Network: Comput. Neural Syst. **9** (1998) 39–71. Printed in the UK PII: S0954-898X(98)87748-8

Theory and implementation of infomax filters for the retina

M Haft† and J L van Hemmen‡

† Corporate Technology, Siemens AG, 81730 München, Germany
‡ Physik-Department, Technische Universität München, 85747 Garching bei München, Germany

Received 22 May 1997

Abstract. In the first part of this paper we discuss a technical visual sensory system, which—in analogy with the retina—includes some preprocessing of visual information. In so doing, we use an information-theoretic criterion, the infomax ansatz, to optimize the response of the sensory system. In particular, it is shown that the lattice structure of the photoreceptor array has to be taken into account. By a discrete Fourier transformation on a triangular lattice we derive the frequency response of the infomax filter within the first Brillouin zone. To illustrate the response properties, infomax filters adapted to different noise levels are applied to images with different signal-to-noise ratios. This clearly demonstrates the necessity of adaptation of the filter properties to the given noise level. Furthermore, it is shown how to efficiently implement infomax-like filters by simple networks with only nearest-neighbour interactions. A two-layered network topology proves to be very advantageous in implementing the desired high-pass or low-pass properties. The network topology allows for adaption of the network to low and high noise levels by simply adjusting the nearest-neighbour couplings.

In the second part of this paper, we compare the previously described information-theoretic requirements on a visual sensory system with biological facts known from the vertebrate retina. The substantial physiological response properties of the vertebrate retina are in agreement with the main features of the infomax filter. Since available experimental data lacks information which is necessary for a more quantitative comparison, we present suggestions for future experiments. Some key anatomical features of the retina of many vertebrates compare well with our two-layered implementation of the infomax filter. The analogy, in particular, concerns the adaption mechanism. To illustrate this point, we summarize some recent experiments which demonstrate that in the retina of some species adaption is based on the release of the neuromodulator dopamine by the interplexiform cells. This causes the horizontal cells to decouple. On the basis of recently gained understanding of the outer plexiform layer of the retina some further hypotheses about the functionality of the retina become obvious and possible future experiments to verify or refute them are suggested.

Finally, we discuss the infomax approach from a more general point of view. In particular, we show that redundancy is essential to obtaining noise robustness of an internal representation of the environment as it is produced by a sensory system such as the retina.

1. Introduction

The amount of information which we are permanently receiving through our sensory systems is considerable. In our eyes, for example, there are about 10^8 photoreceptors at work [63], resulting in Giga-bits of visual information per second [56]. The best modern computers would be vastly overloaded if they had to process this huge stream of information.

In biological sensory systems the processing of information starts in the sensory system itself as soon as the data is received by the receptor cells. This is particularly evident in the retina. The phenomenon of Mach bands [46], for example, is commonly attributed

0954-898X/98/010039+33$19.50 © 1998 IOP Publishing Ltd

to retinal information processing. So a first natural question is: what is the purpose of this early-stage preprocessing? In this paper, we will try to give an answer to this question using concepts from information theory. Information-theoretic approaches to sensory information processing have become very popular and are by now nearly a standard technique. Nonetheless, there is still the need for a better understanding, as will be seen in this paper.

In the above sentences, the keyword 'information ' has occurred repeatedly. If we could get a quantitative concept instead of this catchword, we would be able to formulate the requirements for an optimal sensory information processing system. Sensory systems are mapping the environment onto an 'internal representation'. By using an 'information measure' we may decide 'how much information' concerning our environment is reflected by the internal representation. This enables us to assess the performance of a sensory system and to construct an information-theoretically optimal system. An appropriate information measure is given by the Shannon entropy or by the mutual information, which is a mathematical expression borrowed from information theory. A corresponding approach was first formulated by Linsker [41, 43]. It has since become known as the 'infomax ansatz'. Starting with this ansatz we are able to formulate what an optimal visual information processing system should do, and compare the result with the way the retina works.

Our eye can adapt to very different illumination conditions. For example, the intensity of sunlight is 10 billion times that of starlight, yet the eye works under both conditions surprisingly well [24]. That is to say, between the extreme limits of maximal light and dark adaption the eye changes its sensitivity by a factor of 6×10^6 [63]. With respect to this ability, our eye is by far superior to any technical sensor. This is partially due to using two specialized classes of photoreceptor, the rods and the cones. Moreover, in the retina the photoreceptors are embedded in a neural network that serves as an instantaneous preprocessing system. Given the extreme limits of adaption we are concerned with extremely different signal-to-noise ratios. Of course, this fact calls for very different strategies of preprocessing visual information. It is well known that the information processing of the retinal network is highly adaptive to the illumination condition [20]. The aim of this paper is to understand retinal preprocessing from first principles and, in particular, to understand the adaption to *extremely different noise levels*. The method of information maximization offers a suitable framework to develop optimal strategies for an early stage preprocessing of a *noisy* signal. On the one hand, this information-theoretically optimal strategy may be compared with the information processing of the retina, which leads to a better understanding of the retina. On the other hand, we will show how a technical sensory system may be designed in analogy with its biological counterpart. That is, we will design a sensory system with *on-chip* preprocessing.

A brief comparison with companion literature is in order. There is a fair amount of work which supposes an optimization criterion to explain biological visual information processing [23, 45, 52, 57]. Some authors use an approach based on information-theoretic concepts as well [2, 3, 10, 40, 44, 60, 61, 62]. It was Atick [3] who clearly pointed out the close relation between information theory, the statistics of natural images and the information processing function of the retina. A key role has also been played by the principle of redundancy reduction, which can be traced back to Attneave [5] and Barlow [6, 7]. We will discuss the relationship with some other work at the end and, in particular, we will discuss the meaning of redundancy in the noisy case. As far as is known to us, the infomax approach in conjunction with the processing of natural images was first used by van Hateren [60] to describe the properties of the large monopolar cell of the fly. In so doing, he took advantage of numerical Fourier transformations on a rectangular lattice. Some interesting

general considerations with respect to the infomax approach in a linear network without special reference to image processing can be found in [13].

This paper is organized as follows. In section 2 the technical infomax filter is introduced. There we start by discussing some prerequisites. In particular, we review the statistics of natural images and a physically appropriate power constraint which allows us to incorporate our knowledge of a finite internal signal-to-noise ratio in the theory. Via a short mathematical detour through Fourier transformations on a discrete lattice we then derive the infomax filter and the relation between the power constraint and its Lagrange parameter. The infomax filter is discussed both in the frequency domain and in the space domain. We thereby prove that there is unique phase for the infomax filter in the frequency domain if we use 'local receptive fields'. The filter properties are demonstrated through the effect of the filters on an image. Furthermore, it is shown how the infomax-like filters may be implemented by a simple network with nearest-neighbour interactions alone. This network can be easily adapted to the noise level by simply adjusting the nearest-neighbour couplings. Up to here the first part of the paper might be considered independent of biological visual information processing, i.e. simply a discussion of the infomax filter. Nevertheless, we will partially use the intuitive nomenclature of the retina (receptor cells, ganglion cells, receptive fields, etc) so as to avoid repeating ourselves in the rest of this paper. In the next part, section 3, the previously described technical infomax filter is compared with physiological *and anatomical* features of the vertebrate retina. Some contrast sensitivity measurements are compared with the response properties of the infomax filter. It will thereby become evident that further experimental data is desirable. We then compare the network structure of the infomax filter with the topology of the outer plexiform layer of the retina. At this point we also summarize some recent findings on an adaption mechanism observed in the retinas of some species, which is analogous to the modulation of the nearest-neighbour couplings of the previously suggested network. From our understanding of the retina some further hypotheses about the functionality of the retina follow. Finally, in section 4, we discuss the infomax approach from a more general point of view and relate it to Barlow's principle of redundancy reduction. We close in section 5 by discussing the advantages and possible applications of *on-chip* preprocessing of sensory information. We summarize our main results and give an outlook on possible future work.

2. The infomax filter for natural images

2.1. Some prerequisites

The statistics of natural images. A fundamental prerequisite for constructing optimal information processing systems is some knowledge about the nature of the signals which are to be processed. Here some knowledge about the statistics of natural images is required.

Natural images contain a large amount of redundant information [5, 36]. This redundancy becomes evident very impressively in experiments where a person being tested is asked to estimate missing or noisy parts of an image [33]. In general, man is very successful in this task, which is due to the redundancy of visual information. To a large part the redundancy of natural images is caused by the correlation of nearby image elements. Hence, one goal of investigating the statistics of natural images is the estimation of the correlation function.

For the retina a natural image is a two-dimensional distribution of intensities, which we denote by $\xi(\boldsymbol{x})$. The vector $\boldsymbol{x}$ is the position in the two-dimensional focal plane. We denote by $\langle\cdot\rangle$ the expectation value of any expression within the parentheses. Then

$C_\xi(\boldsymbol{x}, \boldsymbol{y}) \equiv \langle \xi(\boldsymbol{x})\xi(\boldsymbol{y})\rangle$ is the correlation function of the ensemble of natural images. Without loss of generality we assume $\langle \xi(\boldsymbol{x})\rangle = 0$. In addition, we may assume that the correlation of image elements is only dependent on their relative separation, i.e. $C_\xi(\boldsymbol{x}, \boldsymbol{y}) = C_\xi(\boldsymbol{x} - \boldsymbol{y})$. The Fourier transform of the correlation function

$$C_\xi(\boldsymbol{k}) \equiv \frac{1}{2\pi} \int_{\mathbb{R}^2} \mathrm{d}^2\boldsymbol{x}\ \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\, C_\xi(\boldsymbol{x}) \tag{2.1}$$

is called the spectrum of natural images. For the moment we disregard any finite-size or discrete-sampling effects.

By some simplifying assumptions it is possible to determine the *coarse* form of the spectrum C_ξ. First, we assume that C_ξ is rotation invariant, $C_\xi(\boldsymbol{k}) = C_\xi(k)$. Second, the spectrum of natural images is assumed to be invariant with respect to any scaling of the x-axis. A scene on which we focus is scaled by a factor $1/s$ if we change our distance by a factor s. Scaling the space domain by a factor $1/s$ in turn results in scaling the frequency domain by a factor s. This, as well as experimental findings [10, 23], is the motivation for the assumption of scale invariance. Hence, let us take a look at the consequences. We shall see that the only consistent spectrum is $C_\xi(k) \propto 1/k^2$.

Scale invariance means that $C_\xi(sk) \sim C_\xi(k)$ for any factor $s > 0$. Hence we assume

$$C_\xi(sk) = c(s)C_\xi(k) \tag{2.2}$$

where $c(s)$ is to be determined. With this aim in mind we express the variance of natural images at an arbitrary position $\boldsymbol{x}$ in the focal plane by the spectrum $C_\xi(k)$:

$$\begin{aligned}\left\langle \xi(\boldsymbol{x})^2 \right\rangle &= C_\xi(\boldsymbol{0}) \\ &= \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \mathrm{d}^2\boldsymbol{k}\ C_\xi(k) \\ &= s^2 \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \mathrm{d}^2\boldsymbol{q}\ C_\xi(sq) \\ &= s^2 c(s) \left\langle \xi(\boldsymbol{x})^2 \right\rangle . \end{aligned} \tag{2.3}$$

In obtaining the third equality we made the variable transformation $\boldsymbol{k} = s\boldsymbol{q}$ and then used the scale invariance (2.2). From (2.3) we conclude that $c(s) = s^{-2}$, and therefore that

$$C_\xi(sk) = \frac{1}{s^2} C_\xi(k)\ . \tag{2.4}$$

This condition uniquely determines the functional form of the spectrum. We have

$$C_\xi(k) = \frac{g}{k^2} \tag{2.5}$$

with an arbitrary global constant g. Some work [10, 23] has pointed out the scale invariance of natural images, and experimental statistics confirms equation (2.5) by and large for the spectrum of natural images [9, 19, 53, 54]. In [54], for example, the experimental exponent for images in the woods is determined to be 1.81 ± 0.01.

The derivation of equation (2.5) has one weak spot. We end up with (2.5), but because of the second-order pole this expression leads to infinite variance, $\langle \xi(\boldsymbol{x})^2 \rangle = \infty$. Hence equation (2.3) reads $\infty = \infty s^2 c(s)$, which does not admit the conclusion $c(s) \sim s^{-2}$. Physically, infinite variance is not realistic. If necessary we will shift the pole from the real axis by a small amount κ and set

$$C_\xi(k) = \frac{g}{\kappa^2 + k^2}. \tag{2.6}$$

This corresponds to two first-order imaginary poles. The spectrum (2.5) and (2.6) will be the basis of all further investigations.

Perception of visual information by a sensory system starts with sampling the image $\xi(\boldsymbol{x})$ by a photoreceptor array $\mathcal{L}$. We denote the photoreceptor signals by the vector $\boldsymbol{\xi} = (\xi(\boldsymbol{x}), \boldsymbol{x} \in \mathcal{L}) \in \mathbb{R}^N$ and discuss the consequences of this sampling later on (Brillouin zone, Nyquist frequency, aliasing, etc). The correlations of the photoreceptors are now described by the correlation matrix $\mathbf{C}_\xi = \langle \boldsymbol{\xi}\boldsymbol{\xi}^{\mathrm{T}} \rangle$. Not that much is known about higher-order moments. We will therefore estimate the the probability distribution $P(\boldsymbol{\xi})$ of the sampled natural images on the basis of the correlation matrix $\mathbf{C}_\xi$. That is, we assume that $P(\boldsymbol{\xi})$ is a Gaussian distribution (maximum entropy!) with correlation matrix $\mathbf{C}_\xi$:

$$P(\boldsymbol{\xi}) = \frac{1}{\sqrt{2\pi}^N} \frac{1}{\sqrt{\det \mathbf{C}_\xi}} \exp\left[-\frac{1}{2} \boldsymbol{\xi}^{\mathrm{T}} \mathbf{C}_\xi^{-1} \boldsymbol{\xi} \right]. \tag{2.7}$$

The relation of the matrix $\mathbf{C}_\xi$ to the spectrum of the natural images $C_\xi(k)$ will become evident later on. On the basis of the statistics (2.7) we may now construct a corresponding optimal sensory information processing system.

The model. Photoreceptors transform visual information into electrical signals. The photoreceptor signals are processed by a network of different cells. At the end, ganglion cells form the output $\boldsymbol{\sigma}$ of the retina. For the sake of simplicity, we assume that the ganglion cells are arranged in the same lattice as the receptor cells so that $\boldsymbol{\sigma} \in \mathbb{R}^N$ as well. In the fovea this assumption is valid, as here the number of ganglion cells is comparable to the number of cones. Furthermore, we assume that the output $\boldsymbol{\sigma}$ depends linearly on the input $\boldsymbol{\xi}$. This makes sense for the main class of ganglion cells, i.e. for the visual P-system [18, 29] and for the X-ganglion cells of the cat [21]. Finally, we assume that the input $\boldsymbol{\xi}$ as well as the output $\boldsymbol{\sigma}$ is corrupted by Gaussian noise $\boldsymbol{\nu}$ and $\boldsymbol{\mu}$ with variances $\langle \boldsymbol{\nu}\boldsymbol{\nu}^{\mathrm{T}} \rangle = \Delta \mathbb{1}$ and $\langle \boldsymbol{\mu}\boldsymbol{\mu}^{\mathrm{T}} \rangle = \delta \mathbb{1}$, respectively. The significance of noise throughout the whole retina has been discussed elsewhere [16, 20, 51, 56, 59]. The variance δ describes to what extent different output signals may have a different meaning. In other words, δ also has to reflect the accuracy of later internal information processing. The assumption $\delta = 0$ is absolutely unphysical; rather it is reasonable to assume a large amount of internal noise [58]. Altogether, we obtain the following model:

$$\boldsymbol{\sigma} = \mathbf{G}\,(\boldsymbol{\xi} + \boldsymbol{\nu}) + \boldsymbol{\mu} \tag{2.8}$$

for some linear transformation $\mathbf{G}$. According to the noise terms $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$, the output variable $\boldsymbol{\sigma}$ is not a deterministic function of the input variable $\boldsymbol{\xi}$. Instead their mutual dependence is given by the conditional distribution $P(\boldsymbol{\sigma}|\boldsymbol{\xi})$ or by $P(\boldsymbol{\xi}|\boldsymbol{\sigma})$.

In the first part of this paper we are not interested in answering the question of how the network of the retina or a technical sensor manages to represent the linear mapping $\mathbf{G}$. For the moment the sensory system is a black box whose linear response properties are described by the matrix $\mathbf{G}$. The goal we are aiming at is investigating requirements on $\mathbf{G}$, which will follow from an infomax approach.

The infomax ansatz. The task of a sensory system is to map the environment onto an internal representation. The input variable $\boldsymbol{\xi}$ is also called the *external* variable or simply the *environment*. The output variable $\boldsymbol{\sigma}$ is also called the *internal* variable or the *internal representation of the environment*. Of course, the internal representation should reflect as much 'information' about the environment as possible. So far, however, this is just a colloquial formulation and to quantify this we need a so-called information measure.

It is beyond the scope of this paper to discuss information measures. The approach used here is based on the Shannon entropy [14]. The entropy of a random variable is interpreted as a measure of the uncertainty about this variable. We do not want to be uncertain about our environment. Hence we require that, given the internal representation $\boldsymbol{\sigma}$, the entropy of the external variable $\boldsymbol{\xi}$

$$\begin{aligned} H[\boldsymbol{\xi}|\boldsymbol{\sigma}] \;&\equiv\; -\int \mathrm{d}^N\boldsymbol{\xi} \int \mathrm{d}^N\boldsymbol{\sigma}\; P(\boldsymbol{\xi},\boldsymbol{\sigma}) \log P(\boldsymbol{\xi}|\boldsymbol{\sigma}) \\ &= -\langle \log P(\boldsymbol{\xi}|\boldsymbol{\sigma})\rangle \end{aligned} \tag{2.9}$$

be small. Here $P(\boldsymbol{\xi}|\boldsymbol{\sigma})$ is the probability density of $\boldsymbol{\xi}$ given $\boldsymbol{\sigma}$, and we have implicitly assumed that all probability measures are absolutely continuous with respect to Lesbeque measure. The entropy $H[\boldsymbol{\xi}|\boldsymbol{\sigma}]$ can be used as a measure for the quality of a sensory system. For reasons of comparison we subtract the entropy $H[\boldsymbol{\xi}] = -\langle \log P(\boldsymbol{\xi})\rangle$, which is a constant, not depending on the properties of the sensory system. By an extra change of sign and exploiting the Bayes relation $P(\boldsymbol{\xi}|\boldsymbol{\sigma}) = P(\boldsymbol{\xi},\boldsymbol{\sigma})/P(\boldsymbol{\sigma})$ we obtain the mutual information

$$M[\boldsymbol{\sigma};\boldsymbol{\xi}] = \left\langle \log \frac{P(\boldsymbol{\xi},\boldsymbol{\sigma})}{P(\boldsymbol{\xi})P(\boldsymbol{\sigma})} \right\rangle \tag{2.10}$$

which we will use to assess the performance of a sensory system with linear map $\mathbf{G}$. This approach is known as the infomax ansatz, which dates back to Linsker [41, 43].

At this point we have to comment on the concept of the entropy of a continuous variable. Indeed, a solid mathematical foundation of an information measure, e.g., by the above-mentioned Khinchin axioms, is only possible for random variables with a finite number of discrete states. The straightforward generalization (replacing sums by integrals etc) results in a concept which is not without problems. We have to take care of two things. First, the set of real numbers has uncountably many elements, which corresponds to an *uncountably infinite alphabet* in our case. This concept is not within the framework of conventional information theory [1, 14], and, of course, may lead to divergent expressions. The assumption of noise for every random variable leads to a finite resolution of the real axis, and, correspondingly, any entropy becomes finite as well. Finite noise, however small, is absolutely different from zero noise. Finite noise corresponds to a finite number of discriminable states, while zero noise is an unphysical concept corresponding to an infinite number of discriminable states. Second, we may investigate the limits of vanishing noise at the end, e.g., the limit $\Delta \to 0$ or the limit $\delta \to 0$. However, we have to take care that the physical contents which we are aiming at are not lost in this limit. In our case it turns out that the limits $\Delta \to 0$ and $\delta \to 0$, as well as the limits $\Delta \to \infty$ and $\delta \to 0$, do *not* commute, and the different results correspond to different physical assumptions, namely $\delta \ll \Delta$ (coding) or $\delta \gg \Delta$ (noisy information processing). Results which are derived by setting $\delta = 0$ *a priori* must in general be treated with care.

The power constraint. Without any constraints the optimal sensory mapping $\mathbf{G}$ has an infinite determinant, because an infinite determinant results in an infinite signal-to-noise ratio $\langle \boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{\sigma}\rangle/\delta$ at the output. Physically we observe a finite variance or a finite 'number of states' $(\langle \boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{\sigma}\rangle/\delta)^{1/2}$ available to the output. We add this physical constraint to our optimization criterion (2.10) by using a Lagrange parameter λ. The cost function is then given by

$$L = M[\boldsymbol{\sigma};\boldsymbol{\xi}] - \frac{\lambda}{2}\frac{p}{\delta} \qquad \text{where} \quad p \equiv \mathrm{Trace}\left\langle \boldsymbol{\sigma}\boldsymbol{\sigma}^{\mathrm{T}}\right\rangle. \tag{2.11}$$

This is essentially the 'power constraint' which is used by van Hatereen [60, 61] as well as in most technical considerations. Other constraints may be used, but the constraint cannot be chosen arbitrarily. It has to reflect the physics of the system which we want to describe and it has to be experimentally accessible, if we finally want to compare our results with those of a biological system. The variance p, or better the signal-to-noise ratio p/δ, meets these demands.

2.2. Theory of the infomax filter

We derive the infomax filter only briefly. Derivations like this one have become some kind of a standard approach [3, 10, 40, 44, 60]. Here, however, we focus on a rigorous treatment of the foveal lattice (or the lattice of a technical sensor array). This proves to be very important, in particular, if we want to compare the result with physiological experiments.

We have assumed the Gaussian model (2.7) as the distribution of natural images and Gaussian white noise. Hence the marginal distribution of the internal variable $P(\boldsymbol{\sigma})$ and the conditional distribution $P(\boldsymbol{\sigma}|\boldsymbol{\xi})$ are also Gaussians with covariance matrices

$$\mathbf{C}_\sigma \equiv \left\langle \boldsymbol{\sigma}\boldsymbol{\sigma}^{\mathrm{T}} \right\rangle = \mathbf{G}\mathbf{C}_\xi \mathbf{G}^{\mathrm{T}} + \Delta \mathbf{G}\mathbf{G}^{\mathrm{T}} + \delta \mathbb{1} \tag{2.12}$$

and

$$\mathbf{C}_{\sigma|\xi} \equiv \left\langle \boldsymbol{\sigma}\boldsymbol{\sigma}^{\mathrm{T}}|\boldsymbol{\xi} \right\rangle - \langle \boldsymbol{\sigma}|\boldsymbol{\xi}\rangle \left\langle \boldsymbol{\sigma}^{\mathrm{T}}|\boldsymbol{\xi} \right\rangle = \Delta \mathbf{G}\mathbf{G}^{\mathrm{T}} + \delta \mathbb{1}. \tag{2.13}$$

Here the expression $\langle \cdots | \boldsymbol{\xi} \rangle$ denotes a conditional expectation given $\boldsymbol{\xi}$; it is also Gaussian and has the covariance matrix $\mathbf{C}_{\sigma|\xi}$. The entropy H of a multivariate Gaussian distribution with covariance matrix $\mathbf{C}$ is given by $H = \frac{1}{2}\log(\det \mathbf{C}) + \frac{1}{2} N \log(2\pi \mathrm{e})$. For the cost function (2.11) we then obtain

$$L = \frac{1}{2} \log \frac{\det \mathbf{C}_\sigma}{\det \mathbf{C}_{\sigma|\xi}} - \frac{\lambda}{2} \mathrm{Trace}\left(\mathbf{C}_\sigma\right) \tag{2.14}$$

with $\mathbf{C}_\sigma$ and $\det \mathbf{C}_{\sigma|\xi}$ given by (2.12) and (2.13).

We have already discussed the statistics of natural images in terms of the assumption of translation invariance, namely $\mathbf{C}_{\xi,\boldsymbol{xy}} = C_\xi(\boldsymbol{x} - \boldsymbol{y})$. We will also assume this symmetry property for the structure of the retina, i.e. we will assume $\mathbf{G}_{\boldsymbol{xy}} = G(\boldsymbol{x} - \boldsymbol{y})$. This means that the coupling $\mathbf{G}_{\boldsymbol{xy}}$ of an input neuron at the position $\boldsymbol{y} \in \mathcal{L}$ to an output neuron at the position $\boldsymbol{x} \in \mathcal{L}$ is only dependent on the relative separation of these two neurons in the retina. Because of translation invariance and, additionally, assuming periodic or infinite boundary conditions we can diagonalize $\mathbf{G}$ and $\mathbf{C}_\xi$ simultaneously by a discrete Fourier transformation. For example, for the matrix $\mathbf{G}$ this discrete Fourier transformation reads

$$\mathbf{M}^* \mathbf{G}\, \mathbf{M} = \mathrm{diag}\,[G(\boldsymbol{k})] \qquad \mathbf{M}_{\boldsymbol{kx}} \sim \exp\left(-\mathrm{i}\boldsymbol{x}\cdot\boldsymbol{k}\right). \tag{2.15}$$

The vector $\boldsymbol{x}$ is a vector on the chosen lattice structure $\mathcal{L}$, and the frequency vector $\boldsymbol{k}$ has to be defined on the corresponding reciprocal lattice [35]. Using the coordinate transformation (2.15), we rewrite the cost function (2.14) as

$$L = \frac{1}{2} \sum_{\boldsymbol{k}} \log \frac{|G(\boldsymbol{k})|^2 C_\xi(\boldsymbol{k}) + |G(\boldsymbol{k})|^2 \Delta + \delta}{|G(\boldsymbol{k})|^2 \Delta + \delta} - \frac{\lambda}{2\delta} \sum_{\boldsymbol{k}} \left(|G(\boldsymbol{k})|^2 C_\xi(\boldsymbol{k}) + |G(\boldsymbol{k})|^2 \Delta + \delta \right) \tag{2.16}$$

where $\boldsymbol{k}$ runs through the reciprocal lattice.

Using the cost function (2.16), we determine the 'optimal' $|G(\boldsymbol{k})|^2$, taking into account the boundary condition $|G(\boldsymbol{k})|^2 > 0$. The standard optimization procedure leads to the result

$$2\frac{\Delta}{\delta}|G(\boldsymbol{k})|^2 = \frac{1}{1+\Delta/C_\xi(\boldsymbol{k})}\left[\left(1+\frac{4}{\lambda}\frac{\Delta}{C_\xi(\boldsymbol{k})}\right)^{\frac{1}{2}}+1\right]-2 \tag{2.17}$$

or $|G(\boldsymbol{k})|^2 = 0$ if the right-hand side is less than zero. This is the Kuhn–Tucker condition of an optimization task with the constraint given as an inequality [12]. Expressions which are more-or-less equivalent to (2.17) can be found in [44, 52, 60]. The filter gain (2.17) is non-zero, if

$$\frac{C_\xi(\boldsymbol{k})}{\Delta} > \frac{\lambda}{1-\lambda}. \tag{2.18}$$

This is a condition on the signal-to-noise ratio $\mathrm{SNR}_\xi(\boldsymbol{k}) \equiv C_\xi(\boldsymbol{k})/\Delta$ of the external signal $\boldsymbol{\xi}$ corresponding to the frequency component $\boldsymbol{k}$ and means that $|G(\boldsymbol{k})|^2 = 0$, if there is too much noise in the corresponding frequency component. Because $G(\boldsymbol{k})$ corresponds to the eigenvalue of the linear mapping $\mathbf{G}$, it may well be that the optimal mapping is not invertible. In other words, for large noise the weight vectors of the output neurons become linearly dependent. Linsker noted a corresponding property when he originally stated the infomax approach [41, 42].

The *global* signal-to-noise ratio at the input is given by $\mathrm{SNR}_\xi = \sum_{\boldsymbol{k}} \mathrm{SNR}_\xi(\boldsymbol{k})$, which is proportional to g/Δ if we assume the spectrum (2.5) or (2.6). Hence we will also use the ratio g/Δ to characterize the signal-to-noise ratio SNR_ξ at the input. The SNR_ξ will be mainly determined by the illumination intensity. We will therefore associate low SNR_ξ with 'vision at night'†.

For mathematical convenience we now introduce an additional assumption, namely, that the photoreceptor array has an infinite extension. This assumption is valid, if the extent of the 'receptive fields' investigated here is much smaller than the extent of the photoreceptor array. This will be confirmed at the end. In the retina the density of the photoreceptors is decaying rapidly with increasing distance from the fovea [29, 63]. Within the fovea, however, the density of the photoreceptors is almost constant. This justifies the assumption (at least as an approximation) of an 'infinite fovea' with single lattice constant. In this limit $\boldsymbol{k}$ becomes a continuous variable, which is defined in the first Brillouin zone.

The Brillouin zone of a square lattice with lattice constant a is the square $[-\pi/a, \pi/a]\times[-\pi/a, \pi/a]$. The frequency π/a corresponds to what is known as Nyquist frequency of one-dimensional discrete Fourier transformations. For the retina, however, the assumption of a square lattice is not adequate. In the fovea the cones form a fairly regular triangular lattice [63], which corresponds to the densest packaging in two dimensions. This lattice and the corresponding Brillouin zone, which is a hexagon, are shown in figure 1. The hexagon is nearly a disc. Because any highly regular lattice structure breaks rotational symmetry, which is of advantage when manipulating mathematical expressions, we replace the hexagon by a disc of equal volume. The radius of this disc is the frequency k_B, which is given by

$$k_B = 12^{\frac{1}{4}}\pi^{\frac{1}{2}}\frac{1}{a} \approx \frac{\pi}{a}. \tag{2.19}$$

The assumption of a rotation invariant Brillouin zone may also be an adequate strategy in the case of a 'random lattice'. At least in the parafoveal region of the retina the receptors seem to be distributed in a partially random manner [63, 66]. In general, the lattice structure

† Note that for very low illumination conditions the noise is Poisson and not Gaussian [16].

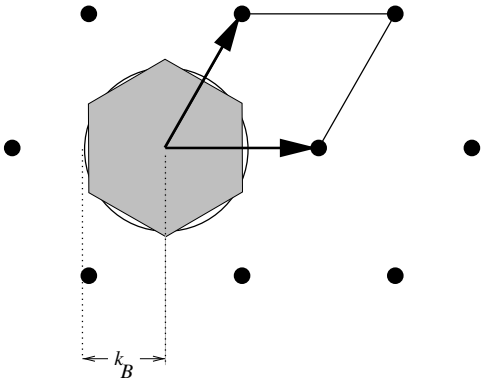


Figure 1. The triangular lattice and the corresponding Brillouin zone as found, for instance, in the fovea; cf [63, figure 3.4a]. Although the elementary cell of this lattice is a parallelogram, this lattice is commonly termed triangular. The two arrows show one possible choice of primitive translations. The Brillouin zone of this lattice is the shaded hexagon. To restore rotational symmetry, the hexagonal Brillouin zone is approximated by a disc of equal volume. The radius of the disc is the frequency k_B, which may by viewed as the analogue of the Nyquist frequency.

may not be disregarded. That is, we are not allowed to assume $a = 0$ or $k_B = \infty$, because it will finally turn out that the extent of the receptive fields is of the order of the lattice constant a.

Due to the assumption of an infinite lattice, the Fourier transforms are given by

$$G(\boldsymbol{k}) = \sum_{\boldsymbol{x}\in\mathcal{L}} \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} G(\boldsymbol{x}) \qquad \text{and} \qquad G(\boldsymbol{x}) = \frac{1}{V_B} \iint_B \mathrm{d}^2\boldsymbol{k}\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\, G(\boldsymbol{k}). \tag{2.20}$$

Here B denotes the Brillouin zone and V_B is its volume. Analogous expressions apply to C_ξ. Later on we will assume the spectrum (2.5) or (2.6) for $C_\xi(\boldsymbol{k})$ on the first Brillouin zone, i.e. we do not take into account any aliasing effects. Aliasing means that the discrete sampling by the photoreceptor array does not discriminate between frequencies which differ by a vector of magnitude $2k_B$ along any lattice direction and, hence, the higher Brillouin zones become superimposed on the first Brillouin zone. For the moment it is sensible to disregard aliasing for the following reasons. First, because of the $1/k^2$ dependence of the spectrum the main contribution always comes from the *first* Brillouin zone. Second, as we will see later on, when we introduce a structure function to take into account the finite extent of the cones and the point spread function of the optical apparatus [63], the system is practically insensitive to frequencies of the higher Brillouin zones, above $k_B \sim 1/a$. Finally, an aliasing argument does not apply to noise as far as the noise emerges *a priori* at the discrete lattice sites.

2.3. Discussion and demonstration of the filter properties

For a discussion of infomax filters and filters with similar characteristics, we refer the reader to [60, 61] and also to [3, 4].

Discussion in the frequency domain. For low input noise (small Δ) the infomax filter (2.17) is approximately given by

$$\frac{\lambda}{\delta}|G(\boldsymbol{k})|^2 \approx \frac{1}{C_\xi(\boldsymbol{k}) + \Delta}\left[1 - \lambda - \frac{2}{\lambda}\frac{\Delta}{C_\xi(\boldsymbol{k})}\right] \tag{2.21}$$

which follows by expanding the square root in (2.17). For vanishing input noise $\Delta = 0$ we finally obtain

$$|G(\boldsymbol{k})|^2 \sim \frac{1}{C_\xi(\boldsymbol{k})}. \tag{2.22}$$

This result corresponds to 'whitening' the input signal, i.e. $C_\sigma(\boldsymbol{k})$ is constant. To obtain a flat output spectrum, a large filter gain $|G(\boldsymbol{k})|$ is needed for frequencies with small $C_\xi(\boldsymbol{k})$. In general, $C_\xi(\boldsymbol{k})$ decreases monotonically with increasing $k = \|\boldsymbol{k}\|$, and, hence, $G(\boldsymbol{k})$ is high-pass.

If there is finite receptor noise $\Delta > 0$, this whitening results mainly in boosting the noise in the high-frequency components of an image. In equation (2.21) this is taken into account to a first approximation by the term $-2\Delta/[\lambda C_\xi(\boldsymbol{k})]$, which mainly damps frequencies with low $\mathrm{SNR}_\xi(\boldsymbol{k}) \equiv C_\xi(\boldsymbol{k})/\Delta$, i.e. high frequencies. With increasing receptor noise Δ or decreasing signal amplitude, the filter properties gradually change from high-pass to low-pass. Finally, high frequencies may even be cut off by the Kuhn–Tucker condition (2.18).

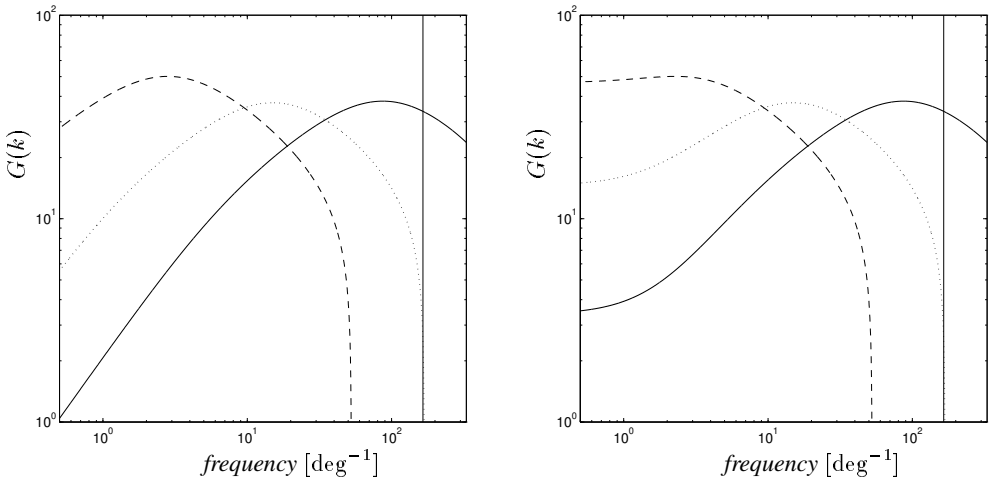


Figure 2. Some infomax filters for large, medium, and small signal-to-noise ratio SNR_ξ. Full line, $g/\Delta = 3.3$, $(p-\delta)/\delta = 10.0$; dotted line, $g/\Delta = 0.1$, $(p-\delta)/\delta = 1.0$; broken line, $g/\Delta = 0.0033$, $(p-\delta)/\delta = 0.1$. The internal noise contributes to the total variance p of the internal signal by an amount δ. Hence it is adequate to characterize the internal noise by the parameter $(p-\delta)/\delta$. The relation between p and λ is evaluated in appendix A.1. The filters in (*a*) correspond to the spectrum (2.5), those in (*b*) to the spectrum (2.6). The lattice constant is $a = 0.02^\circ$ and the corresponding k_B is marked by the vertical line. In (*a*), we have chosen $\kappa = 10^{-2}k_B$.

Some filters are plotted in figure 2. The filters in figure 2(*a*) correspond to the spectrum (2.5), those in 2(*b*) to the spectrum (2.6). The vertical line shows the frequency k_B corresponding to the lattice constant $a = 0.02^\circ$. In figure 2(*a*) the position of this line is somewhat arbitrary, as a rescaling of the frequency axis may be absorbed into the unknown constant g of the spectrum (2.5). This is simply the scale invariance of the natural images discussed previously so that the ratio g/Δ cannot be fixed. Only a finite $\kappa > 0$ fixes the scale of natural images in relation to the lattice constant. For this κ it is adequate to assume at least

$$\kappa \approx \frac{2\pi}{\text{extent of the fovea}}. \tag{2.23}$$

For the infinite lattice, the possible frequencies $\boldsymbol{k}$ are dense in the Brillouin zone and the sum over the reciprocal lattice becomes an integral over the Brillouin zone. Because of the finite extent of the fovea the smallest frequency of physical meaning on the foveal lattice is

given by (2.23). By choosing κ according to (2.23), we have changed the original spectrum essentially only below this frequency. The fovea extends to about 100 lattice constants a only, so that $\kappa = 10^{-2}$ is an adequate assumption.

Back to the space domain. The cost function (2.16) depends only on $|G(\boldsymbol{k})|^2$, and, hence, the infomax ansatz does not determine the phase of the filters. Given any $G(\boldsymbol{k})$, we may form a related filter $G(\boldsymbol{k})\mathrm{e}^{\mathrm{i}f(\boldsymbol{k})}$ with an arbitrary function $f(\boldsymbol{k}) \in \mathbb{R}$, which always results in the same costs (2.16). The only condition on $f(\boldsymbol{k})$ is $\mathrm{e}^{\mathrm{i}f(\boldsymbol{k})} = \mathrm{e}^{-\mathrm{i}f(-\boldsymbol{k})}$, because $G(\boldsymbol{x})$ has to be real, i.e. $G(\boldsymbol{x}) \in \mathbb{R}$. Demanding rotational symmetry leads to the additional condition $\mathrm{e}^{\mathrm{i}f(\boldsymbol{k})} = \mathrm{e}^{\mathrm{i}f(-\boldsymbol{k})}$. This additional condition leaves us only with two possible choices $\mathrm{e}^{\mathrm{i}f(\boldsymbol{k})} = \pm 1$.

In what follows we will always choose $G(\boldsymbol{k}) = |G(\boldsymbol{k})|$. This choice of the phase is distinguished by the additional property that the corresponding 'receptive fields' are 'local and topology preserving'. Only through this property the term receptive field is appropriate since a 'receptive field' is defined as the retinal area of receptors that provide input to a ganglion cell. In appendix A.2 we show that, without rotation symmetry, the two possibilities for the phases, $\mathrm{e}^{\mathrm{i}f(\boldsymbol{k})} = \pm 1$, follow from requiring the receptive fields to be local. In so doing we use a cost function which enables us to assess the locality of a phase $f(\boldsymbol{k})$, given a fixed $|G(\boldsymbol{k})|$.

Having determined the phase of the filters, we may now use the inverse Fourier transform to calculate the receptive field $G(\boldsymbol{x})$. The infomax filter (2.17) depends on $\boldsymbol{k}$ only via the signal-to-noise ratio $\mathrm{SNR}_\xi = C_\xi(\boldsymbol{k})/\Delta$. We suppose the spectrum (2.5), i.e. $C_\xi(\boldsymbol{k}) = C_\xi(k) = g/k^2$. In this case the infomax filter may be viewed as a function of $k^2\Delta/g$ only, so that $G(k) = H(k^2\Delta/g)$. The inverse Fourier transform therefore reads

$$\begin{aligned} G(x) &\sim \int_0^{k_B} k\mathrm{d}k \int_0^{2\pi} \mathrm{d}\varphi\ \mathrm{e}^{-\mathrm{i}kx\cos(\varphi)}\ H\left(k^2\Delta/g\right) \\ &\sim \int_0^{k_B\sqrt{\Delta/g}} q\mathrm{d}q\ J_0\left(q\frac{x}{\sqrt{\Delta/g}}\right) H\left(q^2\right) \end{aligned} \tag{2.24}$$

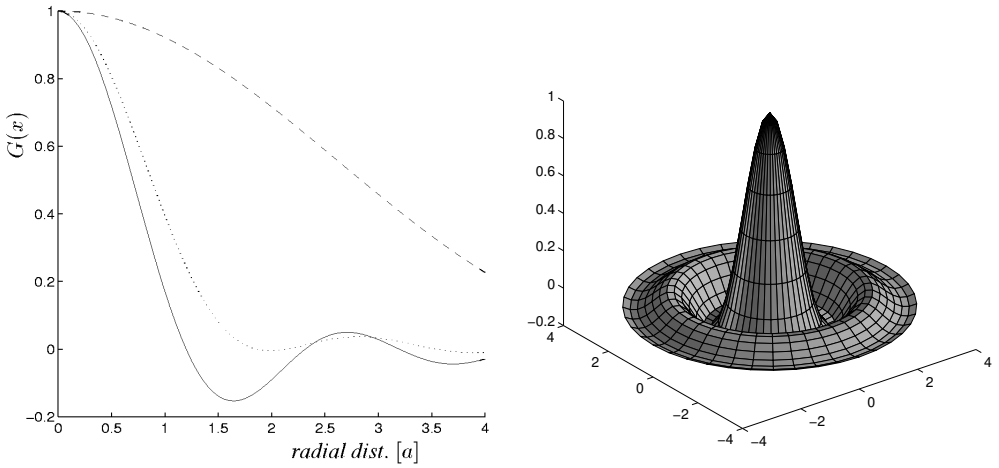
where we have omitted irrelevant global constants. In the second line we have switched to the integration variable $q = k\sqrt{\Delta/g}$ and introduced the zero-order Bessel function

$$J_0(x) = \frac{1}{\pi}\int_0^{\pi} \mathrm{d}\varphi\ \mathrm{e}^{-\mathrm{i}kx\cos\varphi}. \tag{2.25}$$

Integral transformations with Bessel functions as their kernels are called Hankel transformations. They are difficult to perform analytically [64]. Furthermore, the integral in (2.24) has the finite upper bound $k_B\sqrt{\Delta/g}$ so that we are bound to evaluate the integral over q in (2.24) explicitly. Figure 3 shows some results for $G(x)$, which follow from a numerical evaluation of the integral over q in (2.24). For large SNR_ξ we find the typical 'centre-surround receptive fields', which are known from many experiments on the retina. The illumination of the surroundings of an output cell has an inhibitory effect on that cell. This inhibitory effect decreases with increasing noise Δ. Finally, the extent of the optimal receptive fields increases as Δ increases.

The size increase of the receptive fields can be proven explicitly as soon as the Kuhn–Tucker condition (2.18) is not met in the first Brillouin zone. In this case the effective upper bound of the integral (2.24) is not given by $k_B\sqrt{\Delta/g}$ but rather by the value q_0 with $H(q_0^2) = 0$. Hence $G(x)$ may be written as

$$G(x\sqrt{\Delta/g}) \sim \int_0^{q_0} q\mathrm{d}q\ J_0(qx)\, H(q^2). \tag{2.26}$$



(*a*) (*b*)

Figure 3. (*a*) Local infomax filters $G(x)$ according to the parameters of figure 2(*a*). As SNR_ξ decreases, first the inhibitory surround disappears, then the size of the receptive fields increases. To enable better comparison, any filter has been scaled so as to give $G(x = 0) = 1$. The units of the x-axis are given in terms of the lattice constant a. Strictly speaking, the function $G(x)$ is only defined for discrete values. Nearest neighbours have a distance a. On a triangular lattice, however, the next-nearest-neighbour lattice points are at a distance of $\sqrt{3}a$. For this, and because in the retina the lattice will not be absolutely regular, we have decided to draw a continuous line. (*b*) Three-dimensional plot corresponding to the filter for large SNR_ξ in (*a*), namely that represented by the solid line.

We have shifted the factor $\sqrt{\Delta/g}$ into the argument of G so as to show that the right-hand side is now independent of Δ/g for a fixed Lagrange parameter λ. That is to say, increasing Δ/g is now just a rescaling of the argument on the left-hand side, which corresponds to a form-invariant increase of the receptive field.

The fraction Δ/g is proportional to the signal-to-noise ratio. Equation (2.26) is then in agreement with experiments on the retina where a decrease of the inhibitory surroundings [8] and a subsequent increase of the size of the receptive fields [30] has been observed as the average illumination decreases and, hence, Δ/g increases.

Demonstration of the filter properties. The effect of the infomax filters is demonstrated in figure 4. However, one must be careful. Simply because an image is of subjectively poor quality this does not mean that it has no informative value. The following illustration, however, shows that our subjective view is consistent with the 'objective' infomax approach.

The image in figure 4(*a*) is the original image; in 4(*b*) white noise has been added to the gray values of 4(*a*) according to a signal-to-noise ratio of $\mathrm{SNR}_\xi = \langle \xi^2 \rangle / \langle \nu^2 \rangle = 1.0$. The images 4(*c*) and 4(*d*) are results that follow by applying an infomax filter appropriate for a large $\mathrm{SNR}_\xi \gg 1$ to the images 4(*a*) and 4(*b*). The image in 4(*c*) shows that, compared with the original non-noisy image, the contrast is enhanced, which is due to the centre-surround structure of the receptive fields. This is evident in particular at the inserted black and white squares. At the boundary of these squares we find the so-called 'Mach band phenomenon' [20, 46]. Fine details, e.g., the drops of water on the petals, stand out more clearly.

Applying the same filter to the noisy image 4(*b*) leads to the catastrophic result 4(*d*). The fine details of the original image 4(*a*) are already lost in the noisy image 4(*b*): in 4(*a*)

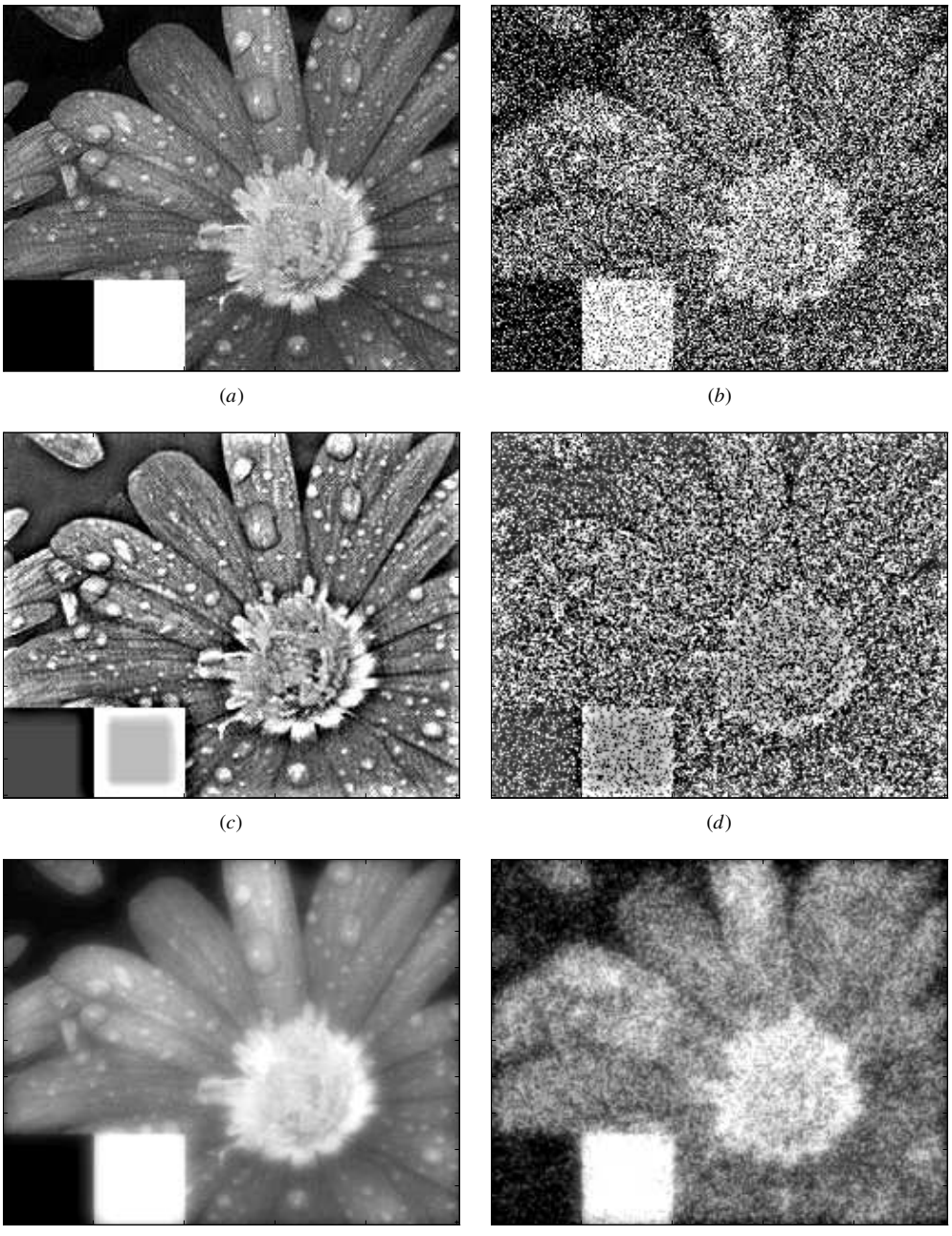


Figure 4. Demonstration of the filter properties. Image (a) is the original image; in (b) white Gaussian noise with $\mathrm{SNR}_\xi = 1.0$ has been added. The images (c) and (d) result by applying an infomax filter adapted to a low noise level, $g/\Delta = 0.0003$, $(p-\delta)/\delta = 3.0$, to the images (a) and (b), respectively. The images (e) and (f) follow by applying a filter which is adapted to $\mathrm{SNR}_\xi = 1.0$, i.e. $g/\Delta = 0.2$, $(p-\delta)/\delta = 3.0$. The results clearly demonstrate that it is crucial to accurately adapt the filter properties to the amount of internal and external noise. Otherwise the consequence will be the catastrophic result (d) or a loss of contrast (e).

we see the drops, in 4(b) we do not. Hence there is no use in trying to enhance the fine structure of a noisy image. This just enhances the noise, which is white and hence 'fine structured'. The image in 4(d) impressively demonstrates the necessity of *adaption* of the filter properties.

A filter which is adapted to a signal-to-noise ratio $\mathrm{SNR}_\xi = 1.0$ results in the images 4(e) and 4(f). The corresponding large receptive fields sum signals from a large area of the image. By this 'averaging' white noise can be partially eliminated at the expense of fine details, which are in any case lost in the noisy image 4(b). In 4(f), however, coarse structures stand out more clearly as before in 4(b). By comparing the squares in image 4(e) and 4(c) the inverse characteristic of the filters for low and high SNR_ξ becomes obvious. The edge is smoothed in 4(e) instead of being sharpened in 4(c). The infomax filter is enhancing structure if possible and, conversely, this filter corresponds to a kind of regularization (smoothing) in the presence of noise.

2.4. Implementation of infomax-like filters by simple networks

Lateral inhibition. Phenomena analogous to Mach bands are observable in many sensory systems and are commonly referred to as the appearance of 'lateral inhibition'. This notion expresses the suggestion that corresponding effects are evoked by an interaction extending in lateral direction between sensory elements. Indeed, from his observations Mach himself immediately drew the conclusion that there must be a lateral inhibitory interaction in the retina [46]. Hence a possible simple attempt to implement filters of the desired form is to couple the nearest neighbours in a photoreceptor array by an inhibitory interaction. This should result in high-pass properties. For adaption to low SNR_ξ the interaction can be switched continuously from negative to positive. This should change the characteristic of the network from high-pass to low-pass.

The dynamics of a linear network with external input $\xi_{\boldsymbol{x}}$ and an nearest-neighbour interaction is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\sigma_{\boldsymbol{x}} = -r\sigma_{\boldsymbol{x}} + w\sum_{\boldsymbol{y}\in\mathcal{N}(\boldsymbol{x})}\sigma_{\boldsymbol{y}} + \xi_{\boldsymbol{x}}. \tag{2.27}$$

Here $\mathcal{N}(\boldsymbol{x})$ denotes the set of nearest neighbours of $\boldsymbol{x}$ on the triangular lattice. Here w determines the strength of interaction and r is a time constant. We pass to the frequency domain, where equation (2.27) reads

$$\frac{\mathrm{d}}{\mathrm{d}t}\sigma(\boldsymbol{k}) = -\,r\sigma(\boldsymbol{k}) + w\sigma(\boldsymbol{k})\sum_{\boldsymbol{a}\in\mathcal{N}(\boldsymbol{0})}\mathrm{e}^{-\mathrm{i}\boldsymbol{a}\cdot\boldsymbol{k}} + \xi(\boldsymbol{k}). \tag{2.28}$$

In the sum the vector $\boldsymbol{a}$ connects $\boldsymbol{0}$ to its nearest neighbours on the triangular lattice, written as $\sum_{\boldsymbol{a}\in\mathcal{N}(\boldsymbol{0})}$. These vectors are displayed in figure 5. They are located on a circle of radius a. To restore rotational symmetry we replace the sum over the six neighbours by an integral over the circle:

$$\frac{1}{|\mathcal{N}|}\sum_{\boldsymbol{a}\in\mathcal{N}(\boldsymbol{0})}\mathrm{e}^{-\mathrm{i}\boldsymbol{a}\cdot\boldsymbol{k}} \approx \frac{1}{2\pi}\int_0^{2\pi}\mathrm{d}\varphi\;\mathrm{e}^{-\mathrm{i}ka\cos\varphi} = J_0(ak) \tag{2.29}$$

where $|\mathcal{N}| = 6$ is the number of nearest neighbours. In the frequency domain the nearest-neighbour interaction is represented by a Bessel function J_0 of order zero. The lattice constant a enters the argument of the Bessel function.

We are solely concerned with a spatial theory and, therefore, we set the time constant r to unity and consider only the stationary points of the dynamic equation (2.28). Together

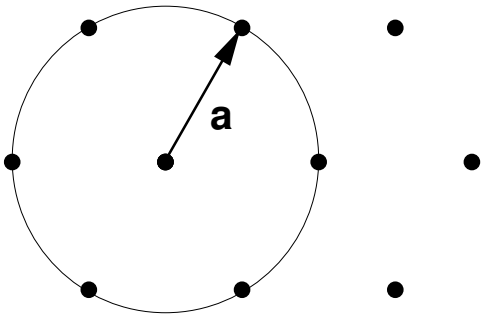


Figure 5. In a triangular lattice, every site has six nearest neighbours, which are located on a circle with radius corresponding to the lattice constant a. The sum over the six neighbours is approximately replaced by the integral over the displayed circle.

with equation (2.29), we obtain

$$\sigma(\boldsymbol{k}) = \frac{1}{1 - wJ_0(ak)}\xi(\boldsymbol{k}). \tag{2.30}$$

In the frequency domain the filter corresponding to the nearest-neighbour network is then given by

$$G_{\mathcal{N}}^{w}(k) \equiv \frac{1}{1 - wJ_0(ak)} \approx \frac{1}{1 - w\left[1 - \frac{1}{4}(ak)^2\right]}. \tag{2.31}$$

We have denoted the interaction parameter w as an upper index of G_N. The second approximate equality follows by expanding the Bessel function in the first Brillouin zone.

Some of these filters are displayed in figure 6. It is obvious that through a positive interaction w a low-pass filter can be implemented very efficiently. The high-pass filter, however, which is implemented by an inhibitory, negative w, is very unsatisactory; cf the desired filters of figure 2. The high-pass is the result of the pole of (2.31), which at best can be placed just beyond the boundary of the first Brillouin zone. But even in this case this does not affect the low frequencies efficiently enough. The corresponding results, which

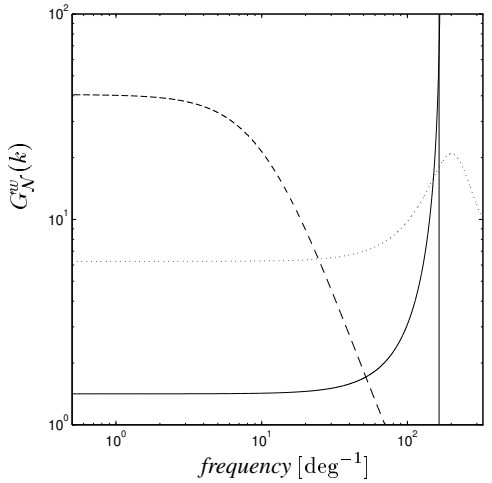


Figure 6. The filters $G_{\mathcal{N}}^{w}$ generated by a nearest-neighbour interaction according to equation (2.31). Full line, $w = -3.0$; dotted line, $w = -1.0$; broken line, $w = +0.99$. Obviously the desired high-pass properties of the infomax filters for low noise levels cannot be implemented through plain lateral inhibition of nearest neighbours.

we obtained by filtering the test images of figure 4, were disappointing from our subjective point of view.

Two-layered implementation. Plain lateral inhibition does not seem to be an appropriate method of high-pass filtering; on the other hand, however, a low-pass filter can be implemented very efficiently by a positive interaction. In what follows we describe a two-layered implementation scheme, which turns out to be very practical and uses only low-pass layers. The desired high-pass effect is implemented as 'one minus the equivalent low-pass'.

The corresponding interaction scheme is illustrated by figure 7. First, the signal is low-pass filtered by coupling nearest-neighbour receptor cells ξ. The signal of the first layer is fed to an additional interaction layer, which is also low-pass by means of the positive couplings u. The signal of this layer is subtracted from the signal of the first layer and this difference forms the output σ of the filter. This scheme results in the filter function

$$G^{uvw}(k) = G^{w}_{\mathcal{N}}(k)\left(1 - vG^{u}_{\mathcal{N}}(k)\right) \tag{2.32}$$

where the lower index $\mathcal{N}$ refers to the nearest-neighbour interaction. Figure 8 shows some filters $G^{uvw}(k)$ for different parameters and the effect of $G^{uvw}(k)$ on the test images figure 4(a) and 4(b) is demonstrated in figure 9. These results show that with the scheme of figure 7 the desired high-pass and low-pass properties can be implemented very efficiently. The parameters in the caption of figure 8 show that we can change the filter characteristic from high-pass to low-pass by simply switching off the couplings u of the second layer and switching on the couplings w of the receptor layer. We did not change the parameter v, but this parameter may also be used to control the filter properties.

With the scheme of figure 7 we have succeeded in efficiently implementing high- and low-pass filters by using *only nearest-neighbour interactions*. This is very important if such a scheme is to be realized on a silicon chip. Long-range interactions cannot be disentangled in two dimensions and, hence, would be very expensive to implement.

Finally, a comment on spatio-*temporal* extensions of the presented concepts is in order. Aiming at a spatial theory, we have considered only stationary points of the dynamics and

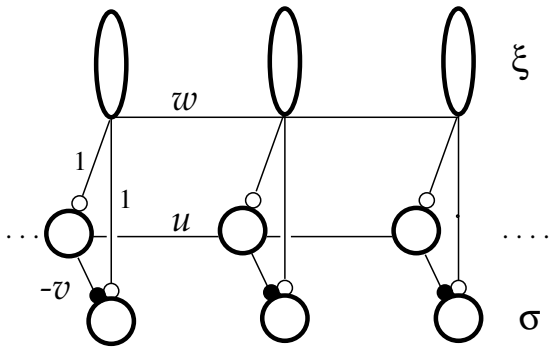


Figure 7. Two-layered implementation scheme with nearest-neighbour interactions alone. The receptors ξ are coupled through a positive interaction w. Their signals are the input to a second layer, which is also low-pass because of a positive interaction u. The final output is a weighted difference between the receptor layer and the second layer. The corresponding 'inhibitory synapse' has the strength v and is displayed by a small full circle. These synapses act only in one direction, as do the excitatory synapses, which are displayed by small open circles. Any other connexions within the two layers are symmetric.

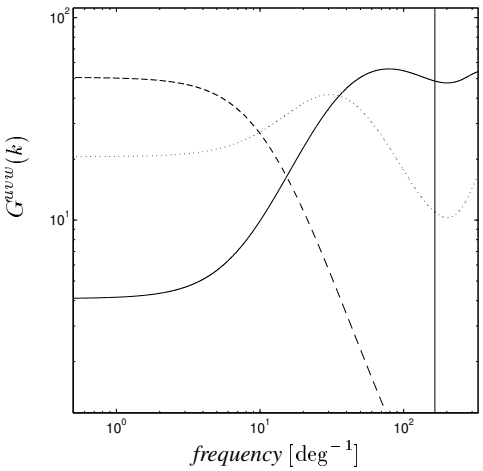
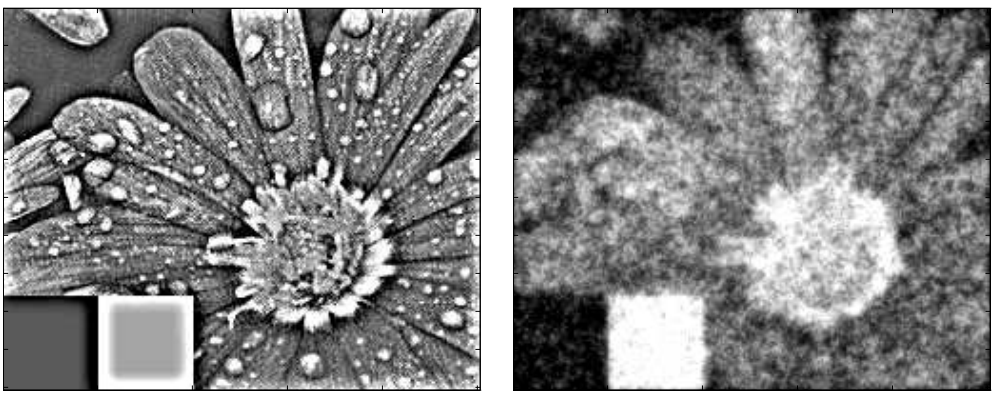


Figure 8. Examples of filters which follow from the two-layered scheme of figure 7. Again note the logarithmic scales. Full line, $u = 0.9$, $w = 0.3$; dotted line, $u = 0.89$, $w = 0.9$; broken line, $u = 0.2$, $w = 0.99$. The parameter $v = 0.095$ is the same for all three filters. We have $G^{uvw}(k) \to 0$ for $k \to 0$ if $v = 1 - u$. Thus to implement a high-pass filter efficiently, one needs a sensible interplay between the parameters u and v. Furthermore, the figure clearly shows that by decreasing u and increasing w one can turn a high-pass filter into a low-pass one.



(*a*) (*b*)

Figure 9. Demonstration of the filter G^{uvw} according to the two-layered scheme of figure 7. (*a*) The original image in figure 4(*a*) after applying the filter G^{uvw} with the parameters $u = 0.97$, $v = 0.028$ and $w = 0.5$. (*b*) The filtered noisy picture in figure 4(*b*) with the parameters $u = 0.1$, $v = 0.028$ and $w = 0.9$. The parameters were determined in a subjective way. It is obvious that with the two-layered scheme of figure 7 it is possible to implement diverse useful filters by simply adjusting the nearest-neighbour interactions.

dropped irrelevant time-integration constants, e.g., the parameter r in equation (2.28). We expect that time-integration constants are crucial parameters of a spatio-temporal theory and will have to be adapted to the signal-to-noise ratio as well. They determine the amount of temporal smoothing of a signal in one layer in the same way as nearest-neighbour interactions determine the spatial smoothing of the signal. It may be well to realize, however, that finite time-integration constants presuppose capacitance facilities (condensers), which are expensive to implement on a silicon chip.

3. The retina as an infomax filter

We have extensively discussed the infomax filter and ways to implement a corresponding 'on-chip' preprocessing by a simple network. A 'natural' issue is now to ask what kind of solution Nature has offered to us through the long process of evolution. First, we will compare the response properties of the vertebrate retina with those of the infomax filter. So far the retina is just a black box. Next, we will have a look at the network of the retina and compare the retinal network with the scheme that we found advantageous for our on-chip implementation of infomax filters.

3.1. The retina as a black box

One aspect of the response properties of the retina is described by the so-called contrast sensitivity, which is determined experimentally at the retina; see, e.g., [21]. The contrast sensitivity function may be compared directly with the filter function $G(k)$; cf the reasoning in [21]. Only one global constant remains undetermined.

One example of a contrast sensitivity function is shown in figure 10 (full line), which is measured at the parvocellular part of the lateral geniculate nucleus (LGN) of a macaque. Given the quality of the data available at the moment, the *spatial* response properties of cells of the LGN do not differ significantly from the properties of corresponding ganglion cells [30, 29]. Some minor differences are reported in [25, 30]. Figure 10 additionally shows an infomax filter corresponding to equation (2.17). The parameters of this filter were determined by eye to obtain agreement with the experiment. For the lattice constant we had to use $a = (1/30)^\circ$, which does not correspond to central foveal vision. In [18] the foveal eccentricity of the measured neuron is not specified, and we have no information about the lattice constant. Hence it does not make sense to speculate about noise levels in the retina under experimental conditions as long as we are not sure about the most important

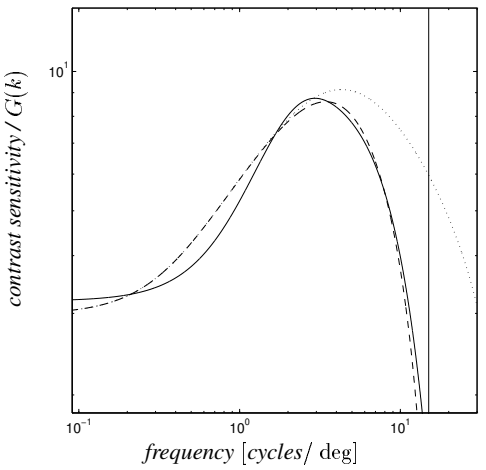


Figure 10. The contrast sensitivity of a neuron in the parvocellular part of the LGN of a macaque compared with the infomax filter. The full line corresponds to experimental data of [18, table 1, parameters of cell number 15M]. The dotted line is an infomax filter with the parameters $g/\Delta = 1.0$, $(p-\delta)/\delta = 4.6$, $a = (1/30)^\circ$ and $\kappa = 2.5 \times 10^{-2} k_B$. For the dashed line the structure function for the finite extent of the cones and the point spread of the optical apparatus has been taken into account according to equation (3.2).

parameter, namely the lattice constant.

In the fovea the cones are packed densely, i.e. their extent corresponds to the lattice constant a. The point spread of the optical apparatus [63] is also of the same order of magnitude as the lattice constant. Hence, the system does not support frequencies of the higher Brillouin zones above $k_B \sim 1/a$ and, therefore, aliasing is, on the whole, prevented [49]. The finite extent of the cones and the optical point spread lead to an additional modulation of the input spectrum within the first Brillouin zone as well, which we did not take into account in the first part of the paper, for the sake of simplicity. We now introduce a structure function $\sim \exp(-k^2/k_B^2)$ to describe the finite extent of the cones and the optical point spread, which corresponds to a 'soft cut-off' at the boundary of the Brillouin zone. In the infomax filter (2.17) this structure function is take into account by the replacement

$$C_\xi(k) \to C_\xi(k)\, \mathrm{e}^{-2k^2/k_B^2}. \tag{3.1}$$

The total system of optical apparatus and neural filter then has response properties according to the product of a structure function and a neural filter

$$G_{tot}(k) = G(k)\mathrm{e}^{-k^2/k_B^2}. \tag{3.2}$$

The dashed line of figure 10 corresponds to this filter function.

So far we have not discussed what happens at the boundary to the higher Brillouin zones, i.e. at the vertical line displayed in any figure showing filter functions. If we have an absolutely regular, infinite lattice, the neural response function of the first Brillouin zone has to be continued periodically to the higher Brillouin zones. Thereafter, to obtain the response properties of the total system, we have to multiply the neural filter and the structure function again, as in equation (3.2). An example for a filter which emerges by this procedure, is shown in figure 11. We find an additional bump beyond the first Brillouin zone, if the filter function crosses the boundary of the Brillouin zone at a very high filter gain, i.e. only if we

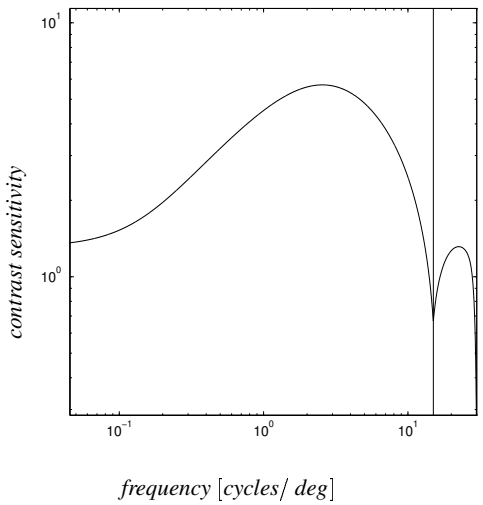


Figure 11. At low noise levels the discrete sampling by the photoreceptor array may result in an additional bump *beyond* the first Brillouin zone, which finally is the reason for possible aliasing effects. There are only very few hints at such a bump in experimental contrast sensitivity measurements; see, e.g., [29].

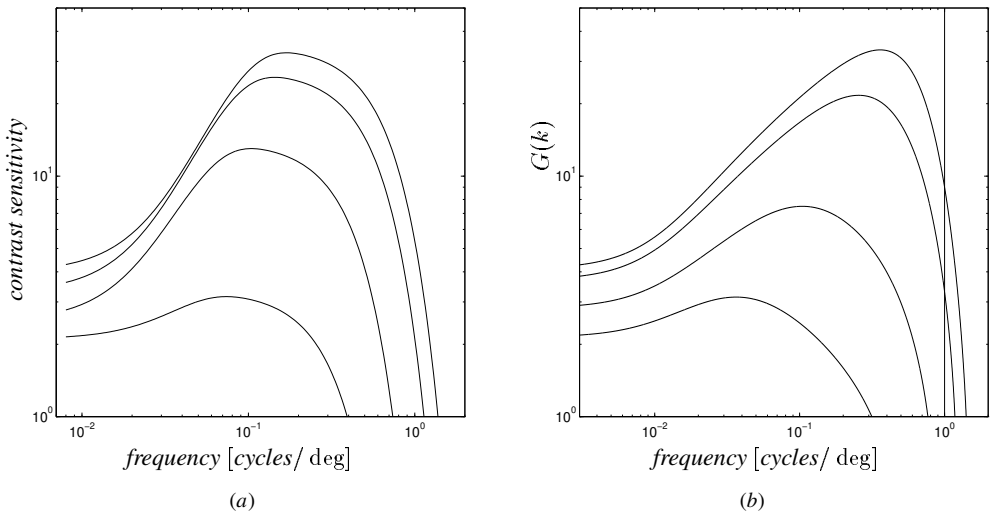


Figure 12. Contrast sensitivity of (a) the cat retina for different mean background illuminations [21] compared with (b) the infomax filter according to equation (3.1) and (3.2). The experimental data correspond to an X-ganglion cell. This class of cells of the cat retina may be viewed as equivalent to the visual P-system of other species [29, 39]. The upper contrast sensitivity corresponds to the highest background illumination. Common to all the theoretical filters are the parameters $\kappa = 10^{-2}k_B$, $a = 0.5^\circ$, $\Delta = 0.3$ and $\delta = 0.01$. For the uppermost curve to the lowest the remaining parameters are respectively $g/\Delta = 4.3$, 1.3, 0.14, 0.017 and $(p-\delta)/\delta = 25.0$, 10.0, 3.0, 0.5. The parameter λ has to be determined from p via equation (A.5), although this equation does not apply here exactly because of the additional structure function.

have a very low noise level. Aliasing effects are the final consequence of this additional structure. We found only one experimental contrast sensitivity function where indications for such a bump appear; see [31] or [29]. However, in [31] and [29] this bump is interpreted as a fine substructure of the receptive field, namely, as an 'on-centre subunit'. A correct interpretation is possible only if the corresponding 'local' lattice constant is determined simultaneously in the experiment. Supposing that the interpretation given here is correct, we may conversely determine the lattice constant a by observing a renewed increase of the contrast sensitivity at high frequencies. By using additional technical supplements to improve the quality of the optical apparatus of the eye, psychophysical effects may indeed be evoked which correspond to aliasing, e.g., impressions resembling Moiré patterns [65, 67]. However, this mainly concerns parafoveal vision.

Figure 10 shows that a single contrast sensitivity measurement can be fitted very well. However, an essential feature of the retina is the ability to adapt to illumination conditions. In figure 12(a) contrast sensitivity measurements of the cat retina for different background illuminations are presented [21]. Obviously the filter characteristics of the retina change drastically from high-pass to low-pass as one decreases the illumination conditions (logarithmic scales!). We have obtained a corresponding variation for the infomax filter in figure 12(b) by assuming decreasing ratios g/Δ and $(p-\delta)/\delta$. Again we have used the parameters g/Δ and $(p-\delta)/\delta$ as free parameters to fit the experiment. Corresponding information on noise levels in the retina is not available from [21]. Moreover, we have used the lattice constant $a = 0.5^\circ$ to fit the experiment. This would correspond to a position in the retina far away from the fovea. For different reasons which we already hinted at above, it is more difficult to investigate parafoveal vision. Our theory mainly applies to foveal vision.

On experimental data. In many other work we find receptive fields of the magnitude of 1° and frequency scales of the magnitude $\sim 1/1^\circ$ [17, 25, 26, 28]. Mostly no information is given about the foveal eccentricity, and we have to assume parafoveal measurements. Receptive fields of the magnitude of 1° are not consistent with, e.g., the human foveal resolution of about 0.01°. It is known that in the central fovea there exist receptive fields with a corresponding small extent [18, 29, 63], which is also obvious from psychophysical contrast sensitivity measurements [15, 32]. Psychophysical measurements closely resemble the measurements of figure 12(a), with the exception that the scales of the frequency axis differ by two orders of magnitude. In the central fovea we expect contrast sensitivities similar to psychophysical measurements, cf [15, 32]. However, psychophysical measurements cannot be used for comparison with a theory of the retina.

It is obvious that further experimental data are desirable. As far as known to us, there are no experimental data available *showing the dependence on the background illumination for a midget ganglion cell in the foveal centre*. To compare with the theory presented here it would be even better to use midget bipolar cells to exclude effects of the inner plexiform layer. Simultaneously to those measurements the 'local' lattice constant should be determined. Furthermore, it should be possible to access the signal-to-noise ratios of the different cells in the experiment (dependent on the illumination conditions and the stimulus frequency), and to estimate the accuracy of further information processing in the LGN and the visual cortex. The accuracy of further information processing should also enter the parameter δ, because the coding of the retina should also take into account the differences that can be resolved in later stages of information processing. Finally, it would be interesting to perform psychophysical measurements under exactly the same conditions. Agreements and disagreements with the retinal contrast sensitivity may give a hint at to what extent the firing rates are indeed the 'code words' of the retina.

3.2. Comparison of the network structures

The retina is a thin nervous tissue at the back of our eye, at most 0.5 mm thick, like cellophane. This membrane is the interface to our visual environment and, hence, corresponds to our most important sense. Nevertheless, up to now our understanding of the retina is superficial, and the complexity of the retina still retains some puzzles. For a comprehensive treatment of the retina we refer the reader to Dowling [20].

Information processing in the vertebrate retina may be viewed as being carried out at two different levels: that of the inner plexiform layer (IPL) and that of the outer plexiform layer (OPL). These two layers perform different tasks. The OPL seems to correspond to the P system of the later stages of visual information processing [39]. The IPL seems to be the basis of the M system. The OPL is mainly concerned with the *static* and *spatial* aspects of illumination. Neurons of this layer respond linearly through sustained, graded potentials. Conversely, the IPL is concerned primarily with the *dynamic* and *temporal* aspects of visual stimuli. Neurons of this layer respond in a transient way through a mixture of graded potentials and action potentials. They are vigorously excited by moving stimuli. Because of their nonlinear response properties, linear system analysis cannot be applied in a straightforward way. Nevertheless, we may characterize this layer by strong high-pass properties, in particular, with respect to its temporal information processing. This property of the IPL will be important later on, when we discuss a further hypothesis concerning the interplay of the IPL and the OPL.

In what follows we are mainly concerned with the OPL of the retina, which is much better understood than the IPL. The receptive fields with centre-surround organization

reflect the information processing of this layer. They already appear at the level of the bipolar cells. Their excitatory centre corresponds to the direct path RECEPTOR CELL $\longrightarrow$ BIPOLAR CELL, while the surroundings are mediated via the horizontal cells [48, 50] and its realization is often illustrated by the path RECEPTOR CELL $\longrightarrow$ HORIZONTAL CELL $\longrightarrow$ BIPOLAR CELL. The inhibitory surround of a bipolar cell is due to the large receptive fields of the horizontal cells, which, in turn, is mainly a consequence of the electrical gap junctions between the horizontal cells [38]. In view of these facts, it is obvious how the scheme of figure 7 can be mapped onto corresponding structures of the OPL:

- The inter-receptor couplings of figure 7 correspond to the gap junctions, which can be found selectively between cones or between rods. These symmetric electrical junctions are mediated via 'short processes' [20] of the receptor cells. In particular, the rods are strongly coupled via gap junctions, which is consistent with the interpretation that this electrical coupling serves to suppress noise at low ambient illumination. At least in some species we find extensive electrical coupling, e.g., it is reported that in the toad retina up to 9000 receptor cell are coupled effectively [20, 22]. It has been suggested that these inter-receptor couplings serve to damp the membrane fluctuation [37]; see also [20] and references therein.
- The bipolar cells correspond to the output layer σ of figure 7. They are fed by the receptor cells via directed chemical synapses (small open circles of figure 7). The midget bipolar cells of the fovea contact only one cone by and large [20], just as one midget ganglion cell contacts only one midget bipolar cell. This arrangement is responsible for the small receptive fields and the high resolution of the fovea.
- Finally, the horizontal cells correspond to the additional layer of figure 7. They are fed by the cones via chemical synapses too, and they propagate signals mainly in lateral direction within the OPL. This lateral propagation is on the one hand due to the wide-spread ramification of the horizontal cells. On the other hand, as mentioned above, the extensive gap junctions between the horizontal cells give rise to an additional lateral propagation and large receptive fields. These symmetric electrical synapses correspond to the symmetric interactions u of the additional layer of figure 7. The horizontal cells act on the bipolar cells via directed chemical synapses in an inhibitory way, i.e. with a different sign on the *on* and the *off* bipolar cells. These synapses correspond to the full circles of figure 7.

Of course, there is plenty of structure and detail of the OPL, which has no adequate counterpart in the scheme of figure 7. In particular, we have to mention the feedback of the horizontal cells to the receptors, which is occasionally viewed as the main reason for the centre-surround structure of the receptive fields [51]. The extent of the feedback HORIZONTAL CELL $\longrightarrow$ RECEPTOR CELLS and its dependence upon the illumination conditions [34] seems to be unclear. We suppose that this feedback is mainly important in a spatio-*temporal* formulation. It has been shown here that the scheme in figure 7 is the minimal structure which is necessary for a efficient *spatial* on-chip information processing, and we hope that with the above processing scheme we have gained the essentials of the spatial information processing of the OPL. This view is underpinned by some recent experimental observations which shed some light on a mechanism, which, possibly among others, contributes to the network adaption of the retina. It is analogous to the reduction of the interaction u of the additional layer which we used to switch from a high-pass to a low-pass filter. The corresponding recent experimental findings, which are mainly a case study of the teleost fish retina, are comprehensively summarized in [20]. See this book for references and many interesting further details.

In addition to the commonly known five basic classes of retinal neurons there exists an additional class, that of the so-called interplexiform cells. Since they are refractory to Golgi staining they were recognized as a distinct class of retinal neurons only 80 years after the pioneering work of Cajal. The interplexiform cells play an essential role in dark adaption of the retina. They appear to be a kind of feedback path, carrying information from the IPL back to the OPL. All of the input to these cells comes from in the IPL, whereas most of the synapses originating from the interplexiform cells are found in the OPL. The interplexiform cells contain the neuromodulator dopamine. Neuro*modulators* are neuroactive agents which act in quite a different way as compared with neurotransmitters. A neuromodulator does not directly affect the postsynaptic membrane potential. Rather its effects are mediated biochemically. Dopamine, for instance, appears to activate an intracellular enzyme system and the physiological changes induced by dopamine are multiple and long-lasting.

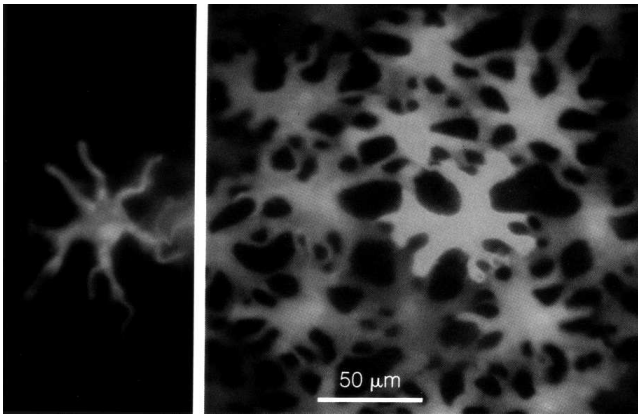


Figure 13. Horizontal cells stained by a fluorescent dye. The cell on the left is from a retina treated with dopamine, which restricts the diffusion of dye from the injected cell to the surrounding cells. The micrograph on the right is from a control preparation, in which dye spreads from the injected cell (centre) to the neighbouring cells via the gap junctions. Obviously dopamine 'narrows the pores' of the electrical gap junctions and, as a consequence, increases the resistance between horizontal cells. Note that horizontal cells are analogue neurons so that 'resistance' is a well-defined notion. (The original micrographs are attributed to T Tonquist and X L Yang, and are reproduced from [20] by permission of Harvard University Press.)

Mostly horizontal cells are postsynaptic to the interplexiform cells and, accordingly, the effect of dopamine on the horizontal cells has been investigated in some detail. From a series of experiments [20] at least two important effects of dopamine became evident. First, dopamine changes the responsiveness of the horizontal cells. Second, the lateral propagation of signals between horizontal cells is depressed by dopamine, which is due to a decreased electrical coupling between adjacent horizontal cells. In figure 13 we have reproduced an experimental result from [20] which impressively demonstrates the effect of dopamine on the gap junctions. Substantial evidence that dopamine alters the electrical coupling between horizontal cells has been gained from a variety of further experiments; for details see [20].

The release of dopamine is controlled by the interplexiform cells. This is confirmed in experiments where selective destruction of the interplexiform cells broadens the receptive fields of horizontal cells and enhances dye diffusion between them. The only question now is: when do interplexiform cells release dopamine? Evidently the release of dopamine is correlated with the adaption to darkness. This is inferred from the observation that

application of dopamine causes changes in the response properties of horizontal cells analogously to changes caused by dark adaption.

Altogether, there is a strong analogy between the adaptation mechanism in the retina and the decrease of the interaction parameter u to adapt the filters (2.32) to a noisy signal. Besides the direct effect on the gap junctions, dopamine seems to cause some further changes in the OPL [34] so that a dependence of the inhibition v upon the illumination conditions might be consistent with biological facts as well.

3.3. Three further hypotheses on the functionality of the retina

We have mentioned that the two-layered scheme of figure 7 should not be taken as an attempt to build a detailed model of the OPL. Instead, we hope that by mapping the scheme of figure 7 onto the topology of the retina we have got to the heart of the spatial, static information processing of the OPL. More refined models should be considered, but yet, on the basis of this temporary understanding of the retina, some further hypotheses about the functionality of the retina suggest themselves. They are formulated as follows.

- First, it is evident from our theory that the inter-receptor couplings may depend on the illumination conditions or on the noise level as well. There are some experiments which do not confirm this hypothesis [37]. Nevertheless, it should be verified in detail, at least for species which depend on the ability to see *in daylight as well as at night*. In the toad retina it has been found that each rod receives input from as many as 8000–9000 rods distributed over an area of 0.5 mm^2 [20, 22]. What is the use of such an extensive receptor coupling if the receptors cannot function independently under circumstances other than those of the experiment?
- The task of the *interplexiform* cells is to control the adaptation of the *outer plexiform* layer (OPL). The interplexiform cells get all their input from the *inner plexiform* layer (IPL). An interesting question is therefore: how do the interplexiform cells manage to access the illumination conditions through their input couplings in the IPL? Our hypothesis is that the interplexiform cells do not ascertain the illumination conditions, but, moreover, they directly determine the signal-to-noise ratio. We have mentioned above the response properties of the IPL cells. Their nonlinear, transient responses mainly code temporal variations of an image so that this layer may be viewed as a high-pass filter, at least with respect to its temporal properties. But this also means that the IPL is very sensitive to noise; white noise means 'fast changing'. Hence the noise level may be estimated by observing the response at the IPL. This is what the dendritic arbors of the interplexiform cells might do. We have summarized this hypothesis in figure 14. If it is correct, it should be possible to bring about a release of dopamine and consequently a dark adaption (increase) of the receptive fields *in spite of high average illumination conditions* by stimulating the retina with white noise or with a high-frequency spectrum. An experiment which verifies of refutes this hypothesis can easily be set up. A positive result would also have implications for our every-day lives in an unnatural environment, e.g., for the design of computer screens. A computer screen should be designed so as to avoid stimulating the IPL.
- It is fair to speculate about the purpose of the invaginations, which still constitute a mystery; see figure 15. Within the invaginations in the fovea the processes of two horizontal cells and one bipolar cell are coupled to the terminal of a receptor cell in a highly specific way [20]. We suggest that this construction may be used to regulate the relative amplitude of the horizontal-cell layer and the bipolar-cell layer. In our scheme of figure 7 the high-pass filter has been implemented as 'one minus the equivalent low-pass filter'. This

strategy depends in a sensitive way on the accurate matching of the relative amplitudes of the two signals. In particular, if we demand $G(k) \to 0$ as $k \to 0$ so as to produce a very efficient high-pass filter, we obtain the condition $v = 1 - u$ for the parameters u and v. Our hypothesis is now that the invaginations serve this matching of parameters, i.e. the correct relative amplitude v of the bipolar and the horizontal cell layer.

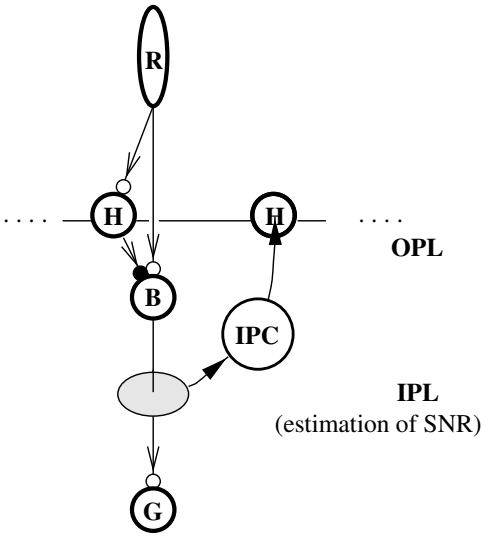


Figure 14. Illustration of the hypothesis on the interplay between the inner (IPL) and outer plexiform layer (OPL). The OPL corresponds to the processing scheme of figure 7: receptors (R), horizontal cells (H), and bipolar cells (B). The bipolar cells feed the ganglion cells (G) and the IPL, where the dendritic arbors of the interplexiform cells (IPC) can be found (pictured schematically; there are probably fewer interplexiform cells than horizontal cells). The interplexiform cells appear to be a kind of feedback path to the OPL [20]. It is suggested that in the IPL the signal-to-noise ratio is 'estimated' by the interplexiform cells, which, in turn, regulate the properties of the horizontal cells. This means that the retina is not adapted to the illumination conditions, but rather to the amount of noise. It should be possible to verify or refute this hypothesis by a simple experimental set-up.

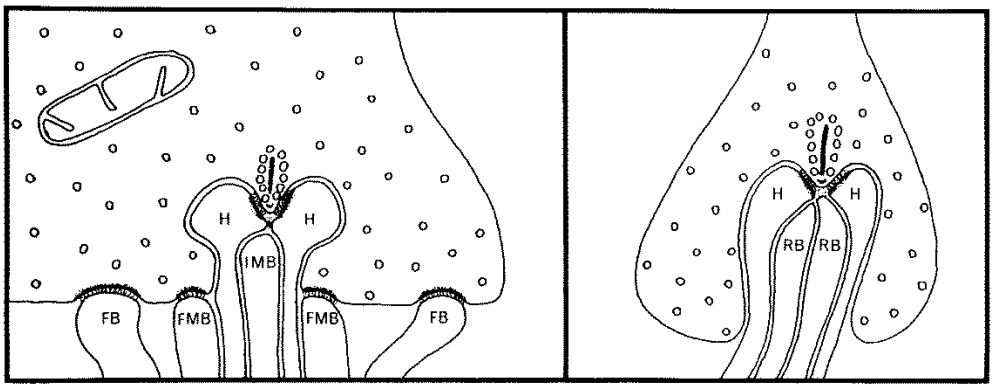


Figure 15. Within the *invaginations* the processes of horizontal and bipolar cells are coupled to the terminal of a receptor cell in a highly specific way. Left: in the fovea two horizontal cells (H) are placed laterally to one bipolar cell (IMB, *invaginating midget bipolar cell*; FB, *flat bipolar cell*; FMB, *flat midget bipolar cell*). Right: coupling scheme in the periphery (RB, *rod bipolar cell*). It is suggested that this obscure construction regulates the amplitude of the horizontal cell layer relative to the amplitude of the bipolar cell layer. The successful implementation of infomax-like filters by the scheme of figure 7 depends in a crucial way on the correct matching of these amplitudes. (©1987 J E Dowling, reprinted from [20] by permission of Harvard University Press.)

4. Redundancy and noise-robustness of an internal representation

The infomax approach is based on probabilities only and, therefore, is very universal. Its drawback, however, is that this approach is also very abstract. Hence, some more concrete interpretations would be desirable, at least for some special cases. We will now offer an additional characterization of the infomax ansatz which is possible in the case of Gaussian distributions.

By our sensory information processing an external input $\boldsymbol{\xi}$ is mapped onto an internal representation $\boldsymbol{\sigma}$. For better intuition we may imagine $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ to be the faces of two different people. To discriminate these two people it is necessary that their internal representations $\boldsymbol{\sigma}_1$ and $\boldsymbol{\sigma}_2$ are different. Otherwise we will confuse them. For $\boldsymbol{\xi}_1 \neq \boldsymbol{\xi}_2$ we demand $\boldsymbol{\sigma}(\boldsymbol{\xi}_1, \boldsymbol{\mu}_1, \boldsymbol{\nu}_1) \neq \boldsymbol{\sigma}(\boldsymbol{\xi}_2, \boldsymbol{\mu}_2, \boldsymbol{\nu}_2)$, *although* the noise $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ may be different from case to case. *Noise is just what makes the whole thing interesting*. Otherwise *any* bijective function would do equally well. We count the number of mistakes by the following δ-function:

$$\#\text{ mistakes} = \left\langle \delta^N \left[\boldsymbol{\sigma}(\boldsymbol{\xi}_2, \boldsymbol{\mu}_2, \boldsymbol{\nu}_2) - \boldsymbol{\sigma}(\boldsymbol{\xi}_1, \boldsymbol{\mu}_1, \boldsymbol{\nu}_1) \right] \right\rangle \tag{4.1}$$

where δ^N is a δ-function whose argument is an N-vector. The number of mistakes may be used as a measure of the performance of a sensory system. It is as good as any function which increases or decreases monotonically with the number of mistakes (4.1). We, therefore, choose minus the logarithm of the average number of mistakes, $-\log\langle\delta^N[\boldsymbol{\sigma}_2 - \boldsymbol{\sigma}_1]\rangle$, as a quality measure which should be maximized.

In the same way we construct a second term with the aim to ensure that the *same* external signal $\boldsymbol{\xi}$ is repeatedly mapped onto the *same* internal representation $\boldsymbol{\sigma}$ in spite of the noise $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$. An appropriate term is given by $\log\langle\delta^N[\boldsymbol{\sigma}(\boldsymbol{\xi}, \boldsymbol{\mu}_2, \boldsymbol{\nu}_2) - \boldsymbol{\sigma}(\boldsymbol{\xi}, \boldsymbol{\mu}_1, \boldsymbol{\nu}_1)]\rangle$, which may be viewed as the diagonal elements of the first term. Our final measure for the performance of a sensory system is

$$\begin{aligned} K[\boldsymbol{\sigma}; \boldsymbol{\xi}] \equiv & -\log\left\langle \delta^N \left[\boldsymbol{\sigma}(\boldsymbol{\xi}_2, \boldsymbol{\mu}_2, \boldsymbol{\nu}_2) - \boldsymbol{\sigma}(\boldsymbol{\xi}_1, \boldsymbol{\mu}_1, \boldsymbol{\nu}_1) \right] \right\rangle \\ & + \log\left\langle \delta^N \left[\boldsymbol{\sigma}(\boldsymbol{\xi}, \boldsymbol{\mu}_2, \boldsymbol{\nu}_2) - \boldsymbol{\sigma}(\boldsymbol{\xi}, \boldsymbol{\mu}_1, \boldsymbol{\nu}_1) \right] \right\rangle . \end{aligned} \tag{4.2}$$

This cost function measures the robustness of an internal representation against internal and external noise. It is an easy exercise to explicitly evaluate the expression (4.2) for Gaussian distributions, which are given by their covariance matrix $\mathbf{C}$. In the present case we obtain

$$K[\boldsymbol{\sigma}; \boldsymbol{\xi}] = \log\det(\mathbf{C}_\sigma) - \log\det(\mathbf{C}_{\sigma|\xi}). \tag{4.3}$$

Comparing equation (4.3) with (2.14) we find $K[\boldsymbol{\sigma}; \boldsymbol{\xi}] = 2M[\boldsymbol{\sigma}; \boldsymbol{\xi}]$, i.e. we have recovered the mutual information. By this equivalence we now have an additional interpretation of the infomax approach in the Gaussian case. High mutual information helps us to discriminate external stimuli within our internal representation despite disturbing noise. In other words, the infomax approach guides the way to a 'noise-robust internal representation of sensory information'.

In general, noise-robust, fault-tolerant systems use 'redundancy' to overcome possible errors. We, therefore, investigate the statistical dependences of the internal units $\boldsymbol{\sigma}$ in dependence upon the noise parameter Δ/g supposing an infomax-optimal linear map $\mathbf{G}$. These dependences may be best measured by a corresponding information-theoretic expression, namely the redundancy

$$R[\boldsymbol{\sigma}] = \langle \log P(\boldsymbol{\sigma}) \rangle - \sum_{x \in \mathcal{L}} \langle \log P(\sigma_x) \rangle \, .$$

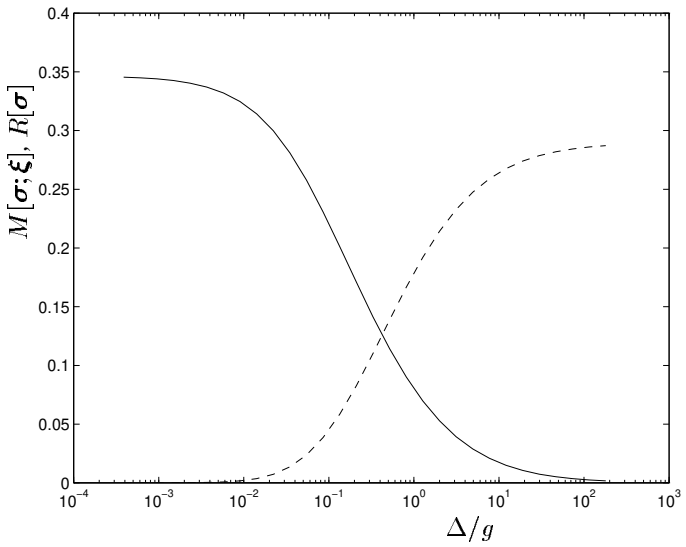


Figure 16. The mutual information (full line) and the redundancy (broken line) in dependence upon Δ/g, which is equivalent to the inverse signal-to-noise ratio SNR_{ξ}^{-1} at the input. The variance of the internal signal and the variance of the internal noise has thereby been fixed to obey the ratio $p/\delta = 2.0$. This is guaranteed by choosing an appropriate Lagrange parameter λ for every Δ/g; see appendix A.1. The redundancy increases monotonically with increasing Δ/g.

In appendix A.1 the mutual information $M[\sigma;\xi]$ and the redundancy at the internal representation $R[\sigma]$, which result from the infomax-optimal mapping, are evaluated in dependence upon the noise parameters and the Lagrange parameter λ, which fixes the power constraint; see equation (2.11). In figure 16 we have plotted the mutual information and the redundancy against Δ/g. Thereby, the variance p of the internal representation is kept constant by choosing an appropriate Lagrange parameter λ for every value of Δ/g; see appendix A.1. Figure 16 shows that at low noise levels there is no necessity to use redundancy. However, as Δ/g increases, so does the redundancy. This increasing redundancy is necessary for an infomax-optimal filter, at least in our case of translation invariance. Translation invariance also means that every neuron has the same signal to noise ratio p/δ or every output element is represented with the same accuracy (number of bits in technical applications), which is a physically reasonable assumption†.

Under physically relevant conditions the infomax-optimal mapping will in general demand increasing redundancy of the internal representation with increasing noise. That is to say, the infomax ansatz differs from the redundancy reduction principle stated by Barlow [6, 7] and used in other work [2, 3, 4, 57]. In the noisy case the redundancy reduction approach has to be supplemented by an additional noise filter [2], which has to be put in by hand. On the other hand, the infomax ansatz can handle *any* noise level on its own. If the Kuhn–Tucker condition (2.18) is not met in the first Brillouin zone, the weight

† There is no constant signal-to-noise ratio $p(\boldsymbol{x})/\delta$ independent of $\boldsymbol{x}$ for the PCA-like solution discussed in [13]. This is a special solution with zero redundancy among a large class of infomax-optimal mappings which differ in redundancy. This large degeneracy of solutions is due to the distinguishing property of Gaussian distributions (namely, its being determined uniquely by its first and second moments alone) and allowing different ratios $p(\boldsymbol{x})/\delta$ for different internal neurons $\sigma_{\boldsymbol{x}}$.

vectors even become linearly dependent. This implies that some 'important' properties of the environment are transmitted 'twice' at the expense of some less important or very noisy properties. This is a strategy which we use in everyday life when we communicate information to other people. We mainly communicate important things, may be even twice at the expense of less important messages if the communication line is noisy. This of course means that at some noise level we decide to transmit redundant information. In dependence upon the noise level, the infomax ansatz trades diversity of information against redundant coding of essential information. The redundancy reduction approach cannot helps us to find that necessary compromise.

Redundancy is at the heart of many systems. As mentioned above, in the retina of some species we find 'receptor pooling', a kind of collective phenomenon, which of course results in a large amount of redundancy. But nevertheless, as we have seen, this pooling may be very useful. Redundancy is also supposed to be essential for the functioning of an associative memory which is designed to restore *noisy* patterns by the collaborative work of many neurons. This restoration process may be compared with the collective activity of the photoreceptors at night. Generally, the information theory of communication channels may be transcribed by the question: 'How much redundancy is *minimally* needed if we want to have a reliable communication over a *noisy* communication channel?' The answer is the famous channel coding theorem [14, 55]. As a consequence engineers have decided to use redundancy even in technical very reliable systems, e.g., the parity bits of our computers. The brain is a very fault-tolerant system with associative properties. Hence, we do not believe that Barlow's redundancy reduction hypothesis may be used as the *only* principle for understanding neural information processing as it has not incorporated the idea of 'collective computation' for noise-robust, associative information processing.

5. Outlook

In this work we have discussed the infomax approach and its essential properties by comparing it with visual information processing. Without any additional assumption this approach may serve as a closed ansatz over the *whole* range of illumination conditions. Hence, this is an appropriate approach to investigate retinal information processing because it is precisely the large range of its functioning which distinguishes the retina from the technical sensors of today. One would do well to remember that from a bright day to a moonless night the ambient illumination changes by many orders of magnitude and consequently the retina increases its sensitivity by a factor of about 6×10^6 [63]. We are thus facing very different signal-to-noise ratios, which require very different ways of preprocessing. Our retinal network equips us with this highly adaptive preprocessing. In the present paper an understanding of the retinal network and, in particular, of its adaptation properties has been derived from information-theoretic considerations. This understanding led us to some further hypotheses, which mainly concern the interplay between the inner plexiform layer and the outer plexiform layer. In particular, we predict that stimulating the retina by white noise might affect the shape of the retinal receptive fields in a way which resembles dark adaption. We hope that our hypotheses may serve as a basis for future investigations. The extension of the theoretical concepts presented here to a spatio-*temporal* formulation and more detailed modelling of the retinal network should also lead to a deeper understanding.

From the technical point of view, our considerations might guide the way to advanced concepts for a silicon retina. Increasing attention has been focussed on the idea of a technical sensory system analogous to its biological counterpart, that is, a sensory system where early-

stage preprocessing is inherent to the system itself [11, 27, 47, 68, 69]. None of the known silicon models, however, is able to adapt to the illumination conditions. There are several advantages of an adaptive sensory system with a simultaneous *on-chip* preprocessing. First, this preprocessing is implemented at the earliest possible level of information processing, i.e. noisy photoreceptors are directly coupled (cf w in figure 7) so as to produce a low-pass filter as early as possible. Second, an adaptive silicon retina might work over an extended range of illumination conditions as compared with conventional technical sensors. Third, a silicon retina may also serve as an image compression module as it reduces the dynamic range of the output signal relative to the internal resolution (internal noise). This property is explicitly built into the infomax filter by the power constraint. Consequently, the filter output may be digitized with only a few bits and restored to the full range by inverse filtering, e.g., after transmission. We expect to demonstrate this image compression properties of a silicon retina in a forthcoming paper. No computational power is necessary for any of these tasks because of the real-time analogue processing of a silicon retina.

Finally, in analogy with this work some fruitful considerations might be possible with respect to other sensory system, such as the auditory one. The phenomenon of adaptation and lateral inhibition is known from other sensory systems as well, e.g., from the auditory system; see [51, chapter 14]. As a consequence a new generation of technical sensory systems may result.

Appendix

A.1. Evaluation of the power constraint, mutual information and redundancy

In equation (2.11) we have used the variance p as a power constraint:

$$p = \mathrm{Trace}\big(\mathbf{C}_\sigma\big) = \sum_{x\in\mathcal{L}} C_\sigma(\boldsymbol{x}) = \sum_{\boldsymbol{k}} C_\sigma(\boldsymbol{k}). \tag{A.1}$$

Before performing the bulk limit $N \to \infty$ we have to normalize any extensive expression by dividing it by the number N of neurons, i.e. any sum has to be supplemented by the a factor $1/N$. Thereby, the variance (A.1) becomes the variance of a single neuron $p = \langle\sigma_{\boldsymbol{x}}^2\rangle$, which is independent of $\boldsymbol{x}$ because of translation invariance. In the limit of an infinite retina, $N \to \infty$, any sum becomes an integral over the first Brillouin zone:

$$\frac{1}{N}\sum_{\boldsymbol{k}} = \frac{1}{V_B}\sum_{\boldsymbol{k}}\left(\frac{V_B}{N}\right) \longrightarrow \frac{1}{V_B}\int_B \mathrm{d}^2\boldsymbol{k}. \tag{A.2}$$

To obtain p we now have to evaluate the integral of $C_\sigma(\boldsymbol{k})$ over the first Brillouin zone. $C_\sigma(\boldsymbol{k})$ is obtained from $C_\xi(\boldsymbol{k})$ according to the general expression (2.12) by $C_\sigma(\boldsymbol{k}) = |\mathbf{G}(\boldsymbol{k})|^2\big(C_\xi(\boldsymbol{k}) + \Delta\big) + \delta$. In our case, $|G(\boldsymbol{k})|$ and $C_\xi(\boldsymbol{k})$ are given by equation (2.17) and (2.5) and we obtain ($k = \|\boldsymbol{k}\|$)

$$C_\sigma(k) = \frac{1}{b}\frac{1}{k^2}\left(\sqrt{1 + \frac{2b\delta}{\lambda}k^2} - 1\right) \tag{A.3}$$

where $b \equiv 2\Delta/\delta g$. The variance of an internal neuron then follows directly:

$$p = \frac{1}{b}\,\frac{2\pi}{V_B}\int_0^{k_B} \mathrm{d}k\;\frac{1}{k}\left(\sqrt{1 + \frac{2b\delta}{\lambda}k^2} - 1\right). \tag{A.4}$$

This integral can be evaluated. We obtain an expression that gives p as a function of m and thus of λ:

$$p = \frac{2\pi}{bV_B}\left[m - 1 - \ln\frac{1+m}{2}\right] \qquad \text{with } m \equiv \sqrt{1 + \frac{2b\delta}{\lambda}k_B^2}. \tag{A.5}$$

We mainly need the inverse relation, namely m as a function of p, but equation (A.5) cannot be inverted explicitly. For a given p we therefore use the following iteration scheme:

$$m_{\text{new}} \equiv \frac{bpV_B}{2\pi} + 1 + \ln\left(\frac{1+m_{\text{old}}}{2}\right). \tag{A.6}$$

This mapping is a contraction and by iteration converges to the m that corresponds to the given p. From m the corresponding λ follows. Of course, the variance δ of the output noise has to be chosen less than the total variance p of the internal representation, i.e. $\delta < p$.

If the Kuhn–Trucker condition (2.18) is satisfied within the first Brillouin zone at $k_0 < k_B$, the variance p is given by the modified expression

$$p = \frac{2\pi}{bV_B}\left[m - 1 - \ln\frac{1+m}{2}\right] + \delta\left(k_B^2 - k_0^2\right)\frac{\pi}{V_B} \qquad \text{with } m \equiv \sqrt{1 + \frac{2b\delta}{\lambda}k_0^2}. \tag{A.7}$$

The second term in this expression takes into account that in the frequency range $k > k_0$ only the output noise contributes to the variance p. By a similar iteration scheme to that of (A.6), equation (A.7) can be inverted to obtain λ as a function of p.

Because of translation invariance, the entropy of a single ganglion $H[\sigma_x]$ is independent of the position $\boldsymbol{x}$ and, up to an additive constant, is given by

$$H[\sigma_x] = \tfrac{1}{2}\log\left\langle\sigma_x^2\right\rangle = \tfrac{1}{2}\log p. \tag{A.8}$$

For the joint entropy per ganglion $H[\boldsymbol{\sigma}]$ determined by the correlation matrix $C_\sigma(k)$, i.e. by (A.3), we find

$$\begin{aligned} H[\boldsymbol{\sigma}] &= \frac{1}{2V_B}\int_B \mathrm{d}^2\boldsymbol{k}\ \log C_\sigma(k) \\ &= \frac{1}{2V_B}\int_B \mathrm{d}^2\boldsymbol{k}\ \log\left[\frac{1}{b}\frac{1}{k^2}\left(\sqrt{1+\frac{2b\delta}{\lambda}k^2} - 1\right)\right]. \end{aligned} \tag{A.9}$$

This integral can be evaluated by a change of variables, $k \to k' = (1 + 2b\delta k^2/\lambda)^{1/2}$. The final result (for the natural logarithm) is

$$H[\boldsymbol{\sigma}] = \frac{1}{2}\ln b + \frac{\pi}{V_B}\left[2k_B^2\left(\frac{\ln k_B}{2} - \frac{1}{4}\right) + \frac{\lambda}{2b\delta}\left(\frac{h^2}{4} - h\ln h + h - \frac{h^2}{2}\ln h\right)\right] \tag{A.10}$$

whereby

$$h \equiv \sqrt{1 + \frac{2b\delta}{\lambda}k_B^2} \ - \ 1.$$

Again, this expression has to be modified in case $k_0 < k_B$. The redundancy is given by the difference of (A.8) and (A.10).

The conditional entropy $H[\boldsymbol{\sigma}|\boldsymbol{\xi}]$ can be calculated in an analogous way by evaluating an integral over the first Brillouin zone. We only give the final result for the case $k_0 > k_B$:

$$\begin{aligned} H[\boldsymbol{\sigma}|\boldsymbol{\xi}] = \frac{1}{2}\ln\frac{\delta}{2} &- \frac{\pi}{2V_B}\left[\frac{g}{\Delta}\left(1 + \frac{\Delta}{g}k_B^2\right)\ln\left(1 + \frac{\Delta}{g}k_B^2\right) - k_B^2\right] \\ &+ \frac{\pi}{2V_B}\,\frac{\lambda}{b\delta}\left[\frac{r^2}{2}\ln r - \frac{r^2}{4} - r\ln r + r - 1\right] \end{aligned} \tag{A.11}$$

where

$$r \equiv \sqrt{1 + \frac{2b\delta}{\lambda} k_B^2} + 1.$$

As always, equation (A.11) is given up to an additive constant. The mutual information is given by the difference of (A.9) and (A.11).

A.2. Fixing the phase by demanding local receptive fields

Given any filter gain $|G(\boldsymbol{k})|$ we want to choose a phase $f(\boldsymbol{k})$ in such a way that we obtain *local receptive fields* as the Fourier transform of $G(\boldsymbol{k}) = |G(\boldsymbol{k})| \exp[\mathrm{i} f(\boldsymbol{k})]$. First, we have to define how locality should be measured. An appropriate expression is given by

$$K = \int \mathrm{d}^2\boldsymbol{x}\ \boldsymbol{x}^2\ |G(\boldsymbol{x})|^2. \tag{A.12}$$

Large, long-ranging weights (i.e. large $|G(\boldsymbol{x})|$ for large $\boldsymbol{x}^2$) are unfavorable for this cost function, which should be minimized. In the Fourier domain the expression $\boldsymbol{x}G(\boldsymbol{x})$ is given by

$$\boldsymbol{x}\, G(\boldsymbol{x}) = \frac{1}{2\pi\mathrm{i}} \int \mathrm{d}^2\boldsymbol{k}\ \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\ \boldsymbol{\nabla} G(\boldsymbol{k}). \tag{A.13}$$

Thus for our cost function (A.12) we obtain

$$\begin{aligned} K &= \frac{1}{(2\pi)^2} \int \mathrm{d}^2\boldsymbol{x} \int \mathrm{d}^2\boldsymbol{k}\ \mathrm{d}^2\boldsymbol{k}'\ \mathrm{e}^{\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{x}}\ \boldsymbol{\nabla} G(\boldsymbol{k}) \left[\boldsymbol{\nabla} G(\boldsymbol{k}')\right]^* \\ &= \int \mathrm{d}^2\boldsymbol{k}\ \boldsymbol{\nabla} G(\boldsymbol{k})\, [\boldsymbol{\nabla} G(\boldsymbol{k})]^* \\ &= \int \mathrm{d}^2\boldsymbol{k}\ \left[|G(\boldsymbol{k})|^2 |\boldsymbol{\nabla} f(\boldsymbol{k})|^2 + |\boldsymbol{\nabla}|G(\boldsymbol{k})||^2\right]. \end{aligned} \tag{A.14}$$

Varying (A.14) with respect to $f(\boldsymbol{k})$, we see directly that the optimal phase is given by

$$\boldsymbol{\nabla} f(\boldsymbol{k}) = 0 \quad \Longrightarrow \quad f(\boldsymbol{k}) = \text{constant}. \tag{A.15}$$

In the space domain $G(\boldsymbol{x})$ has to be real and thus we obtain $\exp[-\mathrm{i} f(\boldsymbol{k})] = \pm 1$. These two solution are equivalent, and we may view these solutions as the *on*- and *off*-centre pathways of the visual system. We also note that we did not make any assumptions on $|G(\boldsymbol{k})|$. Whatever $|G(\boldsymbol{k})|$, the 'most local' receptive field therefore has zero phase (or phase π).

References

[1] Ash R B 1965 *Information Theory* (New York: Wiley) (reprinted 1990 (New York: Dover))

[2] Atick J J 1992 Could information theory provide an ecological theory of sensory processing? *Network: Comput. Neural Syst.* **3** 213–51

[3] Atick J J and Redlich N 1990 Towards a theory of early visual processing *Neural Comput.* **2** 308–20

[4] Atick J J and Redlich N 1992 What does the retina know about natural scenes? *Neural Comput.* **4** 196–210

[5] Attneave F 1954 Some informational aspects of visual perception *Physiol. Rev.* **61** 183–93

[6] Barlow H B 1961 Possible principles underlying the transformation of sensory messages *Sensory Communication* ed W A Rosenblith (Cambridge, MA: MIT Press) pp 217–34

[7] Barlow H B 1989 Unsupervised learning *Neural Comput.* **1** 295–311

[8] Barlow H B, Fitzhugh R and Kuffler S W 1956 Change of organization in the receptive fields of the cat's retina during dark adaption *J. Physiol.* **137** 338–54

[9] Bialek W and Ruderman D L 1994 Statistics of natural images: Scaling in the woods *Phys. Rev. Lett.* **73** 814–7

[10] Bialek W, Ruderman D L and Zee A 1991 Optimal sampling of natural images: A design principle for the visual system? *Advances in Neural Information Processing Systems 3* ed D Touretzky and J Moody (San Mateo, CA: Morgan Kaufmann) pp 363–9
[11] Boahen K A and Andreou A G 1992 A contrast sensitive silicon retina with reciprocal synapses *Advances in Neural Information Processing Systems 4* ed J E Moody, J Hanson and R P Lippmann (San Mateo, CA: Morgan Kaufmann) pp 764–72
[12] Bunday B D 1984 *Basic Optimization Methods* (London: Arnold)
[13] Campa A, Del Guidice P, Parga N and Nadal J 1995 Maximization of mutual information: a detailed study *Network: Comput. Neural Syst.* **6** 449–68
[14] Cover T M and Thomas J A 1991 *Elements of Information Theory* (New York: Wiley)
[15] De Valois R L, Morgan H and Snodderly D M 1974 Psychophysical studies of monkey vision III, Spatial luminance contrast sensitivity tests of macaque and human observers *Vision Res.* **14** 75–81
[16] de Vries H L 1943 The quantum character of light and its bearing on threshold of vision, differential sensitivity and visual acuity of the eye *Physica* **10** 553–64
[17] Derrington A M and Lennie P 1982 The influence of temporal frequency and adaption level on receptive field organization of retinal ganglion cells *J. Physiol.* **333** 343–66
[18] Derrington A M and Lennie P 1984 Spatial and temporal contrast sensitivities of neurons in the lateral geniculate nucleus of macaque *J. Physiol.* **357** 219–40
[19] Dong D W and Atick J J 1995 Statistics of natural time-varying images *Network: Comput. Neural Syst.* **6** 345–58
[20] Dowling J E 1987 *The Retina: An Approachable Part of the Brain* (Cambridge, MA: Harvard University Press)
[21] Enroth-Cugell C and Robson J G 1966 The contrast sensitivity of retinal ganglion cells of the cat *J. Physiol.* **187** 517–52
[22] Fain G L, Gold G H and Dowling J E 1976 Receptor coupling in the toad retina *Cold Spring Harbor Symposia on Quantitative Biology 40* pp 547–61
[23] Field D J 1987 Relations between the statistics of natural images and the response properties of cortical cells *J. Opt. Soc. Am.* A **4** 2379–94
[24] Guyton A C 1991 *Textbook of Medical Physiology* (Philadelphia, PA: Saunders) 8th edn
[25] Hammond P 1973 Contrasts in spatial organization of receptive fields at geniculate and retinal levels: center, surround and outer surround *J. Physiol.* **228** 115–37
[26] Hammond P 1974 Cat retinal ganglion cell: size and shape of receptive field centers *J. Physiol.* **242** 99–118
[27] Heeger J A, Heeger D J, Langan J and Yang Y 1995 Image enhancement with polymer grid triode arrays *Science* **270** 1642–4
[28] Ikeda H and Wright M J 1972 The outer disinhibitory surround of the retinal ganglion cell receptive fields *J. Physiol.* **226** 511–44
[29] Kaplan E and Lee B B 1990 New views of primate retinal functions *Progress in Retinal Reseach* vol 9, ed N Osborne and J Chader (Oxford: Pergamon) pp 273–336
[30] Kaplan E, Marcus S and So Yuen Tat 1979 Effects of dark adaption on spatial and temporal properties of receptive fields in cat lateral geniculate nucleus *J. Physiol.* **294** 561–80
[31] Kaplan E, Shapley R M and Purpura K 1989 Spatial and spectral mechanism of the primate retinal ganglion cell *Seeing Countour and Colour* ed J J Kulikowski (Oxford: Pergamon) pp 36–40
[32] Kelly D H 1977 Visual contrast sensitivity *Optica Acta* **24** 107–29
[33] Kersten D 1987 Predictability and redundancy of natural images *J. Opt. Soc. Am.* A **4** 2395–400
[34] Kirsch K, Wagner H and Djamgoz B A 1991 Dopamin and the plasticity of the horizontal cell function in the teleost retina: regulation of the spectral mechanism through D1-receptors *Vision Res.* **31** 401–12
[35] Kittel C 1987 *Quantum Theory of Solids* (New York: Wiley)
[36] Kretzmer E R 1952 Statistics of television signals *Bell Syst. Tech. J.* **31** 751–63
[37] Lamb T D and Simon J E 1976 The relation between intercellular coupling and the electrical noise in the turtle photoreceptors *J. Physiol.* **263** 257–86
[38] Lankheet M J M, Frens M A and van de Grind W A 1990 Spatial properties of the horizontal cell responses in the cat retina *Vision Res.* **30** 1257–75
[39] Lennie P 1993 Roles of the M and P pathways *Contrast Sensitivity* ed R Shapley and D M-K Lam (Cambridge, MA: MIT Press) pp 202–13
[40] Li Z 1992 Different retinal ganglion cells have different functional goals *Int. J. Neural Sys.* **3** 237–48
[41] Linsker R 1988 Self-organization in a perceptual network *Computer* **21** 105–17
[42] Linsker R 1989 An application of the principle of maximum information preservation to linear systems *Advances in Neural Information Processing Systems 1* ed D S Touretzkey (San Mateo, CA: Morgan

Kaufmann) pp 186–94
[43] Linsker R 1990 Perceptual neural organization: Some approaches based on network models and information theory *Ann. Rev. Neurosci.* **13** 257–81
[44] Linsker R 1993 Deriving receptive fields using an optimal encoding criterion *Advances in Neural Information Processing Systems 5* ed J Hanson, J Cowan and C L Giles (San Mateo, CA: Morgan Kaufmann) pp 953–60
[45] Liu Y and Shouval H 1994 Localized principal components of natural images—an analytic solution *Network: Comput. Neural Syst.* **5** 317–24
[46] Mach E 1865 Über die Wirkung der Raumlichen Vertheilung des Lichtreizes auf die Netzhaut I *Akademie der Wissenschaften in Wien, Sitzungsberichte, math. nat. Cl.* 52 Abt.2, 303–22
[47] Mahowald M and Mead C 1991 The silicon retina *Scientific American* **264** 40–6
[48] Mangel S C and Miller R F 1987 Horizontal cells contribute to the receptive field surround of ganglion cells in the rabbit retina *Brain Res.* **414** 182–6
[49] Miller W H and Bernard G D 1983 Averaging over the foveal receptor aperture curtails aliasing *Vision Res.* **23** 1365–9
[50] Naka K 1972 The horizontal cells *Vision Res.* **12** 573–88
[51] Nicholls J G, Martin A R and Wallace B G 1992 *From Neuron to Brain* (Sunderland, MA: Sinauer)
[52] Ruderman D L 1994 Designing receptive fields for highest fidelity *Network: Comput. Neural Syst.* **5** 147–55
[53] Ruderman D L 1994 The statistics of natural images *Network: Comput. Neural Syst.* **5** 517–48
[54] Ruderman D L and Bialek W 1994 Statistics of natural images: scaling in the woods *Advances in Neural Information Processing Systems 6* ed J Hanson, J Cowan and C L Giles (San Mateo, CA: Morgan Kaufmann) pp 551–58
[55] Shannon C E and Weaver W 1949 *The Mathematical Theory of Communication* (Urbana, IL: University of Illinois Press)
[56] Snyder A W, Laughlin S B and Stavenga D G 1977 Information capacity of the eye *Vision Res.* **17** 1163–75
[57] Srinivasan M V, Laughlin S B and Dubs A 1982 Predictive coding: a fresh view of inhibition in the retina *Proc. R. Soc.* B **216** 427–59
[58] Tolhurst D J, Movshon J A and Dean A F 1983 The statistical reliability of signals in single neurons in cat and monkey visual cortex *Vision Res.* **23** 775–85
[59] Troy J B and Enroth-Cugell C 1989 Signals and noise in the mammalian retina *Proc. IEEE Int. Conf. on Systems, Man and Cybernetics* (Piscataway, NJ: IEEE) pp 443–47
[60] van Hateren J H 1992 Theoretical predictions of spatiotemporal receptive fields of fly LMCs, and experimental validation *J. Comp. Physiol.* A **171** 157–70
[61] van Hateren J H 1992 A theory of maximizing sensory information *Biol. Cybern.* **68** 23–9
[62] van Hateren J H 1993 Spatiotemporal contrast sensitivity of early vision *Vision Res.* **23** 257–67
[63] Wandell B A 1995 *Foundations of Vision* (Sunderland, MA: Sinauer)
[64] Watson G N 1944 *A treatise on the Theory of Bessel Functions* (Cambridge: Cambridge University Press) 2nd edn
[65] Williams D R 1985 Aliasing in human foveal vision *Vision Res.* **25** 195–205
[66] Williams D R 1986 Seeing through the photoreceptor mosaic *Trends Neurosci.* **9** 193–8
[67] Williams D R and Coletta N J 1987 Cone spacing and visual resolution limit *J. Opt. Soc. Am.* **4** 1514–23
[68] Wolpert S and Micheli-Tsanakou E 1993 Silicon models of lateral inhibition *IEEE Trans. Neural Networks* **4** 955–61
[69] Wu C and Chiu C 1995 A new structure for a 2-D silicon retina *IEEE J. Solid State Circuits* **30** 890–97