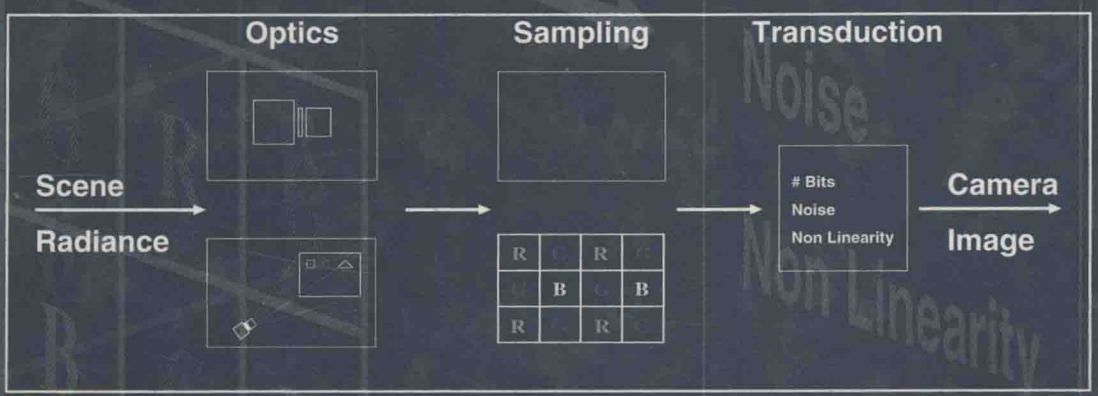
VISION MODELS AND APPLICATIONS TO IMAGE AND VIDEO PROCESSING

Edited by
Christian J. van den Branden Lambrecht



Kluwer Academic Publishers

Vision Models and Applications to Image and Video Processing

Edited by

Christian J. van den Branden Lambrecht

EMC Media Solutions Group

江苏工业学院图书馆
藏书章

KLUWER ACADEMIC PUBLISHERS
BOSTON / DORDRECHT / LONDON

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN 0-7923-7422-3

Published by Kluwer Academic Publishers,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

Sold and distributed in North, Central and South America
by Kluwer Academic Publishers,
101 Philip Drive, Norwell, MA 02061, U.S.A.

In all other countries, sold and distributed
by Kluwer Academic Publishers,
P.O. Box 322, 3300 AH Dordrecht, The Netherlands.

Printed on acid-free paper

All Rights Reserved
© 2001 Kluwer Academic Publishers, Boston
No part of the material protected by this copyright notice may be reproduced or
utilized in any form or by any means, electronic or mechanical,
including photocopying, recording or by any information storage and
retrieval system, without written permission from the copyright owner.

Printed in the Netherlands.

Contributors

Karl-Heinz Bäuml
Institut für Psychologie, Universität Regensburg, 93040 Regensburg, Germany

Martin Boschman
IPO - Center for User-System Interaction, 5600 MB Eindhoven,The Netherlands

Geoffrey M. Boynton
The Salk Institute for Biological Studies, La Jolla, CA 92037, USA

David H. Brainard
Department of Psychology, UC Santa Barbara, Santa Barbara, CA 93106-9660

Thom Carney
Neurometrics Institute,Berkeley, CA 94704

Lisa J. Croner
The Salk Institute for Biological Studies, La Jolla, CA 92037, USA

Scott Daly
Sharp Laboratories of America, Camas, WA 98607

Hans du Buf
Vision Laboratory, Dept. of Electronics and Computer Science, University of Algarve, Faro, Portugal

Stanley A. Klein
School of Optometry, University of California, Berkeley, CA 94720

Murat Kunt
Signal Processing Laboratory, Swiss Federal Institute of Technology, 1015 Lausanne, Switzerland

Philippe Longère
Department of Psychology, UC Santa Barbara, Santa Barbara, CA 93106-9660

Jean-Bernard Martens
Center for User-System Interaction, 5600 MB Eindhoven,The Netherlands

Lydia Meesters
Center for User-System Interaction, 5600 MB Eindhoven,The Netherlands

Amnon Silverstein
Imaging Technology Department,Hewlett Packard Laboratories,Palo Alto, CA 94304

Christian J. van den Branden Lambrecht
EMC Media Solutions Group, Hopkinton, MA 01748, USA

Thomas Wachtler
The Salk Institute for Biological Studies, La Jolla, CA 93027, USA.and Institute for Biology III, Neurobiology & Biophysics, University of Freiburg, 79104 Freiburg, Germany

Brian Wandell
Psychology Department, Stanford University, Stanford, CA 94305, USA

Stefan Winkler
Signal Processing Laboratory, Swiss Federal Institute of Technology, 1015 Lausanne, Switzerland

Xuemei Zhang
Agilent Laboratories, Palo Alto, CA 94304, USA

此为试读,需要完整PDF请访问: www.ertongbook.com

Preface

I came to vision science trying to solve an engineering problem: I was trying to come up with test and measurement methodologies for digital video systems. One of the metrics I wanted to use was some measurement of image quality. After some experiments and after an overview of the literature, I had to realize that simple computational metrics, such as the mean square error, are not very effective for this purpose. This led me to study notions of vision science and vision modeling. As an engineer, I found it fascinating. Vision science uses computational tools and modeling techniques that are very similar to what we use in signal processing, yet it brings you to a new domain that lies at the intersection of engineering, biology and cognitive psychology.

Over the years, vision science has made tremendous contributions to engineering and to the field of image processing in particular. Such contributions include halftoning matrices for the printing industry, color correction for digital cameras, quantization matrices for image coding. As imaging devices are becoming commodities, the impact of vision science is becoming even more significant.

This book is meant to appeal to an engineering audience. It is an introduction to vision science and to the design and application of vision models in engineering. To achieve this goal, we have chosen to organize the book around the main components of vision models. In our discussion, we assume a basic knowledge of signal and image processing as well as some intimacy with image coding.

The book begins with an overview of the neurophysiology of the primate visual system. This chapter describes the structure of the visual system and highlights the features that will be relevant to the design of vision models. The modeling of brightness and spatial perception is introduced in Chapter 2, followed, in Chapter 3, by an analysis of visual masking. These two notions lay the ground for the next stage of modeling, addressed in Chapter 4: the detection of spatial patterns. At that point, we will have established the general architecture of most vision models and we will have set the ground for a discussion in Chapter 5 on psychophysical experimentation. These are the techniques used to derive model parameters from experiments.

Chapter 6 introduces the notion of color and discusses color spaces and color metrics. Chapter 7 is a direct application of such modeling: the simulation and evaluation of digital cameras. In Chapter 8, we revisit the perception of image quality and discuss multidimensional quality models.

The last two chapters of the book move on to video: Chapter 9 addresses spatio-temporal perception, along with the impact of eye movements. Finally Chapter 10 concludes the book with an overview of spatio-temporal models geared at the evaluation of image quality in digital television applications.

I would like to take a moment to thank all the contributors to this book, who, despite their very busy life, accepted to take on this project. We are all very grateful to the staff of Kluwer Academic Publishers for the constant help they provided us with. Finally, I would like to thank Prof Murat Kunt, my former advisor, who has been the instigator of this project and Joyce Farrell, my manager at HP Labs, for her support in this task.

As I reflect back on this book, I realize how fascinating and thought-provoking vision science is and I remember the excitement of discovering vision models and the contributions they bring to imaging science. I also feel very fortunate to have had the chance to work in this field among a pool of highly talented people at HP Labs.

Christian J. van den Branden Lambrecht

Contents

Chapter 1

INTRODUCTION TO NEUROPHYSIOLOGY OF THE PRIMATE VISUAL SYSTEM

Lisa J. Croner[1] and Thomas Wachtler [2,3]
1. *The Salk Institute for Biological Studies, La Jolla, CA 92037, USA.* 2. *The Sloan Center for Theoretical Neuroscience at The Salk Institute for Biological Studies, La Jolla, CA 93027, USA.* 3. *Institute for Biology III, Neurobiology & Biophysics, University of Freiburg, 79104 Freiburg, Germany*

The visual system is the most studied sensory system. This chapter provides a brief introduction to the neurophysiology of the primate visual system, highlighting at the end some of the questions for which modeling may be useful. The information presented, while reflecting our understanding of the primate brain, comes from research in both primates (primarily Old World monkeys) and, for some processes that are likely conserved across higher taxa, other vertebrates.

1. FUNDAMENTALS

We begin with some fundamental topics: neurons, the stages of visual processing, receptive fields, and stimulus selectivity.

1.1 Neurons: Processing Cells of the Nervous System

The nervous system is composed of specialized cells called neurons. Neurons form a network throughout the body, but are most concentrated in the brain and spinal cord, where neuronal circuits process and integrate sensory information to produce perception, thought, and motor output. While a neuron typically has three regions – a cell body; a region of branching processes, called dendrites, extending from the cell body; and a single or branched process, called the axon, also extending from the cell

C.J. van den Branden Lambrecht (ed.),
Vision Models and Applications to Image and Video Processing, 1–19.
© 2001 *Kluwer Academic Publishers. Printed in the Netherlands.*

body – there is considerable diversity in the morphology and arrangement of these parts (Figure 1A). A neuron receives input from other neurons through junctions, termed synapses, on its dendrites and sends output to other neurons via the axon, which forms synapses with the dendrites of other neurons. Neuronal output is generally in the form of brief, stereotyped changes in the membrane voltage, initiated at the cell body and travelling along the axon, that can be measured with a microelectrode (Figure 1B). These travelling voltage changes are generated by the dynamics of ion channels on the cell membrane, and are called action potentials or spikes.

There are two general classes of synapses: chemical and electrical. When a spike arrives at a chemical synapse, the pre-synaptic neuron releases, at the synapse, a chemical substance (neurotransmitter) that diffuses across the cleft between the neurons and activates specialized proteins (receptors) on the membrane of the post-synaptic dendrite. This causes a small change in the membrane potential of the post-synaptic dendrite. The size of the membrane potential change in the dendrite is related to the number of spikes that arrive at the synapses within a few milliseconds of each other. There are two kinds of chemical synapses: excitatory (driving the post-synaptic neuron's membrane potential towards generating a spike), and inhibitory (driving the membrane potential away from generating a spike). The higher the net excitation, the more likely it is that the post-synaptic neuron will generate a spike. At an electrical synapse, the spike current in the pre-synaptic neuron spreads directly to the post-synaptic neuron through a low-resistance pathway. For both chemical and electrical synapses, the number and strength of synapses between neurons determine the degree to which one neuron influences another.

Neurons have been classified into many types based on morphology, connection patterns, and subcellular components. For this chapter, it is sufficient to appreciate that the brain processes signals with a network of many types of neurons, connected via synapses into complex circuits.

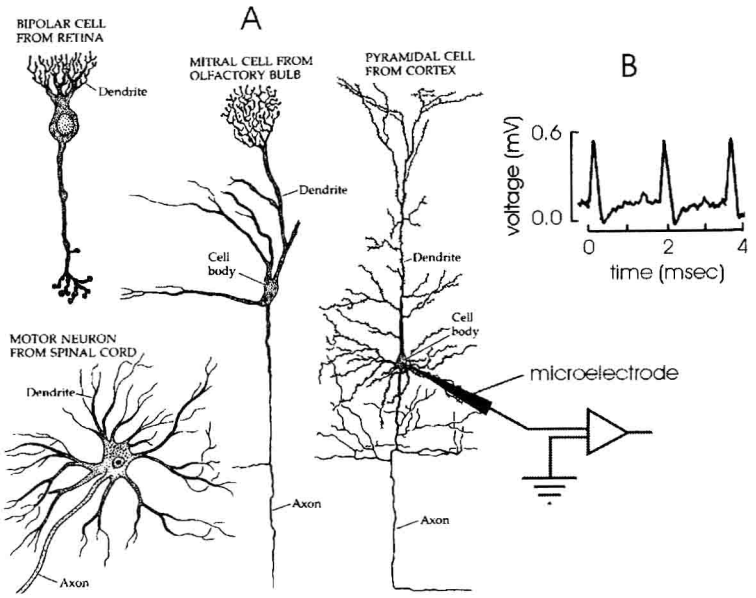


Figure 1. Neurons and action potentials. A) Four neurons illustrate the diversity of neuronal morphology. (Reprinted from [1]). B) A microelectrode measures the voltage at its tip relative to an external reference. Action potentials are recorded as brief, stereotyped changes in the voltage. The time course of three sequential action potentials is shown.

1.2 Stages of Visual Processing and How They are Studied

The visual system is defined by its function: it is the portion of the brain involved in processing information about light to produce visual perception and visually guided behavior. Anatomical and physiological studies have revealed the paths that visual information follows through the brain. The main path, posited to generate conscious vision, is diagrammed in Figure 2. Light enters the eye through the cornea, which, together with the lens, forms an image on the retina. Photoreceptors (a class of retinal neuron) transduce light into neural signals, and other retinal neurons begin to process these signals. Axons projecting from the retina form the optic nerve. While about 10% of these axons project to subcortical pathways, axons in the pathway that concerns us here carry signals to the lateral geniculate nucleus (LGN) of the thalamus. These axons form synapses with LGN neurons whose axons project to the primary visual cortex, also known as striate cortex or V1. There, visual information is further processed and conveyed to higher cortical areas (not shown).

此为试读,需要完整PDF请访问: www.ertongbook.com

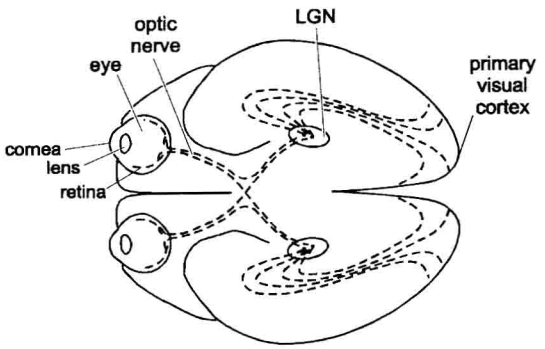


Figure 2. Diagram of the first stages in the main visual pathway (dashed lines) in primates, viewed on a transverse plane.

To study the visual system, neurophysiologists typically examine how neural activity in the visual pathway relates to light stimuli (the input to the system) and/or to visual perception (the output). The predominant technique used to assay neuronal activity is single-cell electrophysiology. This involves using a microelectrode to measure the action potentials produced by single neurons (Figure 1B). To look at activity on a larger spatial scale, several modern approaches have been developed. These include multi-electrode techniques, optical imaging, electroencephalography, magnetoencephalography, positron emission tomography, and functional magnetic resonance imaging. The majority of findings discussed in this chapter stem from single-cell microelectrode studies.

1.3 Visual Receptive Fields and Stimulus Selectivity

Two important concepts are useful in characterizing the responses of a visual neuron. The first concept is that of the receptive field. A visual neuron's receptive field is the region on the retina where a light stimulus elicits a change in the neuron's response (the rate at which it fires action potentials, or for some neurons a change in the graded membrane potential, measured with a microelectrode) [2]. For example, a photoreceptor's membrane potential changes when an appropriate stimulus appears on the small region (the receptive field) of the retina where the photoreceptor lies. Because of the eye's optics, this region corresponds to a particular spot in visual space. If several photoreceptors synapse with another neuron, this neuron's receptive field will be a composite of the contributing receptive fields. The corresponding region in visual space moves with the eye, so the stimulus in the receptive field depends on both the light pattern outside the eye and the eye position (Figure 3).

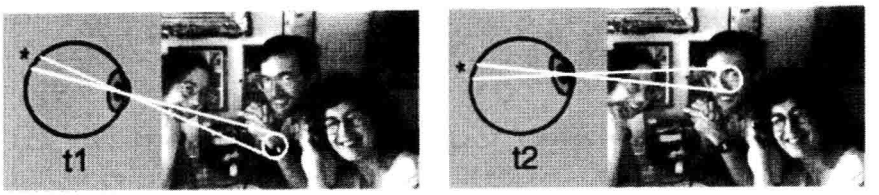


Figure 3. A visual neuron's receptive field is a region on the retina (an asterisk outside the eye indicates the location of one hypothetical receptive field) where light can modulate the cell's response. As the eye rotates (from t1 to t2), the receptive field remains stationary on the retina while the corresponding region in visual space (indicated by white circles on photograph) moves, causing the stimulus in the receptive field to vary.

Receptive fields of neurons beyond the photoreceptors generally have structure – for example, they may have regions where light increases the response (excitation) and regions where light decreases the response (inhibition). Receptive fields tend to increase in complexity further along in the visual processing hierarchy. For instance, receptive fields of retinal and LGN neurons are circular, while those of many V1 neurons are elongated, sometimes with spatiotemporal structure capable of encoding motion [3].

Typically, a visual neuron responds only to some stimuli appearing in its receptive field. This leads to the second important concept: stimulus selectivity, which refers to a neuron's capacity to respond differentially to variation along one stimulus dimension, such as luminance, chromaticity, orientation, motion direction, etc. For example, many V1 neurons are selective along the dimension of orientation; a bar oriented at a particular angle evokes the largest response, while bars at other orientations evoke smaller responses. A given neuron may be selective along more than one stimulus dimension. Neurons at higher levels of processing tend to be selective along a greater number of dimensions [4].

2. THE STAGES OF VISUAL PROCESSING

Armed with the fundamental concepts introduced above, we can now consider each stage of visual processing in more detail.

2.1 Optics

Figure 4A provides a diagram of an eye. The cornea and lens form an image of the environment on the retina, and can be considered the first stage of visual processing. The pupil (surrounded by the iris) is an aperture that dynamically varies between 2 and 8 mm in diameter. The lens has elastic

properties allowing it to expand and contract to focus light onto the retina; lens shape is controlled by the ciliary muscle to allow accomodation up to 10 diopters.

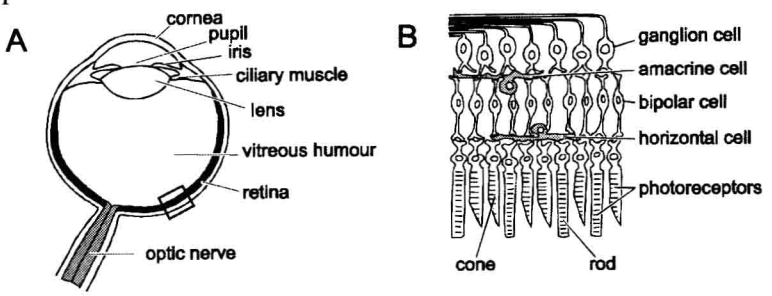


Figure 4. A) Diagrammatic cross-section through a primate eye. The small square indicates the region diagrammed in part B. B) Diagrammatic cross-section through the retina, showing the laminar organization of the five types of retinal neurons.

As in any optical system, the eye's optical quality is limited by diffraction and aberrations. Due to the adaptability of the pupil and lens, these effects depend in part on viewing conditions. The consequence is that the quality of the image on the retina is variable, and this affects a number of visual capacities – including contrast sensitivity and visual acuity.

2.2 Retina

The retina's position at the back of the eye makes it easily accessible and, thus, one of the most studied parts of the brain. We therefore discuss retinal processing at greater length than some of the other stages. The retina serves two main functions. First, retinal photoreceptors transduce light energy into neural signals (changes in membrane potential). Second, other retinal neurons process the information provided by the photoreceptors to produce output (action potentials) for transmission through the optic nerve. Retinal neurons are organized in layers, and are of five distinct types (Figure 4B): the photoreceptors and ganglion cells serve as the retinal input and output cells, respectively, while the horizontal, bipolar, and amacrine cells constitute the intermediate processing circuitry.

2.2.1 Photoreceptors

Photoreceptors transform light energy into changes in membrane potential – a process known as phototransduction. Phototransduction begins when a photon is absorbed by a photopigment molecule on the photoreceptor membrane. This initiates a series of subcellular events that culminate in temporarily blocking an ion current maintained across the membrane in the

dark. The voltage across the photoreceptor membrane changes (in a graded fashion depending on the number of photons absorbed), and this alters the rate of neurotransmitter released at the photoreceptor's synapses. Phototransduction is regulated by intracellular mechanisms so that neurotransmitter release adapts rapidly to the current light level, over several orders of magnitude of intensity. Photoreceptors thus signal *changes* in light intensity.

The human retina contains two kinds of photoreceptors – rods and cones (see Figure 4B). Functionally, rods and cones differ in a number of respects, the most notable being the luminance range over which they provide signals. Figure 5 diagrams the human visual system's operating range, along with the operating ranges of rods and cones. Rods provide information at low light levels (scotopic), but their responses saturate for the high light levels (photopic) at which cones provide informative responses.

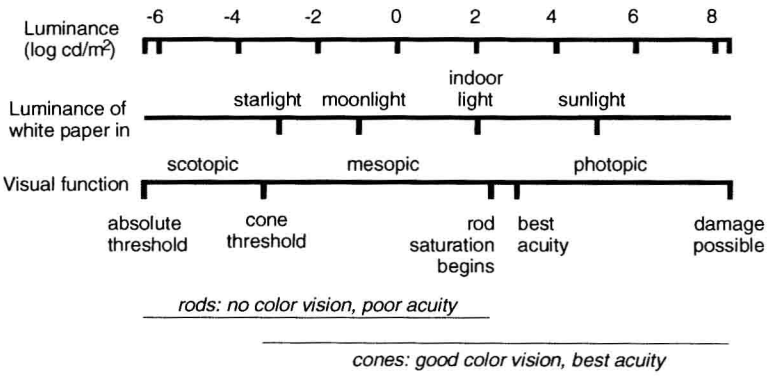


Figure 5. Operating range of the visual system (see [5]).

Our vision has different characteristics at different light levels, and this can be attributed in part to the photoreceptors. Under photopic conditions we can discriminate many different colors. Color vision stems from the presence of three kinds of cones distinguished by three different photopigments, each sensitive to a different range of wavelengths (Figure 6). Each cone type alone is unable to distinguish changes in spectral composition from changes in intensity, but different spectral compositions lead to different *relative* excitations of the three cone types. Thus, color coding is based on a comparison of the signals from the different cone types. By contrast, scotopic vision, which uses rods, lacks color vision because there is only one kind of rod photopigment (Figure 6).

Another characteristic that varies with light level is acuity. Acuity is best at photopic light levels and in the fovea (the retinal region serving the center of vision), which is devoid of rods but densely packed with cones. The high resolution of the foveal cone sampling array gives us our excellent foveal

acuity. Away from the fovea, cones are larger and less densely packed (interspersed with rods), and acuity is worse.

Sensitivity to light increments also varies with light level. At the lowest light levels at which we can see, we are phenomenally sensitive to light – under optimal conditions we can essentially see single photons ([6, 7]; for review see [8]). Electrophysiological studies have confirmed that under these conditions a rod responds to the absorption of a single photon with a large, reproducible signal against a background of very low noise. Through unknown mechanisms, this signal is conveyed through the rest of visual processing to result in a reliable percept of light. At photopic light levels, when cones are used, the ability to detect single photons is lost. It is not known what factors limit the sensitivity of cone vision.

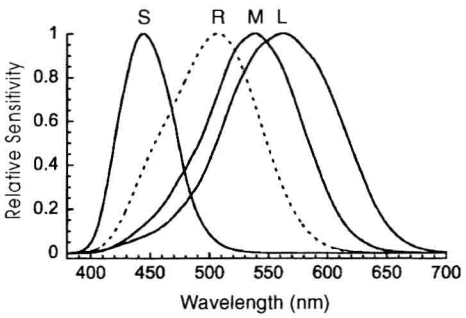


Figure 6. Spectral sensitivity functions of rods (R) [9] and of the three kinds of cones: long (L)-, medium (M)-, and short (S)-wavelength sensitive cones [10].

2.2.2 Horizontal, Bipolar, and Amacrine Cells

Visual signals pass from the photoreceptors to the bipolar cells. The photoreceptor-bipolar synapse can either preserve or invert the direction of the photoreceptor response, leading to two kinds of bipolar cells – those excited by light onset and those excited by light offset. Bipolar cells are also influenced by the activity of horizontal cells, whose dendritic and axonal arbors ramify in the region where photoreceptors synapse with bipolar cells. Horizontal cells receive input from multiple photorecepors and form electrical synapses with neighboring horizontal cells. The horizontal cell signal is thus a measure of net light intensity within a localized region. This signal is fed back to the photoreceptor synapses so as to attenuate the photoreceptor signal. The consequence is that bipolar cells have receptive fields with "center-surround" organization (Figure 7): a small central region in which light causes a response of one polarity (excitation: "on-center;" inhibition: "off-center") is surrounded by a larger region in which light

evokes a response of the opposite polarity. Center-surround receptive fields can be usefully modeled as the sum of two antagonistic components, both with circular profiles on the retina [11] (Figure 7B). Because of the antagonism, bipolar cell responses represent the *contrast* between light in the two regions. In addition, the synapses between cones and bipolar cells are sensitive to temporal contrast: spatial contrasts that do not change over time are suppressed.

Amacrine cells ramify in the layers where bipolar cells synapse with ganglion cells. Although at least 20 different amacrine cell types have been distinguished, their functional roles are poorly understood. Amacrine cells receive input from bipolar cells and other amacrine cells, and send output to bipolar cells, other amacrine cells, and ganglion cells. This circuitry suggests that amacrine cells provide lateral information processing, perhaps strengthening the contrast representation in the center-surround organization of typical ganglion cell receptive fields (see below).

2.2.3 Retinal Ganglion Cells

The neuronal circuitry of the retina leads ultimately to the retinal ganglion cells, whose axons form the optic nerve. These are the first neurons in the visual pathway to produce action potentials, which are the dominant mode of signalling throughout the rest of the visual system. Like bipolar cells, retinal ganglion cells typically have center-surround receptive fields with on- or off-center organization (Figure 7).

2.2.3.1 Main classes of retinal ganglion cells

Two groups of ganglion cells – the midget and parasol cells (Figure 8) – have received the most attention (*e.g.* [12]). These neurons account for about 90% of the ganglion cell population. Midget ganglion cells have relatively small dendritic trees (Figure 8A) that collect input from a small area of the retina; in the fovea, the receptive field center of a midget cell probably corresponds to a single cone. The surrounds are generally much larger, consistent with contributions from the broader dendritic trees of horizontal and/or amacrine cells. Parasol ganglion cells, on the other hand, have broader dendritic trees than do midget cells (Figure 8A), giving parasol cells larger receptive field centers. However, the sampling array of parasol cells is less dense than that of midget cells; for a given retinal area, there is one parasol cell for every ten midget cells. The lower parasol density is compensated by the larger receptive fields, so that both midget and parasol cell arrays tile the retina without gaps.

此为试读,需要完整PDF请访问: www.ertongbook.com