorn et al, 1989; Eckhorn et al, 1990; Engel et al, 1990; Eckhorn & Schanze, 1991; Eckhorn 1991.) and binding in cognition (Shastri & Ajjanagadde, 1993; Tononi et al, 1992). Further, it is possible to reduce the computational power needed to employ oscillatory behaviours in networks. Examples of this can be seen in the case Carpenter and Grossberg's Adaptive Resonance Theory (ART), where the oscillatory behaviour of resonance is represented in a simplified and computationally less demanding form (Carpenter & Grossberg, 1987a & 1987b; Carpenter & Grossberg, 1988; Grossberg, 1988) and in the work of Baldi et al (Baldi et al, 1990), where oscillatory neurons are modelled simply in terms of their frequency and phase relationships.

Grossberg emphasises this point further in a response to Skarda and Freeman's suggestion that chaos is a necessity of biological systems to self-organise (Skarda & Freeman, 1987) Grossberg emphasises the point that although biological data suggests that chaos is evident in the stable self-organisation of sensory recognition codes, notably in olfaction, it is not necessary to build systems that exhibit chaotic activity to achieve the same functional property of the biological system. Grossberg sites ART as an

example of a self-stabilising system that is not dependent on chaos (Grossberg, 1987c). Baldi et al also demonstrate this point, that biological isomorphism is not necessary to achieve the same results in simulation through his use of coupled oscillators which used frequency and phase information to encode information, but do not actually oscillate (Baldi et al, 1990). In summary you can capture the nature of oscillations without the burden of computational expense necessary to simulate oscillatory neurons. Further criticism of the use of synchronous oscillations to explain binding is the limited number of representations that are available per network. In fact, by binding using temporal codes, it is possible to have several representations coexisting in a network. Recent research has shown that it is possible to represent multiple memories in a single neural network of oscillating neurons (Lisman & Idiart, 1995). Lisman and Idiart replicated psychophysical results which showed subjects were capable of storing up to 7 short term memories at any one time. They worked on the principle that different memories were stored as different high frequency (40 Hz) subcycles of a low frequency oscillation. The number of possible representations on any one network is determined by the number of subcycles that fit in any given low frequency cycle.

#### THE RELATIONSHIP BETWEEN SYNCHRONY IN NEURAL FIRING AND SYNCHRONY IN OSCILLATIONS

As discussed in chapter 5 synchronous oscillations result from neurons discharging simultaneously causing oscillations in the 40-60 Hz range.

Temporal firing patterns are closely correlated with this oscillatory activity

(Gray & Singer, 1987; 1989; Eckhorn et al, 1988; Gray et al, 1989; Eckhorn et al, 1989). However, the integration of information solely by synchronous firing patterns is not without its problems. Retinal jitter, and time delays especially in the integration of information between cortical areas would mean that spike trains are randomised. Information arriving say in IT from two parallel streams of information flow would be subject to varying time delays, meaning that information relating to the same object would be represented by firing patterns that were not temporally synchronised. It is possible that oscillatory discharges would enable these problems to be overcome, enabling the integration of information between cortical areas. Within the visual cortex it is thought that cells code for stimulus specific information through a rhythmic firing pattern which correlates closely with oscillatory discharges in the 40-60 Hz frequency band (Gray & Singer, 1989; Eckhorn et al 1989; Engel et al, 1992;). It has been suggested that this oscillatory activity acts as a "carrier signal for a temporal binding mechanism," between cortical areas (Engel et al, 1992).

As discussed in chapter 5 Eckhorn et al acknowledge the relationship between rhythmic firing patterns and synchronous oscillations. They distinguish between stimulus forced and stimulus induced synchronisation. Stimulus forced synchronisation takes the form of averaged visually evoked potentials and are generally not oscillatory. They are involved at all levels of visual processing. Stimulus induced synchronisation's are oscillatory, and facilitate the correlation of irregular discharges. Eckhorn et al go on to suggest that stimulus forced synchronisation's enable crude pre-attentive representations to be formed, whereas stimulus induced synchronisation's

allow for more sophisticated associations to be made involving inter area interaction within the visual cortex and with other cortical processing areas (Eckhorn et al, 1990).

#### EXAMPLES OF THE COMPUTATIONAL ROLES OF SPECIFIC LINKS IN THE MODEL.

In chapter 4 an outline of the nature of the structure and function of the conceptual model is presented. Although simulation of the model was not undertaken, a simple description of visual processing and a possible simplified implementation of the model is outlined (see figures 18 & 19 and corresponding text). It is possible to expand on chapter 4 by looking more specifically at the computational roles played by specific arrows in the model for example the links between the modules representing IT and PARIETAL cortex, and dLGN and V1.

As outlined in chapter 4, processing within the model is carried out in two streams. The "WHAT" pathway processes form information, and is concerned with object identification. The "HOW" pathway processes information which facilitates a response to the visual stimuli. Although there is evidence of integration between these two systems as discussed in chapter 3 it is beyond the scope of the model to include descriptions of this. The model proposed in general, and at this stage in its development is not designed to model specific psychophysical phenomena. However interaction between the two streams is addressed on a more general level. As discussed in chapter 3 interaction in the biological vision system occurs through diffuse feedback connections and very weak specific feedforward

connections. The basis for interaction within the model is via diffuse feedback connections. Such connections occur between all areas, and result in synchronisation of oscillatory discharges. This allows information processed in separate pathways to be integrated via temporal mechanisms. One such point of interaction is made explicit in the model via a reciprocal link between IT and PARIETAL cortex. The purpose of this link is discussed in chapter 4.

Computationally the link consists of diffuse feedback connections between neurons in IT and the PARIETAL cortex. These feedback connections would be not be specific, enabling the integration of information between areas with unrelated receptive field properties. The links between IT and the PARIETAL cortex are associative links. The idea being that if one area is active the other area is also activated by association. It is envisaged that these links would be weak due to the fact they would not share the same receptive fields. This would mean that rather than connecting one to one they would connect groups of neurons in IT to groups of neurons in the PARIETAL cortex. The links would still be strong enough to enable the oscillatory discharges of neurons in IT representing the visual image processed in the "what" pathway, to synchronise with the oscillatory discharges of neurons in the PARIETAL cortex, representing the visual image processed in the "how" pathway. The oscillatory discharges in IT would represent a object hypothesis, whereas the oscillatory discharges in the PARIETAL cortex would represent the spatial layout information relating to that given object hypothesis. The effect of this would be that the oscillatory responses for "what" information would be synchronised with



those for "how" information This link therefore enables object identity to be linked to spatial and motion information which can then be passed to the motor cortex via the "how" pathway to enable the execution of motor responses to the object.

Engel et al confirm that it is possible to record synchronisation of oscillatory responses between different visual processing areas in the cortex that have very different receptive field properties via weak diffuse feedback connections (Engel et al, 1991b), notably between area 17 (V1) and the PMLS. V1 has small receptive fields and is mainly concerned with the processing of detailed form information. The PMLS on the other hand has large receptive fields and is concerned primarily with the analysis of direction of movement, and its capability for processing orientation information is extremely limited.

As mentioned previously in this chapter and in chapters 4 and 5 the relationship between the dLGN and V1 is crucial to the model. Feedforward excitatory connections between the dLGN and V1 serve to relay the retinal image to the cortex for processing. These connections are specific and related to the receptive field properties of neurons in the dLGN and V1. Segmentation in V1 then takes place as described by neurophysiological accounts in chapter 5. As discussed more generally in chapters 3, 4 & 5 and earlier in this chapter feedback connections between V1 and the dLGN perform a different role. They enable an association to be made between the object hypothesis which is arrived at as a product of feedforward processing, and the retinal image. This association enables the labelled object hypothesis to be confirmed.

The mechanism for this association is similar to the mechanism for the integration of information between IT and the PARIETAL region described previously. Feedback connections from V1 to the dLGN will cause neurons in the dLGN to oscillate if they are receiving feedforward input from the RETINA. Neurons in the dLGN which are oscillating will strengthen their feedforward connections to V1 causing a stronger level of excitation in the relevant neurons in V1. This will then be propagated throughout the model affecting all levels. Further to this the feedback connections from V1 to the dLGN will enable the synchronisation of oscillations between the two areas to enable hypothesis confirmation.

## **GENERAL IMPLICATIONS OF THE CONCEPTUAL MODEL**

### **PSYCHOLOGY**

One of the main implications in relation to psychology is the attempt in the conceptual model to link neurophysiological accounts of vision with psychological ones. As mentioned in previous chapters the rationale for this is that through the development of a conceptual model insight into the relationship between visual processing and visual perception will become clear. This is seen by researchers as being fundamental to advancing our understanding of vision:

*"...we know very little about how visual processing leads to perception! A plethora of processing models have been proposed, yet few encompass or account adequately for more than a small fraction of the available neurobiological data."*

(Oram & Perrett, 1994).

In general terms the model represents a theory of perception, which has evolved from combining mainstream psychological theories, and which has been grounded in neurophysiology. This has enabled one of the aims of the thesis, restated in the quotation above to be addressed. A general claim that perception is a combination of data driven and concept driven processing has been supported from neurophysiological evidence, specifically the role of diffuse feedback connections enabling memory and poly sensory information to be integrated into data driven visual processing. The suggestion of the role of the dLGN as a mechanism for confirming a labelled hypothesis of the visual image, again supported by neurophysiological experimentation, adds further support. This then enables psychological theories such as those of Marr and Gregory to be re-evaluated in the light of these suggestions.

It is possible that refinement of the theory and model through further experimentation and simulation may advance understanding of neuropsychological deficits and advance our understanding of the brain.

Through the future development of large scale biologically plausible simulations of the visual system it will be possible to investigate the functional properties of the visual system. This approach has already had a

major impact on our understanding of the brain. By building biologically realistic models it is possible to investigate the dynamical interactions of the nervous system.

More specifically, through making explicit a possible relationship between psychological phenomena i.e. binding and integration, and a neurophysiological solution i.e. synchronisation of oscillatory discharges, the unification of disciplines has been achieved. This unification allows understanding from both disciplines to advance a general understanding of perception. Arguably psychological theory alone has been incapable of providing an adequate solution to the binding problem.

Other researchers have also used the same neurophysiological mechanisms to successfully develop a model capable of dynamic binding in reasoning (Shastri & Ajjanagadde, 1993), attempting to link so called higher level cognitive functions with low level neurophysiological description. Although their work met with much criticism one of the responses to their paper was that they were making a brave attempt to unite psychology and neurophysiology rather than merely replacing psychological accounts of reasoning with neurobiological ones (Ohlsson, 1993). The conceptual model has also attempted to achieve the same results for vision.

Although there is much criticism of oscillation as a mechanism for integration and specifically binding it seems to be a candidate for serious consideration. The emphasis on neurophysiological mechanisms as a possible solution to the binding problem has lead to the development of research interest that attempts to span several disciplines. One of the biggest problems to overcome is the fear of multi-disciplinary research

evident in both the psychological and neurophysiological literature. The conceptual model has endeavoured to provide a framework which spans disciplines, so providing proof that multi-disciplinary research is possible. More specifically, through the development of both a theory and conceptual model that is firmly grounded in both neurophysiology and psychology, it is hoped that the pitfalls of complex computational models such as that developed by Marr can be avoided. Marr's theory of vision, which dominated psychological accounts of vision for so long relied on biologically implausible computational requirements (Werner, 1993). This meant that although it provided the foundations for the development of successful artificial vision systems it did not provide accurate insight into the nature of human visual perception. As mentioned previously in this chapter, Marr's theory of vision suggests that architecture and function can be seen to exist as separate entities. Through taking into account neurophysiological accounts of visual processing it becomes obvious that structure and function are inextricably linked. Therefore it is possible, through multi-disciplinary research, to advance our understanding of perception.

## ARTIFICIAL INTELLIGENCE

The use of oscillations as mechanisms for binding and integration have far reaching implications for AI. In general, input from psychology and neurophysiology provides design constraints which the AI researcher can utilise (Singer & Donoghue, 1988). The design constraints evident from the development of the conceptual model can be considered to be:

- The use of stored information in the analysis of input.
- Simultaneous feedforward processing of information in separate channels
- The use of both spatial and temporal information in the encoding of information rather than simply using spatial encoding as in SNN's.

Specifically the third constraint is of significance in the development of artificial vision systems. As mentioned in chapter 1 the binding problem is a central issue in the development of artificial models of visual processing as well as in biological vision. Solutions to the binding problem within AI generally involves the enumeration of every possible combination of object, and provide a unit or set of units to represent each combination in a SNN (Hummel & Biederman,1990). For example in the recognition of simple one dimensional coloured objects in the visual scene each possible combination of shape, colour and position is calculated and represented by a possible set of outputs in a neural network. This means that in the analysis of complex visual scenes the number of possible enumerations is beyond the capabilities of most networks. Further, more complete information about every possible enumeration has to be built into the network before it is able to process data. If the network is presented with a combination that it has not been previously enumerated then recognition will fail.

Another problem for such a network is highlighted by Hummel and Biedermann as what they refer to as the "scrambling problem" (Hummel &

Biedermann, 1990). This is where a SNN recognises an incorrect combination of features, which fulfils the enumeration criteria but is visually incorrect. For example the right combinations of edges are present say for a square, but in the wrong combination giving rise to a collection of lines that are not identifiable as a square. Hummel and Biederman demonstrate that these problems can be overcome using "dynamic bindings" in the form of synchrony in networks of oscillating artificial neurons. Within a series of linked networks the spatial relations, location and viewpoint of a given object can all be represented separately and bound together through synchronous oscillations facilitated by "fast enabling links" which do not affect excitatory and inhibitory processing within each network. Further it is possible to represent more than one object at a given time using this technique.

Other successful implementations of artificial systems capable of binding information using temporal coding have been mentioned previously and include Skarda and Freeman's model of olfaction and Shastri and Ajjanagadde's model of systematic reasoning. Skarda and Freeman's model is essentially a neurophysiological model, and does not involve accounts of the cognitive aspects of olfaction. Shastri and Ajjanagadde's model related more closely to the approach taken here, concerned with linking higher level cognitive processes to lower level neuronal behaviours.

So far the impact of temporal coding in the form of oscillations has not had a major impact on mainstream AI. Implementations have mainly centred around the development of biologically realistic models by neurophysiologists for the investigation of brain dynamics. However, as

mentioned earlier, Edelman's group have concentrated on the development of biologically realistic models of brain function that drive robots (Tononi et al, 1992).

## **SUGGESTIONS FOR FUTURE WORK.**

### **OVERVIEW.**

It was not intended that the conceptual model and theory would be in anyway a definitive account of visual perception. What was intended by the thesis was the investigation of the nature of visual perception in relation to its architectures and mechanisms. On reflection it seems that one of the main outcomes of this work is that its contribution to understanding visual perception is to provide the foundations for future research. One of the main motivations behind the thesis was that it is essential to have an understanding of perception as a coherent whole before the construction of detailed software models. Bearing this in mind it is also possible to refine that representation of perception through the implementation of a general model in terms of more specific applications.

### **SIMULATE THE MODEL.**

The next stage of the research program would be to attempt a simulation of the model as it stands, the aim being to decide whether it is possible to investigate perception as a complete process or whether it would be more productive to concentrate investigation on a particular aspect of visual perception such as object recognition or visual attention. Through



simulation it is envisaged that possible extensions to the model would become apparent. The architectures and function described by the model may be insufficient to allow for it to be built in software. This may be the case both in relation to alterations that might be necessary due to constraints enforced by the available simulation tools, and also constraints imposed by the neurophysiological and psychological descriptions underlying the conceptual model. This would be the case for the application of the simulation to both the development of artificial vision systems and understanding human visual perception.

Simulation can be regarded as a tool to advance our understanding of visual perception. Arguably in relation to the general theory forming the foundations of the conceptual model, simulation would not provide anymore insight than the conceptual model outlined in the thesis. All simulation would achieve would be a different way of representing what had already been stated. It could almost be likened to taking a written description of the model and then re writing it in diagrammatic form.

In attempting to understand more specific aspects of the conceptual model, simulation may provide a useful tool. For example investigating whether it is possible that synchronous oscillatory activity underlies feature integration within modules, or whether diffuse feedback resulting in synchronous oscillatory activity could explain the integration of information between modules. If the model was to be used in such a way to understand the mechanisms of human visual perception then it would have to correspond to experimental research taking place in neurophysiology. For example the generalisation of neurophysiological findings about neuronal

mechanisms in V1 to explain how integration across modules might occur is speculation. So all the model would achieve is a mechanistic demonstration of a theory of functionality. What the model would do however, is to provide direction for neurophysiological investigations into the possibility of this artificial behaviour existing in the biological brain.

It is possible therefore to view the possible simulation of the conceptual model as either detailed models of modules, investigating the mechanisms underlying perception, or an extremely simplified simulation of the model as a whole. The former would not explain perception, rather visual processing, the latter on the other hand could be expected to explain visual perception, but the fact that it was an extremely simplified model would mean that it would still only provide a generalised interpretation of perception. It is possible that implementation of either the conceptual model as a whole or specific aspects of the model could be applied to help us understand and explain both perceptual deficits and visual behaviour generally. This would obviously be most successful if carried out in parallel with pure experimental research.

Attempted simulation of the model has far greater significance in the development of artificial vision systems. Continued research would demand that the model was implemented in software and initial tests were carried out to see if it was possible in its simple form to demonstrate its ability to perform in the way proposed by the conceptual model. At this point there would be a need for a great deal of work to be conducted into the computational constraints of the model and ultimately a lot of the biological realism would have to be omitted. A simulation for artificial vision systems

would be based upon the conceptual model, but the resultant behaviour of the system would have to be more important than its biological realism. Initially it would be necessary to decide what kind of system was to be built based upon the model. As with its application to understanding human vision there are two possible options. Firstly to try to implement the conceptual model as a complete system capable of human like perception. The possible uses of such a system would include real time robotics. Secondly it is possible to concentrate on a particular aspects of the model to provided solutions for specific vision problems. For example the development of the feature linking mechanisms proposed by the model for use in pattern recognition processes.

The next stage would be to select a possible neuronal model which would behave in the way proposed by the conceptual model. Most neuronal models that produce oscillatory behaviour are based upon experimental data from single cell recordings, therefore it would seem that biological realism could be applied even to artificial vision systems. However the simplification of the behaviours of such models to allow them to be described mathematically means that they can at best be considered approximations of neuronal behaviour. The selection and formulation of the neuronal model is in itself a major research program. This then becomes even more complex when networks of neurons are constructed. The connection of artificial neurons in a network, be it a biologically based architecture or a completely arbitrary one does not guarantee behaviours proposed by experimental data or conceptual models. It may be the case that architecture, function and mechanisms as proposed by the model may not extrapolate to the

development of artificial vision systems. The conceptual model may provide ideas that could then be developed further, but may not rely on the actual constraints of the model.

In sum then, The conceptual model is in fact a hypothesis for a model, it could be argued that simulation at this point would involve a great deal of hard work based on pure speculation. It would be feasible to attempt simulation of certain aspects of the model, but this would not provide support for the theory and conceptual model as a whole.

#### POSSIBLE EXTENSIONS TO THE MODEL.

The development of any model is an iterative process. Once a model has been proposed it is possible to adapt and change that model to fit with new data and to extend it to incorporate a wider range of behaviours. It was not intended that the conceptual model would account for all visual processes. It does not for example include secondary visual pathways, visual attention processes and detailed explanations of connectivity between visual modules. Therefore further work on the model must include a more comprehensive and more detailed account of both architecture and function. What is needed is to look at particular visual behaviours in relation to the model and to extend it incorporate these behaviours, for example how the model explain the recognition of faces, or how action responses are generated from visual stimuli. To a certain degree extensions to the model will be defined by the applications of the model. If for example the model is to be use to help understand blind sight, then secondary visual pathways must be incorporated. In the same way investigation into other visual behaviours

will mean that change to the model will be necessary.

One of the key areas for further work is the development of a more detailed understanding of the more speculative aspects of the model. This would include investigations into the way in which top-down memory was integrated into the system as an aid to hypothesis labelling, and the role of the dLGN in hypothesis confirmation. It would be expected that simulation, especially for the purposes of modelling human visual perception would demand that extensions such as the incorporation of secondary visual pathways were made. The model claims to provide a global account of visual perception. Arguably without incorporating an explanation of attention processes, and by omitting certain known structures relating to visual perception this claim is not entirely valid. However the model provides a point from which perception as a whole process can be investigated. Importantly the development of the model so far has shown that visual perception does not occur in isolation, it includes input from processing across the whole brain. Therefore not including descriptions of certain visual processing functions and particular detailed structures is really a minor omission in relation to the rest of the brain.

The theory underlying the model is not conclusive. Is therefore envisaged that another possible extension to the model would involve conducting both psychological and neurophysiological experiments to investigate the theory in more detail. This may result in changes being made to the model in order for it to be used in attempts at understanding human visual processing.

However as descriptions of structure and function can be separated from the theoretical foundations, explanations of some of the more detailed aspects

of visual processing would not be affected by such modifications. Any changes that were made as a result of such experimentation would not necessarily affect the model in terms of its application to the development of artificial vision systems unless such systems claimed to be biologically plausible.

## **CONCLUSIONS.**

A multi-disciplinary approach to the investigation of visual perception has been successfully applied and the conceptual model developed has benefited from this approach. A better understanding or rather interpretation of human visual perception as well as new possibilities for artificial vision systems has been achieved. Finally has been possible to place the conceptual model in relation to future research.

# GLOSSARY

<b>Local Visual Processing</b>	The processing of sensory information localised to those areas of the brain which process only visual information. Essentially these areas can be defined as; V1, V2, V3, V3a, V4 and V5.
<b>Local Perception</b>	The perception of sensory input localised to a given sensory area, namely vision. Any "perception" at this level would be restricted to visual information only and would not include information processed in other sensory modalities or memory.
<b>Global Perception</b>	The perception of sensory input that includes the result of local processing and local perception in many sensory areas as well as input from memory.
<b>Bottom-Up Approaches</b>	An approach to the study of visual perception that begins at the lowest possible level of explanation, and results in the formation of a theory. More explicitly with regard to the study of visual perception, to begin by trying to understand vision at the neuronal level and using successively higher level descriptions eventually to formulate a theory based on these investigations which also explains high level behavioural accounts of visual perception.
<b>Top-Down Approaches</b>	An approach to the study of visual perception that takes the opposite approach to bottom-up approaches. Initially the formulation of a theory takes place based on a high level behavioural account of visual perception, which is then supported or refuted by the collection of evidence at lower levels of description, ending at the neuronal level.
<b>Data-Driven</b>	An approach to the explanation of visual perception that maintains all the information needed for interpreting the visual image is already present in the visual scene. No prior or stored knowledge is necessary.
<b>Concept-Driven</b>	An approach to the explanation of visual perception that takes the opposite view of data-driven approaches. Concept-driven approaches to understanding visual perception maintain that stored knowledge in the form of memory is necessary for the successful interpretation of the visual image.

# **REFERENCES**

- ADELSON, E. H. & MOVSHON, J. A. 1982. Phenomenal Coherence of Moving Visual Patterns. *Nature*. Vol. 300, pp. 523-525.
- ADELSON, E. H. & MOVSHON, J. A. 1984. Binocular Disparity and the Computation of Two Dimensional Motion. *Journal of the American Optical Society*. pp. 1266.
- ALBRIGHT, T. D. 1991. Color and the Integration of Motion Signals. *Trends In Neuroscience*. Vol. 14, No. 7, pp. 266-269.
- ALLMAN, J. M. & KAAS, J. H. 1971. A Representation of the Visual Field in the Caudal Third of the Middle Temporal Gyrus of the Owl Monkey (*Agoutis Trivirgatus*). *Brain Research*. Vol. 31, pp. 85-105.
- ALLMAN, J. M. & KAAS, J. H. 1974. The Organisation of the Second Visual Area (VII) in the Owl Monkey: A Second Order Transformation of the Visual Hemifield. *Brain Research*. Vol. 76, pp. 247-265.
- ALONSO, A. & LLINAS, R. R. 1989. Na<sup>+</sup>-Dependent Theta-Like Rhythmicity in Stellate Cells of Entorhinal Cortex Layer II. *Nature*. Vol. 342, pp. 175-177.
- BADDELEY, A. D. & HITCH, G. J. 1974. Working Memory. In G. H. Bower (Ed.). *Recent Advances in Learning and Motivation*. Volume 8. Academic Press, New York.
- BAKER, C. L., HESS, R. & ZIHL, J. 1987. The "Motion-Blind" Human Patient: Peripheral Filters. *The Proceedings of the Physiological Society*. 18-19 September, Cambridge, England. In *The Journal of Physiology*. Vol. 396, pp. 63.
- BAKER, C. L., HESS, R. & ZIHL, J. 1991. Residual Motion Perception in a "Motion-Blind" Patient, Assessed With Limited-Lifetime Random Dot Stimuli. *The Journal of Neuroscience*. Vol. 11, No. 2, pp. 454-461.
- BALDI, P., BUHMANN, J. & MEIR, R. 1990. Computing with Arrays of Coupled Oscillators. *Proceedings of the International Neural Networks Conference, Paris, July 1990*.
- BANKS, M. S. 1980. The Development of Visual Accommodation During Early Infancy. *Child Development*. Vol. 51, pp. 646-666.
- BANKS, M. S. & SALAPATEK, P. 1983. Infant Visual Perception. In M. M. Haith & J. J. Campos (Eds.) *Handbook of Child Psychology*. Vol. 2: *Infancy and Developmental Psychobiology*. Wiley, New York.
- BANKS, M. S. & GINSBURG, A. P. 1985. Infant Visual Preferences: A Review and New Theoretical Treatment. In H. W. Reese (Ed.). *Advances in Child Development and Behaviour*. Vol. 19. Academic press, Orlando, Florida.
- BARRERA, M. E. & MAURER, D. 1981. Recognition of Mother's Photographed Face by the Three Month Old Infant. *Child Development*. Vol. 52, pp. 714-716.



- BARLOW, H. B.** 1972. Single Units and Sensation: A Neuron Doctrine for Perceptual Psychology? *Perception*. Vol. 1, pp. 371-394.
- BASAR, E. & BULLOCK, T. H. (Eds.)** 1992. Induced Rhythms in the Brain. *Brain Dynamics Series*. Birkh -user.
- BAYLIS, C. G., ROLLS, E. T. & LEONARD, C. M.** 1985. Selectivity Between Faces in the Response of a Population of Neurons in the Cortex in the Superior Temporal Sulcus of the Monkey. *Brain Research*. Vol. 342, pp. 91-102.
- BEER, R. D.** 1989. Intelligence as Adaptive Behaviour: An Experiment in Computational Neuroethology. *Academic Press INC. Harcourt Brace Jovanovich, London*.
- BERLUCCHI, G. & SPRAGUE, J. M.** 1981. The Cerebral Cortex in Visual Learning and Memory, and Interhemispheric transfer in the Cat. In F. O. Schmitt, F. G. Worden, G. Adelman & S. G. Dennis (Eds.). *The organization of the Cerebral Cortex: Proceedings of a Neurosciences Research Program Colloquium*. 1981. MIT Press, Cam. MA & London, England.
- BIEDERMAN, I.** 1972. Perceiving Real World Scenes. *Journal of Experimental Psychology*. Vol. 97, pp. 22-27.
- BIEDERMAN, I., GLASS, A. L. & STACY, E. W.** 1973. Searching for Objects in Real World Scenes. *Science*. Vol. 177, pp. 77-80.
- BIEDERMAN, I., MEZZANOTTE, R. J. & RABINOWITZ, J. C.** 1982. Scene Perception: Detecting and Judging Objects Undergoing Relational Violations. *Cognitive Psychology*. Vol. 14, pp. 143-177.
- BEST, J. B.** 1986. Cognitive Psychology. *West Publishing Company, USA*.
- BLANK, D. S., MEEDEN, L. A. & MARSHALL, J. B.** 1992. Exploring the Symbolic/Subsymbolic Continuum: A Case Study of RAMM. In J. Dinsmore (Ed.). *Closing the Gap: Symbolism Vs Connectionism*. Erlbaum.
- BLOOM, F. E. & LAZERSON, A.** 1988. Mind, Brain and Behaviour. 2nd edition. *W. H. Freeman, New York*.
- BOWER, T. G. R.** 1966. The Visual World of Infants. *Scientific American*. Vol. 215, No. 6, pp. 80-92.
- BRESSLOFF, P. C. & TAYLOR, J. G.,** 1990. Discrete Time Models of Noisy Networks. *Proceedings of the International Neural Networks Conference, Paris, July 1990*.
- BRODMANN, K.** 1909. Vergleichende Lokalisationlehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues. *J. A. Barth, Leipzig*.
- BRUCE, V. & GREEN, P. R.** 1990. Visual Perception: Physiology, Psychology and Ecology. 2nd edition. *LEA, London*.

- BUGELSKI, B. R. & ALAMPAY, D. A.** 1961. The Role of Frequency in Developing Perceptual Sets. *Canadian Journal of Psychology*. Vol. 15, pp. 205-211.
- BUGMANN, G. & TAYLOR, J. G.** 1994. A Top-Down Model for Neuronal Synchronisation. *Submitted to Neural Computation*.
- BURKHALTER, A.** 1993. Development of Forward and Feedback Connections Between Areas V1 and V2 of Human Visual Cortex. *Cerebral Cortex*. Vol. 3, pp. 476-487.
- BURNOD, Y.** 1988. An Adaptive Neural Network: The Cerebral Cortex. *Prentice Hall, London*.
- BUSHNELL, I. W. R., SAI, F., & MULLIN, J. T.** 1989. Neonatal Recognition of the Mother's Face. *British Journal of Developmental Psychology*. Vol. 7, pp. 3-15
- CARPENTER, G. A. & GROSSBERG, S.** 1987a. A Massively Parallel Architecture for a Self-Organising Neural Pattern Recognition Machine. *Computer Vision, Graphics and Image Processing*. Vol. 37, pp. 54-115.
- CARPENTER, G. A. & GROSSBERG, S.** 1987b. ART 2: Self-Organisation of Stable Category Recognition Codes for Analog Input Patterns. *Applied Optics*. Dec. 1st, pp. 4919-4930.
- CARPENTER, G. A. & GROSSBERG, S.** 1988. The ART of Adaptive Pattern Recognition by a Self-Organizing Neural Network. *Computer*. March, 1988, pp. 77-88.
- CAMPBELL, F. W. & ROBSON, J. G.** 1968. Application of Fourier Analysis to the Visibility of Gratings. *Journal of Physiology*. Vol. 197, pp. 551-556.
- CAMPION, J., LATTO, R. & SMITH, Y. M.** 1983. Is Blindsight an Effect of Scattered Light, Spared Cortex, and Near-Threshold Vision? *Behavioral and Brain Sciences*. Vol. 6, pp. 423-486.
- CARLSON, N. R.** 1986. Physiology of Behavior. Third Edition. *Allyn and Bacon, London*.
- CARPENTER, G. & GROSSBERG, S.** 1987. A Massively Parallel Architecture for a Self-Organizing Neural Pattern Recognition Machine. *Computer Vision Graphics Image Processing*. Vol. 37, pp. 54-115.
- CHALMERS, D. J.** 1990. Syntactic Transformations on Distributed Representations. *Connection Science*. Vol. 1/2, pp. 53-62.
- CHOMSKY, N.** 1965. Aspects of the Theory of Syntax. *MIT Press, Cambridge MA*.
- CHURCHLAND, P. SMITH.** 1986. Neurophilosophy: Toward a Unified Science of the Mind-Brain. *MIT Press, USA*.
- CHURCHLAND, P. SMITH & SEJNOWSKI, T. J.** 1992. The Computational Brain. *MIT Press, Cam. MA & London, England*.

**COHEN, L. B., DELOACHE, J. S. & STRAUSS, M. S.** 1979. Infant Visual Perception. In *J. Osofsky (Ed.). Handbook of Infant Development*. Wiley, New York.

**COTTERILL, R. M. J. & NIELSEN, C.** 1991. A Model for 40 Hz Oscillations Invokes Inter-Area Interactions. *Theoretical Neuroscience*. Vol. 2, No. 5, pp. 289-292.

**COWEY, A.** 1981. Why Are There So Many Visual Areas? In *F. O. Schmitt, F. G. Worden, G. Adelman & S. G. Dennis (Eds.). The organization of the Cerebral Cortex: Proceedings of a Neurosciences Research Program Colloquium*. 1981. MIT Press, Cam. MA & London, England.

**COWEY, A. & STOERIG, P.** 1990. Reflections on Blindsight. In *A. D. Milner & M. D. Rugg (Eds.). The Neuropsychology of Consciousness*. Academic Press, London.

**CRICK, F.** 1984. The Function of the Thalamic Reticular Complex: The Searchlight Hypothesis. *Proceedings of the National Academy of Science (USA)*. Vol. 81, pp. 4586-4590.

**DENHAM, M. J., PATEL, S., TROUP, L. J. & NORMAN, M.** 1991. Oscillatory Neural Networks and their Application to Sensory-Motor Co-ordination and Control in Adaptive Robots. In *Proceedings of the IEE 2nd International Conference on Artificial Neural Networks*. IEE Publishing, London, pp. 328-332.

**DENHAM, M. J. & TROUP, L. J.** 1992. Coherent Oscillatory Responses as a Feature Linking Mechanism in an Artificial Neural Network Model of Visual Perception. In *Proceedings of NEURO NIMES: The 5th International Conference on Neural Networks and their Applications*, Nimes, France. Ec2, Paris, pp. 505-516.

**DEREGOWSKI, J. B.** 1972. Pictorial Perception and Culture. *Scientific American*. Vol. 227, No. 5, pp. 82-88.

**DESIMONE, R., WESSINGER, M., THOMAS, L. & SCHNEIDER, W.** 1990. Attentional Control of Visual Perception: Cortical and Subcortical Mechanisms. *Cold Spring Harbor Symposium. Quantitative Biology*. Vol. 55, pp. 963-971.

**DE VALOIS R. L., ALBRECHT, D. G. & THORELL, L. G.** 1982. Spatial Frequency Selectivity of Cells in Macaque Visual Cortex. *Vision Research*. Vol. 22, pp. 545-559.

**DE YOE, E. A. & VAN ESSEN, D. C.** 1985. Segregation of Efferent Connections and Receptive Field Properties in Visual Area V2 of the Macaque. *Nature*. Vol. 317, pp. 58-61

**DORIZZI, B. & GRAMMATICOS, B.** 1991. Stimulus-Induced Neural Synchronizations. In *Artificial Neural Networks: Proceedings of the 1991 International Conference on Artificial Neural Networks*, Espoo, Finland, June 1991. T. Kohonen, K. Makisara, O. Simula & J. Kangas (Eds.). Elsevier, Amsterdam.

**DOWLING, J. E.** 1968. Synaptic Organisation of the Frog Retina: An Electron Microscopic Analysis Comparing the Retinas of Frogs and Primates. *Proceedings of the Royal Society of London, B*. Vol. 170, pp. 205-228.

**DOWLING, J. E.** 1992. Neurons and Networks: An Introduction to Neuroscience. *The Belknap Press of Harvard University Press*, Cambridge, MA & London, England.

ECKHORN, R., BAUER, R., JORDAN, W., BROSCHE, M., KRUSE, W., MUNK, M. & REITBOECK, H. J., 1988. Coherent oscillations: A Mechanism of Feature Linking in the Visual Cortex? *Biological Cybernetics*. Vol. 60, pp. 121-130.

ECKHORN, R., REITBOECK, H. J., ARNDT, M. & DICKE, P. 1989. A Neural Network for Feature Linking via Synchronous Activity: Results from Cat Visual Cortex and from Simulations. In *Models of brain function*. R. M. J. Cotterill (Ed.), Cambridge University Press, pp. 225-272.

ECKHORN, R., BAUER, R., REITBOECK, H. J., ARNDT, M. & DICKE, P. 1990. Feature Linking via Synchronisation Among Distributed Assemblies: Simulations of Results from Cat Visual Cortex. *Neural Computation*. Vol. 2, pp. 293-307.

ECKHORN, R., & SCHANZE, T. 1991. Possible Neural Mechanisms of Feature Linking in the Visual System: Stimulus-Locked and Stimulus-Induced Synchronisations. In *Self-Organisation, Emerging properties, and Learning*. A. Babloyantz (Ed.), Plenum Press, New York.

ECKHORN, R. 1991. Stimulus-Specific Synchronisations in the Visual Cortex: Linking of Local Features Into Global Figures? In *Neuronal Cooperativity*. J. Kruger (Ed.), Springer Verlag, Berlin, Heidelberg.

ENGEL, E. 1956. The Role of Content in Binocular Resolution. *American Journal of Psychology*. Vol. 69, pp. 87-91.

ENGEL, A. K., KONIG, P., GRAY, C. M. & SINGER, W., 1990. Stimulus-Dependent Neuronal Oscillations in Cat Visual Cortex: Inter-Columnar Interaction as Determined by Cross-Correlation Analysis. *European journal of Neuroscience*. Vol. 2, pp. 588-606.

ENGEL, A. K., KONIG, P., KRIETER, A. K. & SINGER W., 1991a. Interhemispheric Synchronisation of Oscillatory Neuronal Responses in Cat Visual Cortex. *Science*. Vol. 252, pp. 1177-1179.

ENGEL, A. K., KONIG, P., KRIETER, A. K. & SINGER W., 1991b. Synchronisation of Oscillatory Neuronal Responses Between Striate and Extrastriate Visual Cortical Areas of the Cat. *Proceedings of the National Academy of Science (USA)*. Vol. 88, pp. 6048-6052.

ENGEL, A. K., KONIG, P., KRIETER, A. K., SCHILLEN, T. B. & SINGER W., 1992. Temporal Coding in the Visual Cortex: New Vistas on Integration in the Nervous System. *Trends In Neuroscience*. Vol. 15, No. 6, pp. 218-226.

ENROTH-CUGELL, C. & ROBSON, J. G. 1966. The Contrast Sensitivity of Retinal Ganglions of the Cat. *Journal of Physiology*. Vol. 187, pp. 517-552.

FANTZ, R. L. 1961. The Origin of Form Perception. In W. T. Greenough (Ed.). *The Nature and Nurture of Behaviour: Developmental Psychobiology*. W. H. Freeman, London.

FERSTER, D. & LINDSTROM, S., 1983. An Intracellular Analysis of Geniculo-Cortical Connectivity in Area 17 of the Cat. *Journal of Physiology*. Vol. 342, pp. 181-215.

FREEMAN, W. J. 1975. Mass Action in the Nervous System. *Academic Press*.

- FUSTER, J. M. & JERVEY, J. P.**, 1982. Neuronal Firing in the Inferotemporal Cortex of the Monkey in a Visual Memory Task. *Journal of Neuroscience*. Vol. 2, pp. 361-375.
- GAWNE, T. J., KJAER, T. W. K., HERTZ, J. A. & RICHMOND, B. J.** 1994. What are all Those Neurons in the Visual Cortex Really Doing? *Dynamics of Neural Processing. International Symposium (extended abstracts)*. Washington, DC. June 6-8th 1994, pp. 192-194.
- GIBSON, J. J.** 1950. The Perception of the Visual World. *Houghton Mifflin, Boston, USA*.
- GIBSON, E. J. & WALK, R. D.** 1960. "The Visual Cliff". In *W. T. Greenough (Ed.). The Nature and Nurture of Behaviour: Developmental Psychobiology*. W. H. Freeman, London.
- GILBERT, C.D.** 1992. Horizontal Integration and Cortical Dynamics. *Neuron*. Vol. 9, 1-13, July, pp. 1-13.
- GOODALE, M. A. & MILNER, D. A.** 1992. Separate Pathways for Perception and Action. *Trends in Neuroscience*. Vol. 15, No. 1, pp. 20-25.
- GOULD, S. J.** 1981. The Mismeasure of Man. *Penguin, England*.
- GRAY, C. M. & SINGER, W.**, 1987. Stimulus-Specific Neuronal Oscillations in the Cat Visual Cortex: A Cortical Functional Unit. *Society for Neuroscience Abstract*. Vol. 13, pp. 404-403.
- GRAY, C. M. & SINGER, W.**, 1989. Stimulus-Specific Neuronal Oscillations in Orientation Columns of the Cat Visual Cortex. *Proceedings of the National Academy of Science (USA)*. Vol. 86, pp. 1698-1702.
- GRAY, C. M., KONIG, P., ENGEL, A. K. & SINGER, W.**, 1989. Oscillatory Responses in Cat Visual Cortex Exhibit Inter-Columnar Synchronisation which Reflects Global Stimulus Properties. *Nature*. Vol. 338, pp. 334-337.
- GREGORY, R. L.** 1972. Eye and Brain. *Second edition. World University Library*.
- GREGORY, R. L.** 1973. The Confounded Eye. In *R. L. Gregory & H. E. Gombrich (Eds.). Illusion in Nature and Art*. Duckworth, London.
- GREGORY, R. L.** 1995. Seeing Backwards in Time. *Nature*. Vol. 373, pp. 21-22.
- GROSS, C. G., ROCHA-MIRANDA, C. E. & BENDER, D. B.** 1972. Visual Properties of Neurons in Inferotemporal Cortex of the Macaque. *Journal of Neurophysiology*. Vol. 35, pp. 96-111.
- GROSSBERG, S.** 1988. Nonlinear Neural Networks: Principles, Mechanisms, and Architectures. *Neural Networks*. Vol. 1, pp. 17-61.

- GROSSBERG, S. & SOMERS, D.**, 1991a. Synchronised Oscillations During Co-Operative Feature Linking in Visual Cortex. In *Artificial Neural Networks: Proceedings of the 1991 International Conference on Artificial Neural Networks, Espoo, Finland, June 1991*. T. Kohonen, K. Makisara, O. Simula & J. Kangas (Eds.). Elsevier, Amsterdam.
- GROSSBERG, S. & SOMERS, D.**, 1991b. Synchronised Oscillations During Co-Operative Feature Linking in a Cortical Model of Visual Perception. *Neural Networks*. Vol. 4, pp. 453-466.
- HABER, R. N. & HERSHENSON, M.** 1965. Effects of Repeated Brief Exposures on the Growth of a Percept. *Journal of Experimental Psychology*. Vol. 69, pp. 40-46.
- HAITH, M. M.** 1980. Visual Competence in Early Infancy. In R. Held, H. Liebowitz & R. Teuber (Eds.). *Handbook of Sensory Physiology*. Vol. 8. Springer Verlag, Berlin.
- HARDCASTLE, V. G.** 1994. Psychology's Binding Problem and Possible Neurobiological Solutions. *Journal of Consciousness Studies*. Vol. 1, No. 1, pp. 66-90.
- HARTLINE, H. K. & GRAHAM, C. H.** 1932. Nerve Impulses from Single Cell Receptors in the Eye. *Journal of Cellular and Comparative Physiology*. Vol. 1, pp. 227-295.
- HARTLINE, H. K., WAGNER, H. G. & RATLIFF, F.** 1956. Inhibition in the Eye of the Limulus. *Journal of General Physiology*. Vol. 39, pp. 651-673.
- HINTON, G. E. McCLELLAND, J. L. & RUMELHART, D. E.** 1986. Distributed Representations. In D. E. Rumelhart & J. L. McClelland (Eds.). *Parallel Distributed processing: Explorations in the Microstructure of Cognition*. Volumes 1 & 2. MIT Press, USA
- HIRSCH, H. V. B. & SPINELLI, D. N.** 1971. Modification of the Distribution of Receptive Field Orientation in Cats by Selective Visual Exposure During Development. *Experimental Brain Research*. Vol. 13, pp. 509-527.
- HODGKIN, A. L. & HUXLEY, A. F.** 1952. A Quantitative Description of Membrane Current and its Application to Conduction and Excitation in Nerve. *Journal of Physiology*. Vol. 117, pp. 500-544.
- HUBEL, D. H. & WIESEL, T. N.** 1962. Receptive Fields, Binocular Interactions and Functional Architecture in the Cat's Visual Cortex. *Journal of Physiology*. Vol. 160, pp. 106-154.
- HUBEL, D. H. & WIESEL, T. N.** 1970. The Period of Susceptibility to the Physiological Effects of Unilateral Eye Closure in Kittens. *Journal of Physiology*. Vol. 206, pp. 419-436.
- HUBEL, D. H. & WIESEL, T. N.** 1977. Functional Architecture of the Macaque Monkey Visual Cortex. *Proceedings of the Royal Society of London. Series B*. Vol. 198, pp. 1-59.
- HUBEL, D. H., WIESEL, T. N. & LE VAY, S.** 1977. Plasticity of Ocular Dominance Columns in Monkey Striate Cortex. *Philosophical Transactions of the Royal Society, B*. Vol. 278, pp. 377-409.

- HUBEL, D. H. & LIVINGSTONE, M. S.** 1983. Blobs and Color Vision. *Canadian Journal of Physiology and Pharmacology*. Vol. 61, pp. 1433-1441.
- HUBEL, D. H. & LIVINGSTONE, M. S.** 1985. Complex-Unoriented Cells in a Subregion of Primate Area 18. *Nature*. Vol. 315, pp. 325-327.
- HUMMEL, J. E. & BIEDERMANN, I.** 1990. Dynamic Binding: A Basis for the Representation of Shape by Neural Networks. In *Proceedings of the 12th Annual Conference of the Cognitive Science Society*. Erlbaum, Hillsdale, N.J. USA, pp. 614-621.
- HUMMEL, J. E. & BIEDERMANN, I.** 1992. Dynamic Binding in a Neural Network for Shape Recognition. *Psychological Review*. Vol. 99, No. 3, pp 480-517.
- KOCH, C.** 1990. Biophysics of Computation: Toward the Mechanisms Underlying Information Processing in Single Neurons. In *E. L. Schwartz (Ed.). Computational Neuroscience*. Bradford Books, MIT Press, CA MA.
- KRAUSKOPF, J. & FARELL, B.** 1990. Influence of Colour on the Perception of Coherent Motion. *Nature*. No. 348, pp. 328-331.
- KREITER, A. K. & SINGER, W.** 1992. Oscillatory Neuronal Responses in the Visual Cortex of the Awake Macaque Monkey. *European Journal of Neuroscience*. Vol. 4, pp. 369-375.
- KUFFLER, S. W.** 1953. Discharge Patterns and Functional Organisation of the Mammalian Retina. *Journal of Neurophysiology*. Vol. 16, pp. 37-68.
- KURRER, C., NIESWAND, B. & SCHULTEN, K.** 1991. A Model for Synchronous Activity in the Visual Cortex. In *Self-Organisation, Emerging properties, and Learning*. A. Babloyantz (Ed.), Plenum Press, New York.
- KURRER, C., NIESWAND, B. & SCHULTEN, K.** 1992. Dynamics of Synchronous Neural Activity in the Visual Cortex. In *Artificial Neural Networks 2: Proceedings of the 1992 International Conference on Artificial Neural Networks, (ICANN 1992) Brighton, England, September 1992*. I Aleksander & J. G. Taylor (Eds.). Elsevier, Amsterdam.
- LA BERGE, D., CARTER, M. & BROWN, V.** 1992. A Network Simulation of Thalamic Circuit Operations in Selective Attention. *Neural Computation*. Vol. 4, pp. 318-331.
- LEEPER, R.** 1935. A Study of a Neglected Portion of the Field of Learning: The development of Sensory Organisation. *Journal of Genetic Psychology*. Vol. 46, pp. 41-75.
- LE VAY, S., WIESEL, T. N. & HUBEL, D. H.** 1981. The Postnatal Development and Plasticity of Ocular-Dominance Columns in the Monkey. In *F. O. Schmitt, F. G. Worden, G. Adelman & S. G. Dennis (Eds.). The organization of the Cerebral Cortex: Proceedings of a Neurosciences Research Program Colloquium*. 1981. MIT Press, Cam. MA & London, England.
- LINDSAY, P. H. & NORMAN, D. A.** 1972. Human Information Processing. *Academic Press, New York*.

- LIVINGSTON, M. S. & HUBEL, D. H.** 1982. Thalamic Inputs to Cytochrome Oxidase-Rich. Regions in the Monkey Visual Cortex. *Proceedings of the National Academy of Sciences (USA)*. Vol. 79, pp. 6098-6101.
- LIVINGSTON, M. S. & HUBEL, D. H.** 1984. Anatomy and Physiology of a Color System in the Primate Visual Cortex. *Journal of Neuroscience*. Vol. 4, pp. 309-356.
- LIVINGSTONE, M. S.** 1991. Visually Evoked Oscillations in Monkey Striate Cortex. *Society for Neuroscience Abstracts*. Vol. 17, 73.3.
- LLINAS, R. R., GRACE, A. A. & YAROM, Y.** 1991. In Vitro Neurons in Mammalian Cortical Layer 4 Exhibit Intrinsic Oscillatory Activity in the 10- to 50- Hz Frequency Range. *Proceedings of the National Academy of Sciences (USA)*. Vol. 88, pp. 897-901.
- MAHOWALD, M. A. & MEAD, C.** 1991. The Silicon Retina. *Scientific American*. May 1991, pp. 40-46.
- MARR, D.** 1976. Early Processing of Visual Information. *Philosophical Transactions of the Royal Society, B*. Vol. 275, pp. 483-524.
- MARR, D.** 1982. Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. *W. H. Freeman, San Francisco, (USA)*
- MARR, D. & HILDRETH, E.** 1980. Theory of Edge Detection. *Proceedings of the Royal Society, London B*. Vol. 207, pp. 187-217.
- MASLAND, R. H.** 1986. The Functional Architecture of the Retina. In R. R. Llinas (Ed.). *The Biology of the Brain: From Neurons to Networks. Readings From Scientific American*. W. H. Freeman, England.
- MASSARO, D. W.** 1988. Some Criticisms of Connectionist Model of Human Performance. *Journal of Memory and Language*. Vol. 27, pp. 213-234.
- MAUNSELL, J. H. R. & NEWSOME, W. T.** 1987. Visual Processing in Monkey Extrastriate Cortex. *Annual Review of Neuroscience*. Vol. 10, pp. 363-401.
- MEAD, C. & MAHOWALD, M. A.** 1990. A Silicon Model of Early Visual Processing. In E. L. Schwartz (Ed.). *Computational Neuroscience*. MIT Press, Cam. MA.
- MILLER, K. D. & STRYKER, M. P.** 1990. Ocular Dominance Column Formation: Mechanisms and models. In S. J. Hanson & C. R. Olson (Eds.) *Connectionist Modelling and Brain Function: The Developing Interface*. Cambridge, MA. MIT Press.
- MISHKIN, M., UNGERLEIDER, L. G. & MAKO, K. A.** 1983. Object Vision and Spatial Vision: Two Cortical Pathways. *Trends In Neuroscience*. Vol. 6, pp. 414-417.
- MORAN, J. & DESIMONE, R.** 1985. Selective Attention Gates Visual Processing in the Extrastriate Cortex. *Science*. Vol. 229, pp. 782-784.
- MUMFORD, D.** 1991. On the Computational Architecture of the Neocortex I: The Role of the Thalamo-Cortical Loop. *Biological Cybernetics*. Vol. 65, pp. 135-145.



- MUMFORD, D.** 1992. On the Computational Architecture of the Neocortex II: The Role of the Thalamo-Cortical Loop. *Biological Cybernetics*. Vol. 66, pp. 241-251.
- MURTHY, V. N. & FETZ, E. E.** 1991. Synchronized 25-35 HZ Oscillations in Sensori-Motor Cortex of Awake Monkey's. *Society of Neuroscience Abstracts*. Vol. 17, pp. 310.
- NAKAYAMA, K.** 1990. The Iconic Bottleneck and the Tenuous Link Between Early Visual Processing and Perception. In C. Blakemore (Ed.), *Vision: Coding and Efficiency*. Cambridge University Press, Cambridge, England.
- NEWELL, A. & SIMON, H. A.** 1961. GPS: A Program that Simulates Human Thought. In H. Billing (Ed.), *Lernende Automaten*. R. Oldenbourg, Munich.
- OLSON, G. M. & SHERMAN, T.** 1983. Attention, Learning and Memory in Infants. In M. M. Haith & J. J. Campos (Eds.) *Handbook of Child Psychology*. Vol. 2: *Infancy and Developmental Psychobiology*. Wiley, New York.
- OHLSSON, S.** 1993. Psychological Implications of the Synchronicity Hypothesis. *Open Peer Comentary in: SHASTRI, L. & AJJANAGADDE, V.* 1993. *From Simple Associations to Systematic Reasoning: A Connectionist Representation of Rules, Variables and Dynamic Bindings Using Temporal Synchrony*. *Behavioral and Brain Sciences*. Vol. 16, pp 417-494.
- ORAM, M. W. & PERRETT, D. J.** 1994. Modelling Visual Recognition from Neurobiological Constraints. *Neural Networks*. Vol. 7, No. 6/7, pp. 945-972.
- O' SHAUGHNESSEY, F. P.** 1995. Towards a Model of Declarative Memory. *Internal Publication, Neurodynamics Research Group, School of Computing, University of Plymouth, Devon. UK*
- PERRETT, D. J., ROLLS, E. T. & CAAN, W.** 1982. Visual Neurones Responsive to Faces in the Monkey Temporal Cortex. *Experimental Brain Research*. Vol. 47, pp. 329-342.
- PERRETT, D. J., MISTLIN, A. J., POTTER, D. D., SMITH, P. A. J., HEAD, A. S., CHITTY, A. J., BROENNIMANN, R., MILNER, A. D., & JEEVES, M. A.** 1986. Functional Organisation of Visual Neurones Processing Face Identity. In *Aspects of face processing*. H. D. Ellis, M. A. Jeeves, F. Newcombe & A. Young (Eds.), Dordrecht: Martinus Nijhoff.
- PERRETT, D. J., MISTLIN, A. J., CHITTY, A. J.** 1987. Visual Neurones Responsive to Faces. *Trends In Neuroscience*. Vol. 10, No. 9, pp. 358-364.
- POLLACK, J. B.** 1990. Recursive Distributed Representations. *Artificial Intelligence*. Vol. 46, pp. 77-107.
- RAMACHANDRAN, V. S. & COBB, S.** 1995. Visual Attention Modulates Metacntrast Masking. *Nature*. Vol. 373, pp. 66-68.
- RATCLIFF, G. & NEWCOMBE, F.** 1982. Object Recognition: Some deductions From the Clinical Evidence. In A. W. Ellis (Ed.). *Normality and Pathology in Cognitive Functions*. Academic Press, London.

- RATCLIFF, R.** 1990. Connectionist Models of Recognition Memory: Constraints Imposed by Learning and Forgetting Functions. *Psychological Review*. Vol. 97, pp. 285-308.
- RIBARY, U., IOANNIDES, A. A., SINGH, K. D., HASSON, R., BOLTON, J. P. R., LADO, F., MOGILNER, A. & LLINAS, R.** (1991). Magnetic Field Tomography of Coherent Thalamocortical 40 Hz Oscillations in Humans. *Proceedings of the National Academy of Science (USA)*. Vol. 88, pp. 11037-11041.
- ROBERTS, L. G.** 1965. Machine Perception of Three-Dimensional Solids. In J. T. Tippett et al (Eds.). *Optical and Electro Optical Information Processing*, MIT Press, Cambridge, MA.
- ROSENQUIST, A. C.** 1985. Connections of visual Cortical Areas in the Cat. In A. Peters & E.G. Jones. *Cerebral cortex*. Vol. 3. *Visual cortex*. New York: Plenum, pp. 81-117.
- RUMELHART, D. E. & McCLELLAND, J. L.** 1986. PDP Models and General Issues in Cognitive Science. In D. E. Rumelhart & J. L. McClelland (Eds.). *Parallel Distributed processing: Explorations in the Microstructure of Cognition*. Volumes 1 & 2. MIT Press, USA
- SCHANK, R. C. & ABELSON, R. P.** 1977. Scripts, Plans, Goals and Understanding: An Inquiry into Human Knowledge Structures. *Lawrence Erlbaum, Hillsdale NJ*.
- SCHILLER, P. H.** 1982. Central Connections of the Retinal ON and OFF Pathways. *Nature*. Vol. 297, pp. 580-583.
- SCHNEIDER, G. E.** 1969. Two Visual systems. *Science*. Vol. 163, pp. 895-902.
- SEGEV, I.** 1992. Single Neurone Models: Oversimple, Complex and Reduced. *Trends in Neuroscience: Special Issue: Modelling the Nervous System*. Vol. 15, No. 1, pp. 414-421.
- SEJNOWSKI, T. J., KOCH, C. & CHURCHLAND, P. SMITH.** 1988. Computational Neuroscience. *Science*. Vol. 241, pp. 1299-1306.
- SERVICE, R. F.** 1993. Making Modular Memories. *Science*. Vol. 260, pp. 1876.
- SHASTRI, L. & AJJANAGADDE, V.** 1993. From Simple Associations to Systematic Reasoning: A Connectionist Representation of Rules, Variables and Dynamic Bindings Using Temporal Synchrony. *Behavioral and Brain Sciences*. Vol. 16, pp 417-494.
- SHEPPARD G. M.** 1979. The Synaptic Organisation of the Brain. *Second edition*. OUP, Oxford, England.
- SHIPP, S. & ZEKI, S.** 1985. Segregation of Pathways Leading from Areas V2 to V4 and V5 of Macaque Monkey Visual Cortex. *Nature*. Vol. 315, pp. 322-325.
- SILLITO, A. M., JONES, H. E., GERSTEIN, G. L. & WEST, D. C.** 1994. Feature-Linked Synchronization of Thalamic Relay Cell Firing Induced by Feedback from the Visual Cortex. *Nature*. Vol. 369, pp. 479-482.

- SKARDA, C. A. & FREEMAN, W. J.** 1987. How Brains Make Chaos in Order to Make Sense of the World. *Behavioral and Brain Sciences*. Vol. 10, pp 161-195.
- SLATER, A., MORISON, V., TOWN, C. & ROSE, D.** 1985. Movement Perception and Identity Constancy in the New Born Baby. *British Journal of Developmental Psychology*. Vol. 3, pp. 211-220.
- SOMPOLINSKY, H., GOLOMB, D. & KLEINFELD, D.** 1990. Global Processing of Visual Stimuli in a Neural of Coupled Oscillators. *Proceedings of the National Academy of Science (USA)*. Vol. 87, pp. 7200-7204.
- SPRAGUE, J. M., LEVY, J., DIBERARDINO, A. & BERLUCCHI, G.** 1977. Visual Cortical Areas Mediating Form Discrimination in the Cat. *Journal of Comparative Neurology*. Vol. 172, pp. 441-448.
- STONE, J.** 1983. Parallel Processing in the Visual System: The Classification of Retinal Ganglion Cells and its Impact on the Neurobiology of Vision. *Plenum Press, New York & London*.
- STONE, J. & FUKUDA, Y.** 1974. Properties of Cat Retinal Ganglion Cells: A Comparison of W Cells with X and Y Cells. *Journal of Neurophysiology*. Vol. 37, pp. 722-748.
- STRYKER, M. P.,** 1989. Is Grandmother an Oscillation? *Nature*. Vol. 338, pp. 297-298.
- TAYLOR, J. G. & ALAVI, F. N.** 1993. A Global Competitive Network for Attention. *Neural Networks World*. Vol. 3, No. 5, pp. 477-503.
- TENENBAUM, J. M. & BARROW, H. G.** 1976. Experiments in Interpretation-Guided Segmentation. *Stanford Research Institute Technical Note 123*.
- THOMPSON, R. F.** 1985. The Brain: An Introduction to Neuroscience. *W. H. Freeman, USA*
- TINBERGEN, N.** 1952. Derived Activities. *Quarterly Review of Biology*. Vol. 27, pp. 1-32.
- TROUP, L. J.** 1991. An Investigation into the Dynamics of Oscillatory Neural Networks. *Project Dissertation, MSc Intelligent Systems, University of Plymouth (formerly Polytechnic South West)*.
- UNGERLEIDER, L. G. & MISHKIN, M.** 1982. Two Cortical Visual Systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.). *Analysis of Visual Behaviour*. MIT Press, Cam. MA.
- VAN ESSEN, D. C., FELLEMAN, D., DE YOE, E. & KNIERIM, J.** 1991. Probing the Primate Visual Cortex: Pathways and Perspectives. In A. Valberg & B. B. Lee (Eds.). *Advances in Understanding Visual Process*. Plenum, New York.
- VAN ESSEN, D. C.** 1985. Functional Organisation of Primate Visual Cortex. In A. Peters & E.G. Jones. *Cerebral cortex*. Vol. 3, Visual cortex. Plenum, New York, pp. 259-329.

- VON DER MALSBURG, C., 1981. The Correlation Theory of Brain Function. *Internal report, Department of Neurobiology, Max-Planck Institute for Biophysical Chemistry.*
- VON DER MALSBURG, C., 1990. Considerations for a Visual Architecture. In R. Eckmiller (Ed.). *Advanced Neural Computers. Elsevier Science Publishers B. V.*, pp. 303-312.
- WALTZ, D. 1975. Understanding line Drawings of Scenes with Shadows. In P. H. Winston (Ed.). *The Psychology of Computer Vision, McGraw-Hill, New York.*
- WEHMEIER, U., DONG, D., KOCH, C. & VAN ESSEN, D. 1989. Modelling the Mammalian Visual System. In C. Koch & I. Segev (Eds.). *Methods in Neuronal Modelling. 1989. MIT Press, Cam. MA & London England.*
- WERBLIN, F. S. & DOWLING, J. E. 1969. Organisation of the Retina of the Mudpuppy, *Necturus Maculosus*. II. Intracellular Recording. *Journal of Neurophysiology. Vol. 32, pp. 339-355.*
- WEISKRANTZ, L. 1986. Blindsight: A Case Study and Implications. *Oxford University Press (Clarendon Press) Oxford.*
- WEISKRANTZ, L. 1992. Introduction: Dissociated Issues. In A. D. Milner & M. D. Rugg (Eds.). *The Neuropsychology of Consciousness. Academic Press, London.*
- WERNER, G., REITBOECK, H. J. & ECKHORN, R. 1993. Constructions of Concepts by the Nervous System: From Neurons to Cognition. *Behavioral Science. Vol. 38, pp. 114-123.*
- WILSON, F. A. W., O' SCALCIDHE, F. O. & GOLDMAN-RAKIC, P. S. 1993. Dissociation of Object and Spatial Processing Domains in Primate Prefrontal Cortex. *Science. Vol. 260, pp. 1955-1958.*
- WORGOTTER, F. & KOCH, C. 1991. A Detailed Model of the Primary Visual Pathway in the Cat: Comparison of Afferent Excitatory and Intracortical Inhibitory Connection Schemes for Orientation Selectivity. *The Journal of Neuroscience. Vol. 11, No. 7, pp. 1959-1979.*
- YOUNG & DE HAAN, 1992. Face Recognition and Awareness. In A. D. Milner & M. D. Rugg (Eds.). *The Neuropsychology of Consciousness. Academic Press, London.*
- YOUNG, M. P., TANAKA, K. & YAME, S. 1992. On Oscillating Neuronal Responses in the Visual Cortex of the Monkey. *Journal of Neurophysiology. Vol. 67, pp. 1464-1474.*
- ZEKI, S. M. 1978a. Functional Specialisation in the Visual Cortex of the Rhesus Monkey. *Nature. Vol. 274, pp. 423-428.*
- ZEKI, S. M. 1978b. Uniformity and Diversity of Structure and Function in Rhesus Monkey Prestriate Visual Cortex. *Journal of Physiology. Vol. 227, pp. 273-290.*
- ZEKI, S. M. 1980. The Representation of Colours in the Cerebral Cortex. *Nature. Vol. 284, pp. 412-418.*

- ZEKI, S. M.** 1983. Colour Coding in the Cerebral Cortex: The Reaction of Cells in Monkey Visual Cortex to Wavelength and Colours. *Neuroscience*. Vol. 9, pp. 741-765.
- ZEKI, S. M. & SHIPP, S.** 1988. The Functional Logic of Cortical Connections. *Nature*. Vol. 335, pp. 311-317.
- ZEKI, S.** 1992. The Visual Image in Mind and Brain. *Scientific American*, Special issue September 1992, pp. 24-33.