Introduction

Vision is probably the most complex sensory function the human brain performs. A large part of the brain is dedicated to this task. It involves numerous specific functions such as motion detection, recognition of objects, and colour vision. But before visual input reaches the parts of the brain that are specialized for these tasks, it has already been through a large number of transformation steps. The first transformations are performed by an intricate neural network, which is called the retina. In this thesis, these first steps of processing are studied, in particular light adaptation. In this Introduction I will discuss what light adaptation is, how it can be studied, why it is necessary to have a model of light adaptation, and how well light adaptation and early vision work under natural conditions.

WHAT IS LIGHT ADAPTATION?

One of the basic operations the retina performs is the subject of the first four chapters of this thesis: light adaptation. Light adaptation is the term used for the process that changes the sensitivity of the visual system to different light levels. In natural conditions the ambient light level can vary by a factor of about 10 log units (Shapley & Enroth-Cugell, 1984). The biological hardware of the system does not possess such a large dynamic range (Shapley & Enroth-Cugell, 1984; Wilson, 1997). Therefore, it is necessary that the sensitivity of the system is continually adjusted in order to allow efficient transfer of information about the visual input to the brain. Without such an adjustment, small signals will drown in neuronal noise, and large signals will saturate the system. The purpose of light adaptation is thus to keep the response to rapidly varying visual input within the dynamic range of the neurons in the retina.

Because the range of light levels is so large, adaptation takes place at several different sites in the visual system (Hood & Finkelstein, 1986; Walraven et al., 1990). These different adaptation processes complement each other. Before light reaches the retina its intensity is already regulated by the size of the eye's aperture, the pupil. Another adaptation process is formed by the division of photoreceptors into rods and cones. The rod system is very sensitive and works mainly at low (scotopic) light levels, but it is saturated at photopic (daylight) light levels. This is when the cone system becomes active. At very high light intensities a third adaptation process becomes important, namely bleaching of photopigment. This reduces the amount of photopigment available and thereby prevents saturation of the cone system. Although these three processes are important for obtaining a complete picture of adaptation over the whole range of light intensities, they will not be discussed in this thesis.
Introduction

The present study focuses on the adaptation processes that take place at photopic light levels in the various layers of neurons in the retina. At a certain time of day, it may appear that the luminance level is fairly constant. This is not the case, however, because when shifting gaze the luminance on a certain point on the retina can vary by up to three log units within a time of several tens of milliseconds (van Hateren, 1997). It is therefore necessary to have adaptation processes that adjust the gain of the system quickly and continually. The mechanisms to be discussed must be active all the time. But they often go unnoticed, because of their speed compared to that of, for instance, adaptation to darkness, which can take many minutes to complete. The main topic of this thesis is the dynamics of the continuous adaptation mechanisms, and especially how they work in natural conditions.

HOW CAN WE STUDY LIGHT ADAPTATION?

Electrophysiology and anatomy

The visual system of animals can be studied by using microelectrodes to record the response of single neurons to visual stimuli. In this way we can obtain information about the transformations that the visual input has undergone at a specific cell level in the retina or cortex. The functions of the different types of neurons that are present in the retina and cortex can then be examined. Furthermore, anatomical methods are used to reveal the structure of the visual system, such as what kind of different cell types exist and to which other cells they are connected.

To explain the precise anatomical structure of the retina and what is known about the functions of different neurons is beyond the scope of this thesis. Only an overview of the general structure of the retina will be given here (Kaplan et al., 1990). The first layer consists of photoreceptors: rods and three types of cones with different spectral sensitivities. Here light is transformed into an electrical signal. In intermediate layers we find horizontal, bipolar, and amacrine cells. The final layer leading to the optic nerve consists of ganglion cells. Ganglion cells are connected to the cones via the bipolar cells. Horizontal and amacrine cells provide information from surrounding areas of the retina, and regulate the responses of cones and bipolar cells. The precise role of the various types of horizontal and amacrine cells is not well known.

Functionally, several different parallel information channels can be distinguished in the retina. First, bipolar and ganglion cells are divided into parvocellular (P) and magnocellular (M) neurons, which project to different layers in the lateral geniculate nucleus (LGN). The LGN connects the retina to the visual cortex. In the LGN information from the right and left field of view is combined from the two eyes and sent to the appropriate hemispheres of the visual cortex. The P and M channels have quite different spatial and temporal processing properties (Derrington & Lennie, 1984; Purpura et al., 1990). The P channel is sensitive to higher spatial frequencies, but to lower temporal frequencies than the M channel. The P channel is color sensitive while the M channel is not. Furthermore, the contrast sensitivity of the two channels is quite different. The P channel is thought to be used for detection of fine structure and colour. The M channel is believed to play an important role in motion detection.
Another subdivision of bipolar and ganglion cells is that in ON and OFF cells. ON cells respond to an increase in luminance, OFF cells to a decrease. Finally, P cells can be divided into three different colour channels. The complicated retinal structure makes it very hard to deduce the function of the processes that take place from the responses of individual cells as measured with electrophysiology. It is not only important to know what each cell in the neuronal network does, but even more so how the network functions as a whole. This is where psychophysics can play an important role.

Psychophysics

Due to its invasive nature, it is not customary to study the human visual system with single cell electrophysiology. On the one hand this is a disadvantage, since it is thus not possible to know exactly which transformations take place at which cell levels. On the other hand it has led to the development of a range of alternative techniques measuring the performance of the visual system as a whole, a research field commonly called psychophysics. In psychophysics we seek an objective measure of human performance. For example, when we study light adaptation we would like to have an objective measure of the sensitivity of the visual system to light. To obtain such objective measures a number of psychophysical procedures have been developed. To study light adaptation, we test the ability of the observer to detect a test stimulus presented simultaneously with the adaptation stimulus. During a number of presentations of the adaptation stimulus, the test stimulus is either present or absent. After each presentation, the observer responds whether or not the test stimulus was visible. After repeated stimulus presentations we calculate how often the observer has answered correctly. Next, we change the intensity of the test stimulus. Again, we calculate the percentage of correct answers of the observer. We define the detection threshold for the test stimulus at a certain percentage of correct answers. This threshold is a measure of the sensitivity of the visual system during the presented adaptation stimulus. When the adaptation stimulus changes in time we can follow the time course of the adaptation process by presenting the test stimulus at various times during the adaptation stimulus, and measure the detection thresholds at these times.

Because the detection thresholds provide a quantitative measure of sensitivity at a series of moments during stimulus presentation, we can construct a threshold curve that indicates the change of sensitivity during an adaptation stimulus. This was first done by Crawford (1947), who measured thresholds at large steps in luminance. In Chapter 1 we present psychophysical experiments following the paradigm introduced by Crawford, but with luminance steps smaller than those studied earlier and at higher ambient luminance levels. Furthermore, we do not only study incremental luminance steps, but also decremental steps.

**WHY DO WE NEED A MODEL FOR LIGHT ADAPTATION?**

In Chapter 2 we present a model for light adaptation, based on the results of the psychophysical experiments of Chapter 1, together with new psychophysical experiments with sinusoidally modulated stimuli. The reasons for developing a model are at least twofold. First, it is in general useful to have a model in order to better understand how a system works. It is then possible to make predictions about the response of the
Introduction

system to new inputs. Thus we can design psychophysical experiments that test specific aspects of the model, and investigate whether the model has to be adjusted. An adequate model can, indirectly, provide information about how adaptation is implemented in the visual system. Eventually our aim is to have a model based on anatomical and physiological data that can also describe psychophysical data. Because of the complexity of the system and need for more detailed information on the function of the different cell types it is at this point not yet fully possible to develop such a model. This remains a challenge for the future.

A second reason for developing a model for light adaptation is the possibility of technological applications. For instance, automatic gain control in video cameras can be considered as a form of light adaptation. Presumably, knowledge of human light adaptation can be used to improve the design of such a gain control. Other (future) technical applications could be in robot vision and for artificial retinas.

Contrast gain control

From physiology, it is known that the visual system not only adapts to absolute luminance levels, but also to the amount of fluctuations in luminance, i.e. temporal contrast (Werblin & Copenhagen, 1974; Shapley & Victor, 1978; Kaplan & Shapley, 1986). This contrast gain control has been shown to take place at the level of ganglion cells in the retina (Lee et al., 1994). Benardete and Kaplan (1997, 1999) measured the responses of magnocellular and parvocellular ganglion cells to moving sinusoidal gratings with different modulation depths (or contrast). The responses of the parvocellular ganglion cells increase almost linearly with contrast. However, the responses of magnocellular cells behave quite differently. After a first rapid increase of responses with contrast, responses no longer increase with contrast at higher modulation depths. So the response vs. contrast function is markedly nonlinear.

Furthermore, the temporal contrast sensitivity function (response vs. temporal frequency) of magnocellular cells is different for different contrasts. It is a band-pass function, and its peak shifts to higher frequencies for higher contrasts, with low frequencies more strongly suppressed. Parvocellular cells have a temporal contrast sensitivity function that is more low-pass than band-pass, and it keeps a similar shape at all contrasts, although the absolute response level does change. The dependence of magnocellular cell response on contrast will be referred to in this thesis as contrast gain control. Although there is evidence that further (spatial) contrast adaptation mechanisms are present in the visual cortex (e.g. Carandini et al., 1997), we expect that the contrast mechanisms revealed by our psychophysical experiments have a retinal basis. Contrast gain control is part of the model presented in Chapter 2. Moreover, in Chapter 3 we more extensively study the dynamics of the contrast gain control psychophysically.
HOW DO LIGHT ADAPTATION AND EARLY VISION PERFORM IN A NATURAL ENVIRONMENT?

A theme in vision research which has recently become increasingly popular is to study visual processing of natural stimuli (Field, 1987; Atick & Redlich, 1990; van Hateren, 1997). This interest is fostered by the hypotheses that an animal or a human must survive in its natural environment and that it is, through evolution, optimized for this purpose. The visual system plays an important role in an organism’s ability to survive. It must be able to detect predators or prey, and select appropriate food. A visual system that is better adapted to process natural visual input will increase the chance for survival. It is therefore interesting to study the visual system in its natural environment, and test to what extent it is indeed optimized for this environment. The visual system then receives the input it is built for, instead of artificial stimuli like sinusoidal gratings. It may also be that some mechanisms of visual processing are revealed more clearly with natural stimuli, because these stimuli have a structure that is extremely rich in a statistical sense.

However, it is very difficult to perform controlled experiments in a natural environment. First, the visual surroundings are so complex that it is next to impossible to control or reconstruct the visual input that has reached the eye when an organism behaves naturally in this environment. It is then difficult to correlate the output of experiments with the input. Therefore, two different approaches have been followed in the last two chapters of this thesis that circumvent these difficulties, and that enable us to study how the visual system handles natural input.

In Chapter 4 we focus on stimuli that have the temporal properties of natural visual stimuli, disregarding the spatial properties. We use these stimuli for studying light adaptation under natural luminance conditions. For this purpose light intensities where recorded while walking around in an outdoors environment with a light detector having properties similar to those of a human (para)foveal cone. Although spatial information is lost, the stimulus thus recorded still contains both the rapid dynamics and the large dynamic range of light intensities present in the natural environment. In Chapter 4 we investigate some of the properties and statistics of natural time series of intensities. To study light adaptation to natural luminance conditions we played the recorded natural time series of intensities back on a very bright LED in the laboratory, and performed psychophysical experiments. The results show that the visual system is able to handle these dynamic stimuli quite well.

Second, it is possible to make theoretical predictions about visual processing when it is known what the properties are of the natural visual input, and suitable assumptions are made on physiological constraints and an optimization criterion. For example, it is possible to predict contrast sensitivity functions from the spatiotemporal statistics of natural scenes (van Hateren, 1993). Natural images have a spatial power spectrum that behaves as $1/f^2$, with $f$ spatial frequency (Field, 1987; Ruderman & Bialek, 1994). When we want the transfer of information to be as efficient as possible, redundancy and noise of the input signal have to be reduced. The redundancy of natural images can be
Introduction

reduced by whitening the power spectrum, which means that the low frequencies in the image are suppressed. Furthermore, the high frequencies in the image contain relatively little power and therefore relatively much noise. This means that it is necessary to suppress these high frequencies as well. A combination of these two factors leads to band-pass filtering characteristics for early spatial processing, which are indeed found to be matched quite well with human contrast sensitivity functions.

Of course, these predictions depend both on the physiological constraints and on the criterion that is assumed to be most important for visual processing. The constraints and criterion are presumably different at different stages in the visual system, depending on the visual function at that stage. At the output stage of the retina, efficient information transfer is most important because of the limited information capacity of the optic nerve: it contains a limited number of neurons, with each a limited dynamic range. At a higher stage, the early visual cortex, the information that was packed very densely by the retina, has presumably to be brought into a form suitable for further processing by higher stages of the visual brain. The number of neurons is not as limited here as in the optic nerve. The goal here may be efficient coding, which is to find a code that contains all the information which is available in the optic nerve, but which has coefficient values that are statistically as independent as possible (Bell & Sejnowski, 1997). Such an efficient code has been shown to arise when the criterion of either independent or sparse coding is used (Bell & Sejnowski, 1997; Olshausen & Field, 1996). The underlying assumption is that natural images are built up from only a relatively small number of features. Olshausen and Field (1996) showed that receptive fields similar to those of simple cells in the primary visual cortex emerge for a neural network that implements a sparse code on a set of natural images. This approach has been extended by van Hateren and van der Schaaf (1998). They performed a so-called independent component analysis on a very large set of natural images. This analysis decomposes images into superpositions of components that are as independent as possible by a linear transformation. The amplitude of each component can be extracted from the image with a corresponding filter. Because under natural conditions the visual system does not receive still images, but a full spatiotemporal input, a similar analysis was later performed on spatiotemporal image sequences by van Hateren and Ruderman (1998). The result of this analysis was that natural scenes are composed of components that resemble moving edges or bars. The corresponding filters match the receptive fields of simple cells in the primary visual cortex quite well. They are shaped like moving Gabor patterns, with a fixed envelope, similar to receptive fields measured in cat simple cells by DeAngelis et al. (1993).

In Chapter 5 we investigate the psychophysical consequences of the theory that the spatiotemporal receptive fields of simple cells decompose natural images into small moving components. In the psychophysical experiments, moving Gabors with fixed envelopes are presented to human observers. We show that these patterns produce a strong illusion of visual movement of the envelope, consistent with the independent component interpretation of simple cells in primary visual cortex.
REFERENCES


Introduction


