PERGAMON

Vision Research 39 (1999) 3960–3974

Vision Research

www.elsevier.com/locate/visres

Section 4

Visual adaptation as optimal information transmission

Martin J. Wainwright *

Vision Science Laboratory, Harvard University, 33 Kirkland Street, Cambridge, MA 02138, USA

Received 24 September 1998; accepted 18 March 1999

Abstract

We propose that visual adaptation in orientation, spatial frequency, and motion can be understood from the perspective of optimal information transmission. The essence of the proposal is that neural response properties at the system level should be adjusted to the changing statistics of the input so as to maximize information transmission. We show that this principle accounts for several well-documented psychophysical phenomena, including the tilt aftereffect, change in contrast sensitivity and post-adaptation changes in orientation discrimination. Adaptation can also be considered on a longer time scale, in the context of tailoring response properties to natural scene statistics. From the anisotropic distribution of power in natural scenes, the proposal also predicts differences in the contrast sensitivity function across spatial frequency and orientation, including the oblique effect. © 1999 Elsevier Science Ltd. All rights reserved.

Keywords: Adaptation; Aftereffects; Signal-to-noise ratio; Sensitivity; Natural scenes

1. Introduction

Adaptation is a widespread property of the visual system (and more generally, of sensory systems), occurring in neurons ranging from the retina to the cortex. Given the prevalence of adaptation, an important theoretical question is its functional significance. While the functional importance of light adaptation is well-established (e.g. Shapley & Enroth-Cugell, 1984; Laughlin, 1989), there exist other forms of adaptation for which the analogous question remains unresolved. Examples include the motion, spatial frequency, and orientation aftereffects. In the motion aftereffect, after adapting to a moving stimulus, observers report the illusion of motion in a subsequently presented test stimulus that is physically stationary (Addams, 1834). A similar effect occurs in the orientation domain: adaptation to an inclined grating causes a vertically presented test grating to be perceived as tilted in the opposite direction (Gibson & Radner, 1937).

Traditionally, psychophysicists and physiologists have used adaptation as a tool to probe the structure of the visual system. For instance, it can be used to infer the encoding of a certain attribute, using logic along the following lines: if adaptation specific to an attribute can be demonstrated, then it can be concluded that the given attribute is encoded explicitly by the visual system. Blakemore and Campbell's (1969) classic work on spatial frequency and orientation channels includes examples of this use of adaptation, as does other work (e.g. Gilinksy, 1968; Pantle & Sekuler, 1968; Blakemore & Nachmias, 1971). In fact, with some exceptions (e.g. Barlow, MacLeod & van Meeteren, 1976), most psychophysical work has not explicitly addressed the possible functional significance of these curious phenomena.

Should such pattern aftereffects simply be dismissed as a consequence of poor design? Or can some functional significance be ascribed to these effects? In this paper, we explore the hypothesis that visual adaptation to orientation, spatial frequency and motion can be understood in terms of *the principle of maximizing information*. Our development is based on earlier work that applied ideas of information theory and efficient coding to early vision (e.g. Srinivasan, Laughlin & Dubs, 1982; Atick & Redlich, 1990; Linsker, 1992; van Hateren, 1992a), a portion of which we briefly review in Section 2. The principle of constrained information maximization is described in Section 3. Our proposal applies equally well

* Fax: +1-617-495-3764.
E-mail address: mjwain@mit.edu (M.J. Wainwright)

PII: S0042-6989(99)00101-7

to orientation, spatial frequency, or motion adaptation. However, due to the ready availability of both neurophysiological and psychophysical data addressing orientation coding, we first illustrate it by application to orientation adaptation in Section 4.

Such aftereffects are essentially a laboratory phenomenon, in which the presented signal has very simple statistics. Of more ecological relevance are the statistics of natural scenes, in which power is known to drop off roughly as f^{-2} (Field, 1987; van der Schaaf & van Hateren, 1996). Thus, it is also possible to consider adaptation as operating on a much longer time scale, in tailoring system response properties to the statistics of natural scenes. Along these lines, Section 5 establishes a link between the power spectrum of natural scenes, and a number of classic psychophysical phenomena (Campbell, Kulikowski & Levinson, 1966; Appelle, 1972), including the oblique effect.

Our proposal is similar in spirit but distinct from Barlow and Foldiak (1989) and Barlow (1990)'s decorrelation approach to adaptation. In particular, by explicitly considering noise, our proposal exhibits distinct regimes of behavior, depending on the signal-to-noise ratio (SNR). Links to other work, as well as directions for future work, are provided in Section 6.

2. Background

Natural signals are characterized by a high degree of redundancy, a property that can be exploited by efficient coding techniques. It was proposed by Attneave (1954), and independently by Barlow (1961a,b), that a major task of early vision is to remove the redundancy present in natural signals. Redundancy reduction is desirable, given the constraints on neural processing, because it allows sensory information to be transmitted and represented more efficiently. One technique to reduce redundancy is *predictive coding*, a technique initially studied in the context of TV transmission (Harrison, 1952). This technique exploits the autocorrelation function to estimate the signal at one point as a function of neighboring points. Only the error in this prediction is then transmitted which reduces the dynamic range occupied by the signal. In seminal work, Srinivasan et al. (1982) applied linear predictive coding to the center-surround antagonism commonly observed in peripheral visual pathways.

A strategy of pure redundancy reduction, however implemented, fails to take into account the noisiness in neural systems (Atick & Redlich, 1990; van Hateren, 1992a). In the presence of noise, pure redundancy reduction is no longer the optimal strategy. Rather, it is necessary to actively add redundancy in a controlled fashion in order to combat the noisiness of the system. More redundancy will be required where the signal-to-noise ratio (SNR) is low. Consequently, the optimal encoding strategy strikes a balance between removing and retaining redundancy, depending on the SNR. This optimal balance can be specified in *information-theoretic* terms (Shannon & Weaver, 1949). Both Atick and Redlich (1990, 1992) and van Hateren (1992a,b, 1993), in independent work, put forth theories of early visual processing formulated in terms of information theory. These information-theoretic proposals take into account the effects of noise, and are equivalent to the original proposal of redundancy reduction in the limit of no noise. Using the well-known f^{-2} drop-off of spectral power in natural scenes (Field, 1987; van der Schaaf & van Hateren, 1996), Atick and Redlich (1990) made predictions of ganglion cell receptive fields in the mammalian retina, which showed qualitative agreement with observed retinal ganglion cell properties under conditions of low and high SNR (Enroth-Cugell & Robson, 1966; Enroth-Cugell & Shapley, 1973). On the other hand, van Hateren (1992a,b) considered spatiotemporal power spectra; his theoretical predictions agreed qualitatively with the behavior of fly visual neurons with changing SNR. Subsequent work by both Atick and Redlich (1992) and van Hateren (1993) established links with mammalian contrast sensitivity.

3. Theory

The current paper builds on this previous work by applying the principle of information maximization to commonly observed forms of visual adaptation. The essence of our proposal is that computation should be tailored to the *changing* statistics of the input in an information-theoretically optimal manner. From this theoretical perspective, then, aftereffects are a reflection of optimal changes in response characteristics. The foundation of the proposal involves three elements. It is first necessary to characterize the image statistics. Included in this characterization should be consideration of noise contained in the signal itself. Secondly, it is necessary to specify how the system acts upon this noisy input. Neural systems are inherently noisy, meaning that this operation introduces an additional source of noise into the process, which also must be characterized. The third requirement is to establish the criterion of optimality, which specifies how the system behavior ought to change as a function of the signal statistics. Included in this specification of optimality is a constraint on the output of the system, corresponding to the fact that neural systems have limited dynamic ranges. We address each of these facets in turn, and finish by formulating the constrained optimization problem.

Our development is based on work reviewed in the previous section (e.g. Atick & Redlich, 1990; van Hateren, 1992a).

3.1. Characterization of the signal

A single static image of the world can be described by an intensity value $S(x, y)$ at every point (x, y). The visual environment as a whole can be viewed as an ensemble $\{\mathscr{S}\}$ of such images—that is, a stochastic process defined over the space (x, y). An important characteristic of natural images is that there are statistical relations among the light intensities in subsets of pixels. In terms of second-order characteristics, light intensities tend to be correlated over short distances in the image. These second-order relations are captured by the *spatial autocovariance* function $K((x, y), (x', y'))$, which gives the covariance between luminance values at image locations (x, y) and (x', y'). It is often assumed that the ensemble statistics are *stationary* (i.e. invariant to translations of the coordinate system), in which case the autocovariance function depends only the vector $(u, v) = (x - x', y - y')$. In this case, the autocovariance function can be written more simply as $K(u, v)$. In many cases, this assumption of stationarity is reasonable, because images are likely to be viewed from a range of positions, so that shifted versions of the image are equally likely.[1] In contrast, given our intention to address orientation coding, we do not assume *isotropy*—i.e. invariance to rotations of the coordinate system—as has been done in previous work on natural image statistics (e.g. Field, 1987; Atick & Redlich, 1990; van Hateren, 1992a).

With a stationary ensemble, an equivalent description of the second-order relations is in Fourier space, via the power spectral density or *power spectrum*. By the Wiener–Khinchine theorem, the power spectrum is equal to the Fourier transform of the autocovariance function (Papoulis, 1965). Thus, the visual environment will be characterized by an ensemble power spectrum in two-dimensions $S(f_x, f_y)$ This spectrum can also be expressed in polar frequency coordinates $S(f, \theta)$ where f is the radial spatial frequency, and θ is orientation.

The existence of noise sources is critical to the proposal. The first source is input noise: it consists of all upstream noise, photon shot noise, noise arising from phototransduction (e.g. Lillywhite & Laughlin 1979), as well as noise contributed by earlier neural processing. As with the image ensemble. this noise is characterized by its power spectrum N_1, which is assumed to be flat or white. This assumption of a flat power spectrum is made simply for convenience, and is not essential to the theory.

3.2. Characterization of the response

It is standard to characterize human psychophysical sensitivity by means of the contrast sensitivity function (Campbell & Robson, 1968). The contrast sensitivity function (CSF) captures the sensitivity of human observers to luminance-defined sinusoidal gratings of varying spatial frequency. In the current paper, we illustrate our proposal by applying it to both orientation and radial spatial frequency. It is therefore necessary to consider a two-dimensional CSF in polar co-ordinates, which specifies sensitivity as a function of both orientation (θ) and spatial frequency f. In applying the theory to the tilt aftereffect, we will ignore spatial frequency f, and consider a CSF that depends only on the orientation θ.

Although psychophysical experiments measure the overall shape of the CSF, this envelope is understood to be formed by an underlying set of channels tuned to both orientation and radial spatial frequency. For simplicity in the current work, the optimization argument is performed at the level of the overall shape, and therefore does not explicitly treat the tuning of the individual channels. The issue of individual channels is briefly discussed in Section 6. Note that our choice of a global criterion for optimization is motivated purely by analytic simplicity; this choice does not imply any global assumption about visual processing.

While the system's sensitivity to a given frequency and orientation is fixed, the system response itself will vary, due to the presence of intrinsic noise sources. Noisiness in the input signal itself corresponds to the first noise source (with power spectrum N_1), which was described in the section above. In addition, the channel, being a neural system, also operates in a noisy manner. This second noise source also characterized by its power spectrum N_2, is added to the response. The block diagram in Fig. 1 shows the structure of the system: a noisy signal $S + N_1$ is multiplied by the transfer function F, followed by the addition of channel noise N_2. For a more precise description of this response, see Appendix A.

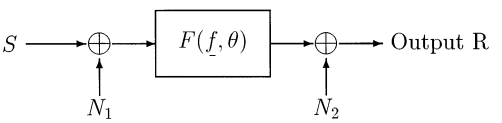


Fig. 1. Schematic block diagram of the system: input is the signal or image S which is contaminated by an additive noise source N_1. The system multiplies by the transfer function $F(f, \theta)$, and then a second noise source N_2 is added.

[1] It should be noted, as pointed out by one reviewer, that stationarity could be compromised by eye movements and other viewing strategies.

3.3. Specification of optimality

As described in Section 1, the optimal coding strategy strikes a balance between removing and adding redundancy. depending on the signal-to-noise ratio (SNR). This optimal balance is captured by the principle of maximizing the mutual information between the signal ensemble $\{\mathscr{S}\}$, and the response ensemble $\{\mathscr{R}\}$, in the presence of the noise sources N_1 and N_2. For each given envelope function F (as defined above), the response ensemble conveys a certain amount of information about the signal ensemble. Specifically, this quantity. known as the mutual information between response and signal, amounts to the reduction in uncertainty about the signal provided by the response. As derived in Appendix A, the rate of information transmission can be expressed as an integral in terms of the signal and noise power spectra:

$$\mathscr{I} = \int_{r_1}^{r_2}\int_0^{\pi} \log\left\{\frac{|F(f,\theta)|^2 \mathrm{S}(f,\theta) + |F(f,\theta)|^2 \mathrm{N}_1 + \mathrm{N}_2}{|F(f,\theta)|^2 \mathrm{N}_1 + \mathrm{N}_2}\right\} \times f\,\mathrm{d}f\,\mathrm{d}\theta \quad (1)$$

Thus given a certain signal and noise ensemble, we can ask what choice of envelope function F *maximizes* the quantity in Eq. (1). That is, we seek the envelope function F that minimizes the uncertainty about the signal given the response.

Given that neurons operate with limited dynamic ranges, the response R of the system must be bounded from above. An appropriate constraint is an upper bound on the response variance $\mathbb{E}[|R|^2]$, where $\mathbb{E}$ denotes expectation over the image ensemble. The following equation for this variance (see Appendix A) follows from Plancherel's formula:

$$\mathbb{E}[|R|^2] = \int_{r_1}^{r_2}\int_0^{\pi} \{|F(f,\theta)|^2[\mathrm{S}(f,\theta) + \mathrm{N}_1] + \mathrm{N}_2\} f\,\mathrm{d}f\,\mathrm{d}\theta \quad (2)$$

We require that this variance be bounded above as $\mathbb{E}[|R^2|] \leq M^2$, for some M. Since information transfer increases with increasing response variance, this condition is equivalent to requiring that the response variance be equal to M^2. It is worthwhile to note that such a constraint is not only biologically plausible, but also formally required in order for the problem to be well-defined. Without such a constraint, letting the envelope function F become arbitrarily large would render the channel noise N_2 irrelevant.

The problem is now one of constrained optimization: given certain signal and noise power spectra, choose the response envelope F so as to maximize the information flow (given by Eq. (1)) through the neural system shown in Fig. 1. subject to the bounded response variance (given by Eq. (2)). The reader can consult Appendix A for the derivation of this pair of Eqs. (1) and (2). Fortunately, the problem is analytically tractable; applying the calculus of variations (Gel'fand & Fomin, 1963) and a Lagrange multiplier λ yields the following explicit solution:

$$|F(f,\theta)|^2 = \left\{\frac{-\mathrm{N}_2[2\mathrm{N}_1 + \mathrm{S}(f,\theta)] + \sqrt{[\mathrm{N}_2\mathrm{S}(f,\theta)]^2 + 4/\lambda[\mathrm{N}_1\mathrm{N}_2 S(f,\theta)]}}{2\mathrm{N}_1[\mathrm{N}_1 + \mathrm{S}(f,\theta)]}\right\} \quad (3)$$

The value of the Lagrange multiplier λ needs to be determined numerically by substitution into Eq. (2).

Turning now to the proposal of interest, we apply this theory first to tilt adaptation in Section 4, and then to natural scenes in Section 5. In all of the following simulations, the height of the noise power spectra N_1 and N_2 are fixed at 0.025 and 0.05 respectively. The absolute values of these parameters are not critical; more relevant, as will be seen, are ratios.

For Section 4 on orientation adaptation, we integrate only over orientation. For natural scenes, integration over spatial frequency ranges from $r_1 = 0.25$ to $r_2 = 61$. As these ranges of integration differ between tilt adaptation and natural scene sections, the upper bound M is set to $20\sqrt{(\pi/180)}$ in Section 4 on orientation adaptation, and $10\,000\sqrt{(\pi/180)}$ in Section 5 on natural scenes. These choice of M give reasonable dynamic ranges for the given signal and noise power spectra; again, the precise choice of M is not critical.

4. Results: orientation adaptation

In a standard experiment on orientation adaptation, the human observer views a grating at a fixed orientation during an initial adaptation period that ranges from 30 seconds to several minutes. Subsequent psychophysical measurements (for example, of contrast thresholds, or orientation perception) are taken between intervening periods of top-up adaptation. Without loss of generality, we consider the case of adapting to a vertical (90°) grating.

We begin by assuming that under unadapted conditions, the power spectrum is flat across orientation. The visual system then updates its estimate of the power spectrum by a running average over time. It is this *covtinual updating* of the power spectral estimate that drives the process of adaptation. The actual input power spectrum during adaptation is (disregarding aperture effects) a delta function at 90°. However, the visual system can only update its estimate based on samples that have been smoothed by orientation-tuned cells. Therefore, during adaptation, the visual system receives samples (expressed in terms of Fourier amplitude) that are proportional to $\exp(-\theta^2/\sigma^2)$ where σ corresponds to the width of orientation tuning. The effect of adapting is to titrate the initially flat (i.e. 'unadapted') power spectrum with the spectrum of the

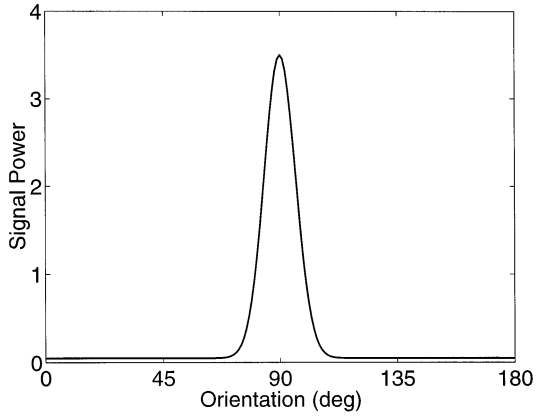


Fig. 2. A plot of power spectrum used to simulate orientation adaptation. It represents the power spectrum estimated by the visual system after adaptation, and it is given by Eq. (4) with $\sigma = 12°$, $a = 0.05$ and $b = 3.5$.

adapting ensemble. As adaptation proceeds, two changes will occur: first, the power spectrum will become increasingly peaked around the adapting orientation of 90°, and second, the power at orientations distant from the adapted location will become progressively weaker. Thus, the signal power spectrum during adaptation can be represented as follows:

$$S(\theta) = a + (b - a)\left[\exp\left(\frac{-\theta^2}{\sigma^2}\right)\right]^2 \quad (4)$$

The parameter σ, which is related to the width of orientation tuning, is set to 12° in all simulations. Such a choice of σ corresponds to an orientation bandwidth of roughly 20°, which is reasonable given psychophysical measurements of orientation bandwidths (e.g. Blakemore & Nachmias, 1971; Movshon & Blakemore, 1973). The parameter b, which determines the maximum height of the power spectrum at the adapted location (90°), will depend on both the length of the adaptation period as well as the contrast of the adapting grating. The level of residual power at non-adapted locations, given by the parameter a, will be related to the length of the adaptation period. These parameters will vary with experimental set-up: for our purposes, a precise choice is not critical. In all simulations, the parameters a and b were hand-chosen as 0.05 and 3.5 respectively, which yield reasonable agreement with the data. A plot of the power spectrum used to simulate the effects of adaptation is shown in Fig. 2. Given such a signal power spectrum, we then compute the optimal transfer function F according to Eq. (3).

4.1. Changes in contrast sensitivity

It is well-known that adaptation causes decrements in contrast sensitivity, or equivalently, increases in the contrast threshold required to detect a grating (e.g. Blakemore & Campbell, 1969). These changes are localized to the orientation (and spatial frequency) of the adapting stimulus. Elevation in contrast threshold is greatest when the orientation of the test grating matches the adapting orientation, and the size of the effect drops to half of this maximal value for test gratings at orientations $\approx \pm 8°$ away from the adapting orientation. For test gratings beyond $\pm 20°$ from the adapting location, contrast thresholds may actually decrease.

By the model specification. the response envelope provides a measure of the overall contrast sensitivity of the system. Given that contrast thresholds are simply the inverse of sensitivity, it is straightforward to calculate changes in contrast threshold following adaptation. Under 'unadapted' conditions, we assume that signal power is flat across orientation. Optimizing with respect to this flat power spectrum (of height fixed at 0.15) gives a baseline measure of unadapted thresholds. Then, optimizing with respect to a peaked power spectrum of the form shown in Fig. 2 yields adapted contrast thresholds, which are then normalized by the baseline to obtain percent changes in contrast threshold. Along the vertical axis of Fig. 3 are percent changes in contrast threshold, relative to the baseline ('unadapted') condition. Positive and negative numbers correspond to threshold elevation and reduction, respectively. Plotted in circles are contrast threshold data from Regan and Beverley (1985), measured following adaptation to a 12 cpd grating at 100% contrast. Also shown in a solid line are theoretical predictions, obtained from the optimized response envelope F. It is clear that the model predictions show good correspondence with the experimental data. The effect peaks at the adapted location (90°), where the threshold elevation is near twofold. On either side of the adapted

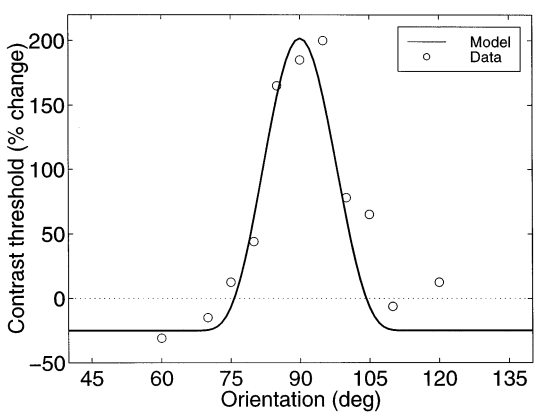


Fig. 3. Predicted changes in contrast thresholds following adaptation to a vertical (90°) grating. Vertical axis shows percent changes in contrast threshold relative to a baseline ('unadapted') condition. Theoretical predictions are given by the solid line, whereas experimental data. taken from Regan and Beverley (1985), are plotted in circles. See text for more details.

orientation, the effect drops off rapidly. For orientations beyond $\pm 20°$ from the adapted location, a slight reduction in contrast threshold (i.e. increase in sensitivity) is predicted. The Regan and Beverley (1985) data, due to asymmetry around the adapted orientation 90°, are not decisive in this regard. However, adaptation data from the analogous domain of spatial frequency (de Valois, 1977) support the notion that adaptation can actually increase contrast sensitivity away from the adapted frequency.

4.2. Tilt aftereffect

Perhaps the most striking effect of adapting to a grating is the tilt aftereffect (Gibson & Radner, 1937), in which the perceived orientation of a test grating is altered. For instance, after adapting to the vertical (90°), slightly off-vertical test gratings are perceived as being more strongly tilted. The direction of this effect is *repulsive*, in that the perception of gratings is pushed away from the adapted orientation symmetrically on both sides. The magnitude of this tilt aftereffect, though it varies with experimental conditions, is on the order of 3–4° at its peak.

It is commonly assumed that orientation is represented in a population code (e.g. Vogels, 1990), where the response-weighted average of the population determines the perceived orientation. With orientation θ_k in radians represented by the complex exponential $\exp(-2j\theta_k)$, the response-weighted vector $\boldsymbol{v}$ is calculated as a sum over the population of orientation units

$$\boldsymbol{v} = \sum_k r_k \exp(-2j\theta_k)$$

where r_k is the response used as a weight in the sum, and θ_k is the preferred orientation of unit k. Finally, the orientation is recovered as $\arctan[\mathrm{Im}(\boldsymbol{v})/\mathrm{Re}(\boldsymbol{v})]$. Assuming this type of coding, we use the optimized response envelope $F(\theta)$ as the response weighting in this calculation. Plotted in Fig. 4 with a solid line is the error in orientation perception predicted by the model, which corresponds to the tilt aftereffect. Also shown in circles are tilt aftereffect data taken from Campbell and Maffei (1971). Again, the model prediction captures the behavior of the data. The peak magnitude of the tilt aftereffect is ≈ 3–4°; it occurs at ± 8–10° from the adapted location and then fades off with $\pm 45°$.

4.3. Orientation discrimination

The main focus of the Regan and Beverley (1985) study was orientation discrimination, which they studied using a temporal two-alternative forced choice task. After adaptation to a vertical (90°) grating, they found that orientation discrimination was most impaired at $\pm 15°$ from the vertical. Interestingly, at the adapted

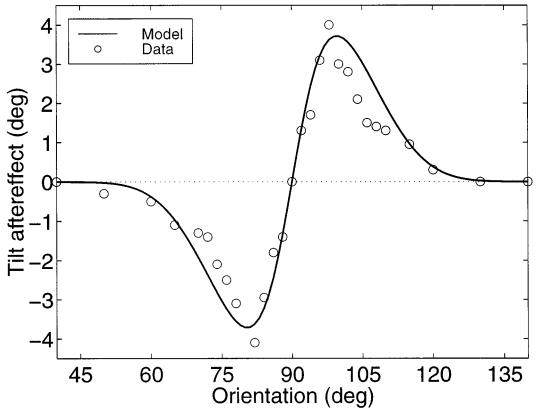


Fig. 4. Predicted tilt aftereffect following adaptation to a 90° grating. Plotted on the horizontal axis is the orientation of the test grating. The vertical axis shows the magnitude of the tilt aftereffect: that is, the difference between the perceived and presented orientations. Experimental data from Campbell and Maffei (1971) are plotted with open circles, whereas the model prediction is plotted with a solid line. See text for more details.

orientation, contrary to earlier results (Barlow et al., 1976), orientation discrimination actually improved by roughly 25%. The authors analyzed these results in the context of two opponent-process models for orientation discrimination. In both models, orientation-tuned units feed their outputs to an opponent unit. In the first model, the opponent unit computes the difference between the two outputs, whereas in the second case, it takes the ratio. Note that taking the ratio between two outputs is equivalent to taking the logarithmic difference. Consequently, the relative merits of these two models, as the authors point out. depend primarily on the contrast dependencies (e.g. saturating, proportional to log contrast, etc.) of the underlying orientation-tuned neurons.

Orientation discrimination thresholds are predicted using the following procedure. From the optimal response F, we used the response-weighted vector method (described in Section 4.2 on the tilt aftereffect) to calculate the curve describing orientation perception under the adapted condition. We then apply the linear difference model (described above) to the output of this orientation stage, which yields a vector of sensitivity changes after adaptation. To make comparisons with the baseline ('unadapted') data of Regan and Beverley (1985) meaningful, these raw values must be normalized, which is done so that the sensitivity changes at 90° are in correspondence. Finally, percent changes in discrimination thresholds are computed as a difference between the 'adapted' thresholds (theoretically predicted) and Regan and Beverley (1985) pre-adaptation measurements of thresholds, all divided by these same preadaptation measurements. Model predictions of threshold changes were symmetrized, in order to eliminate spurious effects of asymmetries in the pre-adaptation measurements.

Fig. 5 shows percent change in orientation discrimination threshold (relative to the unadapted baseline) on the vertical axis versus test orientation. Plotted on the left with circles joined with a dotted line are the measured changes in orientation discrimination threshold from Regan and Beverley (1985), whereas predictions from the theory are plotted on the right in triangles joined by a solid line. In both the experimental data and theory, orientation discrimination is most severely impaired at peaks on either side of the vertical, where it shows a roughly 50% elevation. This threshold elevation drops off by $\pm 30°$ away from vertical. Around the adapted location of 90°, orientation discrimination thresholds actually decrease by 25–30%.

5. Results: natural scenes

In recent years, a great deal of research has focused on first characterizing the statistics of natural images, and then relating these statistics to coding in the visual system. It is well-established that the power spectrum of natural images falls off as roughly f^{-2} with spatial frequency (e.g. Field, 1987; Ruderman, 1994; van der Schaaf & van Hateren, 1996), and this behavior has been used to explain various aspects of processing in both early insect vision (e.g. Srinivasan et al., 1982; van Hateren, 1992a,b), as well as early mammalian vision (e.g. Atick & Redlich, 1992; van Hateren, 1993; Dong & Atick, 1996). Similarly, anisotropies in the distribution of power across orientations have been documented (Switkes, Mayer & Sloan, 1978; van der Schaaf & van Hateren, 1996). Analysis of natural images reveals that greater amounts of power are concentrated at the vertical (90°) and horizontal (0°), relative to oblique (i.e. 45° and 135°) orientations. These differences are understandable, given the prevalence of vertical and horizontal structure in both carpentered and natural environments.[2] Despite this known anisotropy, relatively few solid links have been made between this anisotropic power distribution and the characteristics of early visual processing.

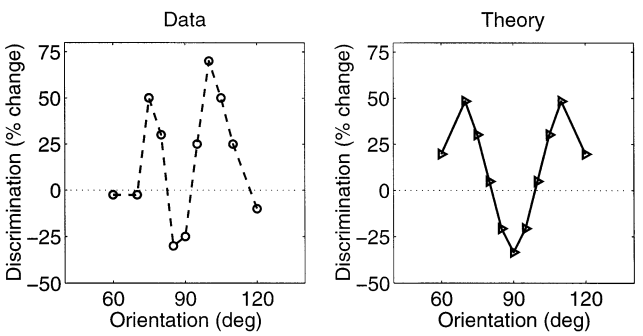


Fig. 5. Percent change in orientation discrimination thresholds as a function of orientation after adaptation to a vertical grating (90°). Plotted on the left in circles joined by a dotted line are experimental data, taken from Regan and Beverley (1985). Model predictions are plotted on the right in triangles joined by a solid line. Note that thresholds are lowered at the adapted orientation, and rise to peaks at $\approx +15°$ away from the vertical.

In this section, on the basis of measurements of the power spectrum of natural scenes, we calculate the optimal response envelope as a function of both radial frequency (f) and orientation (θ). This matching of response to natural scene statistics can be considered as adaptation operating on a long time scale. We show that this optimal response function accounts for several observed differences in contrast sensitivity across orientation and radial spatial frequency.

First of all, taking cross-sections across spatial frequency at a fixed orientation, the optimization predicts a sharper high-frequency fall-off in the CSF for oblique versus vertical/horizontal orientations, as has been demonstrated experimentally (Campbell et al., 1966). Secondly, taking cross-sections across orientation at fixed spatial frequencies reveals the oblique effect. That is, although contrast sensitivity across orientations is roughly constant at low spatial frequencies, greater sensitivity to vertical and horizontal is observed at high spatial frequencies (Campbell et al., 1966; Heeley & Timney, 1988).

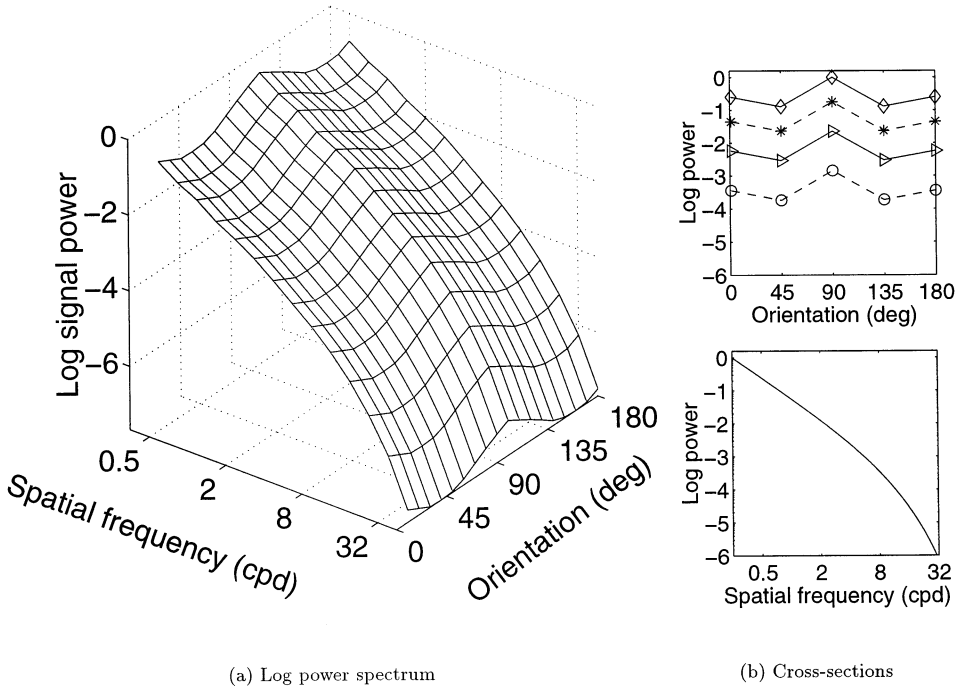
Using measurements of the power spectrum of natural scenes from van der Schaaf and van Hateren (1996), we formed the signal power spectrum as the product of the radial frequency and the orientation measurements. Following van Hateren (1993), we approximate the optics of the human eye, as measured by Campbell and Gubisch (1966), with a transfer function of the form

$$\sqrt{1-(f/f_{\max})}\,\exp[-(f/\mu f_{\max})]$$

where $\mu = 0.28$, and $f_{\max} = 61.2$ cpd. A mesh plot of this power spectrum (normalized) is shown in the left panel of Fig. 6, whereas the right-panel shows cross-sections through orientation (top) and through radial spatial frequency (bottom). The plot of log power versus log spatial frequency is linear through a significant range, reflecting the f^{-2} drop-off in power; deviations from linearity are caused by the transfer function of the optics of the eye.

Again, the optimal response envelope for this two-dimensional signal power spectrum was calculated using Eq. (3). By sectioning this response envelope through different axes, we can make predictions about differences in contrast sensitivity across orientation and radial spatial frequency.

[2] It should be noted that horizontal lines do not, in general, project horizontally onto the retina.



(a) Log power spectrum (b) Cross-sections

Fig. 6. (a) Log power spectrum of the signal, based on measurements of natural scenes by van der Schaaf and van Hateren (1996), as a function of radial spatial frequency and orientation. (b) Top: Cross-sections of this power spectrum as a function of orientation, taken at four fixed spatial frequencies. Note that log power simply shifts downwards with increasing radial spatial frequency. Bottom: Cross-section at a fixed orientation, showing the f^{-2} drop-off in power.

5.1. Contrast sensitivity function: high frequency fall-off

Campbell et al. (1966) measured the contrast sensitivity function (CSF) for horizontal (0°), vertical (90°), as well as two oblique orientations (45 and 135°). The basic finding was that contrast sensitivity at higher frequencies drops off earlier for oblique than for vertical or horizontal gratings. There is also a small difference between vertical and horizontal gratings, with greater sensitivity to the vertical than to the horizontal. This experimental data is plotted in the left panel of Fig. 7 on a log–linear scale in order to emphasize the high frequency behavior. Straight lines represent least squares fits to the data. Shown in the right panel are the CSFs predicted by the optimal response envelope. These predictions were obtained by sectioning the envelope along spatial frequency at the four orientations (0° for horizontal; 90° for vertical: and 45° and 135° for the two obliques). Again, curves are plotted on a linear spatial frequency axis in order to emphasize the high-frequency behavior of the CSFs. Since predictions for the two oblique orientations are identical, only one such curve is plotted. It is clear that the qualitative behavior of the Campbell et al. (1966) data has been captured.

5.2. Contrast sensitivity across orientations

Early psychophysical work (Campbell et al., 1966; Mitchell, Freeman & Westheimer, 1967) revealed an anisotropy in contrast thresholds across orientation. In particular, human observers show greater sensitivity to vertical and horizontal gratings compared to oblique gratings. This effect is a function of retinal orientation, and not of gravitational orientation (Banks & Stolarz, 1975). It is also known that the strength of this effect

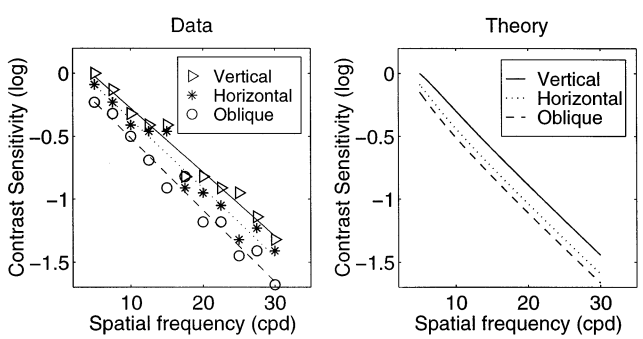


Fig. 7. Plotted in the left-hand panel are data from Campbell et al. (1966), of contrast sensitivity as a function of spatial frequency for three orientations. Straight lines are least squares fits to the data. The right-hand panel shows the drop-off in contrast sensitivity for three orientations as predicted by the theory. See text for more details.

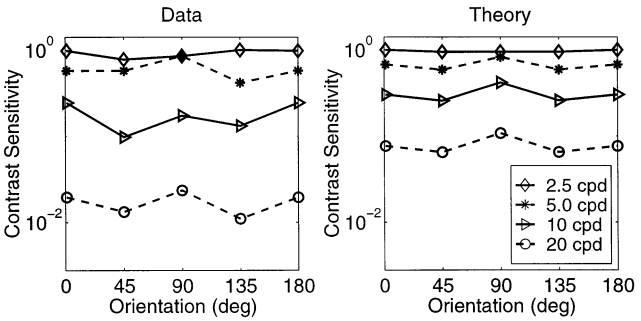


Fig. 8. The left-hand panel shows experimentally measured contrast sensitivity as a function of orientation for four spatial frequencies. Data is taken from Heeley and Timney (1988). The right-hand panel shows the corresponding contrast sensitivities, as predicted from the theory. These model predictions were based on measured properties of the power spectrum of natural scenes (van der Schaaf and van Hateren, 1996).

varies with spatial frequency: absent at low spatial frequencies, it appears only *beyond* the peak of the CSF, at roughly 4–8 cpd (Mitchell et al., 1967).

The data plotted on the left of Fig. 8, taken from Heeley and Timney (1988), are measurements of contrast sensitivity as a function of orientation for four spatial frequencies (2.5, 5, 10, and 20 cpd). The data has been re-scaled so that the maximum contrast sensitivity is unity. Note that contrast sensitivity is roughly constant for 2.5 cpd gratings, whereas an orientation anisotropy sets in for higher spatial frequency gratings. In particular, for high spatial frequencies, human observers show greater sensitivity to vertical (90°) and horizontal (0°) gratings. In the right panel of Fig. 8, we have plotted the normalized predictions of the model. In qualitative agreement with the data, the model predicts isotropic contrast sensitivity at 2.5 cpd, but an anisotropy for higher spatial frequencies.

Earlier attempts have been made to establish a link between the oblique effect and properties of the visual environment (for example, Annis & Frost, 1973). Interestingly, orientation anisotropy sets in only beyond the peak of the CSF (Mitchell et al., 1967: Berkeley, Kitterle & Watkins, 1975). Therefore, as noted by Switkes et al. (1978), because the anisotropy appears only at higher spatial frequencies, these earlier proposals rely on the assumption that horizontal and vertical orientations are disproportionately represented only at higher spatial frequencies. Based on their measurements of image power spectra, Switkes et al. (1978) found that such an assumption is not valid. The current model, by contrast, does not depend on such an assumption, and thus differs from previous proposals. Recall that the input power spectrum was composed as a separable product of measurements across orientation and spatial frequency (taken from van der Schaaf & van Hateren, 1996), so that anisotropy in the signal exists at all spatial frequencies. This fact is seen most clearly in the right-hand side of Fig. 6, where cross-sections at different radial spatial frequencies are seen to simply shift downwards on a logarithmic scale. According to the current proposal, the preference for vertical and horizontal only sets in beyond the peak of the CSF because moving to higher spatial frequencies corresponds to shifting to lower SNR. In this lower SNR regime, signal power at oblique orientations is affected more seriously by noise, and therefore is reduced relative to horizontal and vertical. We elaborate further on this important point in Section 6.

5.3. Other applications

Certain optical defects, among them astigmatism, will lead to differences in the statistics of the images presented to the visual system. For instance, consider an astigmatic subject in which horizontal gratings are defocused. For this subject. power at the horizontal (0°) should drop off more quickly than for a non-astigmatic observer. Therefore, the current proposal predicts reduced contrast sensitivity for horizontal gratings, and a corresponding decrease in the maximum resolvable spatial frequency at that orientation. These differences have been documented by Freeman and Thibos (1975) in an astigmatic subject in whom the horizontal was defocused.

It is a rarer case of astigmatism in which the principal optical meridians are oblique. In such a subject, the orientations of maximal and minimal defocus will lie on oblique axes. Therefore, depending on the degree of astigmatism, the current proposal predicts a reversal in the 'oblique effect'—namely, both the maximum and minimum resolvable frequencies should be observed at oblique orientations. This is in sharp contrast to the normal pattern, where the maximum and minimum are found at the vertical and oblique respectively. Freeman and Thibos (1975) also documented this interesting reversal in such an astigmatic subject.

6. Discussion

6.1. Two regimes of SNR

A comparison of the results on tilt adaptation and natural scenes reveals an interesting feature of the current proposal. First of all, in the context of tilt adaptation, the theory predicts that sensitivity should be *reduced* where the signal power is high. This reduction in sensitivity corresponds to an elevation in contrast threshold, which is experimentally observed (Regan & Beverley, 1985), as shown in Fig. 3. On the other hand, in Section 5 on natural scenes, the model shows the opposite effect. Specifically, it predicts that

for high spatial frequencies, contrast sensitivity should be *greatest* for orientations where the signal power is large (i.e. horizontal and vertical). This prediction, shown in Fig. 8, again is consistent with experimental measurements (Mitchell et al., 1967; Berkeley et al., 1975; Heeley & Timney, 1988).

These two types of behavior emerge, because there are two possible reasons for reducing sensitivity, each corresponding to a different regime of SNR. The first reason applies to regions of high SNR, where the signal is sufficiently robust that sensitivity can be reduced, thereby freeing up resources (i.e. the limited dynamic range) to be allocated more efficiently elsewhere. This reduction in sensitivity corresponds to a tendency towards flattening the power spectrum. (In the absence of noise, a flat power spectrum would be optimal for information transmission.) Sensitivity reduction in such a high SNR regime accounts for the low-frequency drop-off in the CSF as a function of radial frequency (Campbell & Robson, 1968), as has been pointed out previously (Atick & Redlich, 1992; van Hateren, 1993). In the context of the current work, behavior in tilt adaptation also corresponds to a high SNR regime, where the extended presentation of a high contrast grating represents an extremely high power signal. According to the current proposal, the reduction in contrast sensitivity observed following adaptation corresponds to the tendency to flatten the power spectrum, which is strongly peaked around the adapted orientation.

A second reason for reducing sensitivity corresponds to the low SNR regime where the signal begins to fall below the noise. In this regime, it becomes more efficient to reduce sensitivity to signals with low SNR, so that the associated noise does not occupy too much of the dynamic range. This type of sensitivity reduction is involved in the high-frequency drop-off of the CSF (Atick & Redlich, 1992; van Hateren, 1993). In the current proposal, it is this type of sensitivity reduction that underlies the orientation anisotropy in contrast sensitivity shown in Fig. 8. As radial spatial frequency is increased, the oblique orientations are swamped by noise before the vertical and horizontal orientations. According to the current proposal, therefore, at high spatial frequencies, sensitivity to oblique gratings should be reduced (relative to vertical and horizontal gratings).

6.2. Further predictions

With this basic intuition, we can turn to additional theoretical predictions that are not addressed by the data currently under consideration. First of all, for spatial frequencies lower than the peak of the CSF, the theory predicts an anisotropy in contrast sensitivity opposite to that of the oblique effect. In particular, contrast sensitivity should be *greater* for oblique than horizontal/vertical gratings.

Fig. 9 shows cross-sections through orientation (from the optimal response envelope obtained in Section 5) taken at 0.5 cpd and 20 cpd. Note the reversal in the direction of the orientation anisotropy in sensitivity between the lower and higher spatial frequency. The 20 cpd points show the usual oblique effect, whereas the 0.5 cpd points show the reversed pattern of greater sensitivity to oblique orientations.

The intuition underlying this prediction is that the peak of the CSF roughly marks a transition between higher and lower regimes of SNR. Frequencies higher than the peak correspond to a lower SNR regime, so that sensitivity should be reduced for the oblique orientations, where the signal is weaker. As discussed, this 'oblique' effect is experimentally observed (Campbell et al., 1966; Mitchell et al., 1966). On the other hand, spatial frequencies below the peak of the CSF correspond to a high SNR regime, where sensitivity should be reduced for the stronger signals—namely, the horizontal and vertical. In this regard, the data of Heeley and Timney (1988) plotted in Fig. 8 (left panel) are suggestive, but not conclusive, because sensitivity for very low spatial frequencies (i.e. <1 cpd) was not measured. Two caveats should be made regarding this prediction. First of all, it could be difficult to test due to well-known aperture issues in displaying very low frequency gratings. Secondly, the prediction was based on a power spectrum composed as a separable product of measurements across orientation and spatial fre-

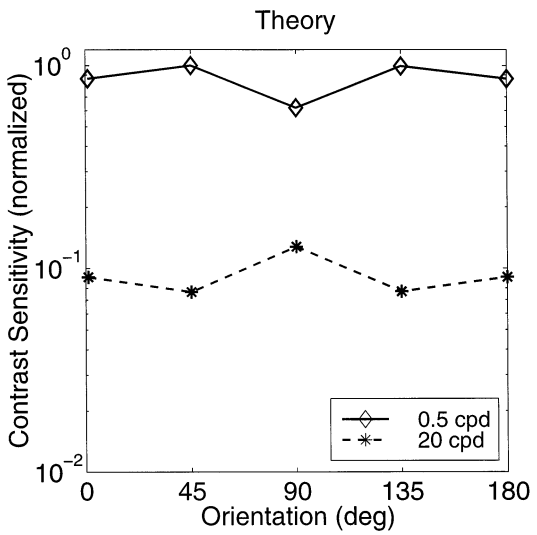


Fig. 9. Theoretical predictions of contrast sensitivity as a function of orientation at two spatial frequencies. Observe that the direction of the anisotropy reverses between the low and high spatial frequency cross-sections. The 20 cpd cross-section shows the usual oblique effect (Campbell et al., 1966; Mitchell et al., 1967), whereas the 0.5 cpd cross-section predicts that contrast sensitivity should be greater for the oblique orientations.

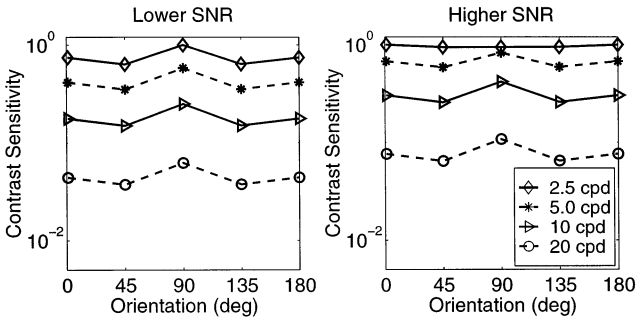


Fig. 10. Theoretical predictions of contrast sensitivity as a function of orientation at four spatial frequencies, for a lower SNR (left) and higher SNR (right). Note that the orientation anisotropy sets in earlier for the lower SNR than the higher SNR.

quency (van der Schaaf & van Hateren, 1996). Serious deviations from this separable form in the two-dimensional power spectra of natural scenes could invalidate this prediction.

The second prediction is related to the known changes in the shape of the CSF as a function of mean luminance. Specifically, peak sensitivity of the CSF shifts to lower spatial frequencies as mean luminance is reduced (e.g. Patel, 1966; Kelly, 1979). Recall that the model predicts that orientation anisotropy in contrast thresholds should set in only beyond the peak of the CSF. Since moving to lower mean luminance shifts the peak of the CSF, the theory predicts that orientation anisotropy should be observed at lower spatial frequencies, when tested under conditions of lower mean illumination. This prediction is illustrated in Fig. 10.

Both plots show contrast sensitivity across orientation at four spatial frequencies. The right-hand plot, a duplicate of the right plot in Fig. 8, displays theoretical predictions at a higher SNR. In this case, the orientation anisotropy is largely absent at the 1ower frequencies (2.5 and 5 cpd), but sets in with the shift to higher frequencies. The left-hand plot corresponds to an SNR lower by a factor of ≈ 4.5. In this lower SNR case, the usual oblique effect—greater sensitivity to vertical and horizontal—is seen more clearly even at the lower spatial frequencies. Again, the single caveat here is that our power spectrum was based on a separable product of measurements across orientation and spatial frequency.

6.3. Relation to previous work

The relation to earlier work on the oblique effect in natural scenes was addressed in Section 5.2. Laboratory aftereffects (e.g. spatial frequency or orientation aftereffects) have also been the focus of previous work. Ross and Speed (1991), studying spatial frequency adaptation, modeled it as a shift in the semi-saturation constant of a Naka–Rushton function. Similarly, Heeger (1992) accounted for adaptation effects by a shift of the semi-saturation constant in a divisive normalization operation. In a gain control network, Wilson and Humanski (1993) applied divisive operation with modifiable synapses to model a range of adaptation effects. Given their divisive (non-linear) nature, these models can be applied to a wider range of data than the model proposed here. The current theory differs primarily from this earlier work in its motivation. Whereas a significant thrust of this earlier work was towards understanding mechanism, the current proposal is motivated more by statistical goals. In the current work, we have shown that a number of adaptation effects (on both short and longer time scales) emerge in a principled fashion from a single goal—namely, that visual system responses are adjusted to the changing statistics of their input in an information-theoretically optimal manner.

In terms of its motivation, the current work is closest in spirit to an earlier proposal by Barlow and Foldiak (1989) that decorrelation among output units could be the principle underlying aftereffects. Whereas this work and Barlow's proposal share a similar motivation, they also differ in a number of ways. In the current proposal, the optimization was performed at a macroscopic level, in terms of an overall envelope of sensitivity. Due to this choice, made primarily for reasons of analytic simplicity, the tuning of individual output units (and therefore possible correlations among them) was not explicitly considered. Rather, the focus was on the balance between reducing and retaining redundancy at the image level, and how this optimal balance changes depending on the SNR. On the other hand, Barlow and Foldiak (1989) begin by assuming that the output units have equal variance, and then advocate complete decorrelation of these units. Thus, the proposal of Barlow and colleagues does not take into account noise. In contrast, by explicitly considering noise sources, the current proposal exhibits different behavior, even when the signal power spectrum is biased in the same way. These differences, which depend on the SNR, were discussed in detail in Section 6.1.

In terms of information transmission, explicit consideration of the tuning of output units leads to a constrained optimization problem more general than either the current approach, or the problem implicitly solved by the decorrelation approach of Barlow and colleagues. Furthermore, consideration of statistical dependencies beyond correlation will undoubtedly prove important (see Section 6.4). Due to considerations of analytic simplicity, we did not address these more general problems here. Statistical dependencies among output units are likely to be an important consideration, especially with adapting ensembles involving more than one stimulus type. Moreover, consideration of these relations is certainly interesting for reasons other than

information maximization. As emphasized by Barlow (1990), there exist other compelling rationales for decorrelation-like strategies, including the detection of novel associations. This topic thus remains a fertile ground for future research (see Wainwright & Simoncelli, 1999).

6.4. Future directions

In the current work, the signal (i.e. images) and the noise sources were characterized only by their power spectra. On one hand, the simplicity of this statistical characterization speaks to the power of the present proposal. Specifically, second-order statistics suffice to predict a number of differences in contrast sensitivity across orientation and spatial frequency, under both tilt-adapted conditions as well as natural scenes. Given only knowledge of power spectra, we followed the principle of least commitment (or maximum entropy (Wu, 1997)) in characterizing the signal and noise as Gaussian. However, previous work has shown that applying a wavelet transform (corresponding to cortical receptive fields) to natural images yields highly non-Gaussian marginal distributions (e.g. Field, 1987; Simoncelli & Adelson, 1996). The non-Gaussian nature of these marginals is one manifestation of higher-order statistical structure in natural images, not captured by the power spectrum. In a related issue, the current model is a linear approximation to a system well-known to be non-linear (e.g. Heeger, 1992; Tolhurst & Heeger, 1997). Non-linearities in the system, as well as the non-Gaussian nature of images, present further interesting questions for research. Such issues are addressed by work currently in progress (Wainwright & Simoncelli, 1999).

7. Conclusions

Building on previous work applying ideas of coding theory to the visual system (e.g. Srinivasan et al., 1982; Atick & Redlich, 1990; van Hateren, 1992a), we have proposed an information-theoretic model for adaptation at the cortical level. As a first illustration, we applied this proposal to orientation adaptation, and showed that it accounted for the standard effects, including changes in contrast thresholds, the tilt aftereffect, and changes in orientation discrimination thresholds. The effects of spatial frequency adaptation are analogous, in that similar changes in contrast threshold and shifts in perceived spatial frequency are well-documented (e.g. Blakemore & Campbell, 1969; de Valois, 1977). Although we have not presented such simulations in this paper, the current proposal could also be applied to the standard effects of spatial frequency and motion adaptation. In a second application, we considered adaptation operating on a longer time scale. On the basis of measurements of the power spectrum of natural images (van der Schaaf and van Hateren, 1996), we showed that the proposal accounts for various differences in contrast sensitivity across spatial frequency and orientation. We also set forth additional predictions, not currently addressed by data, about the shape of the CSF across orientation and radial frequency.

This work provides further support to the enterprise of deriving visual system response properties in a principled fashion—namely, from the statistics of its inputs. From the perspective advocated here, well-documented effects of visual adaptation are reflections of information-theoretically optimal changes in response properties. Thus, the current work simultaneously emphasizes the dynamic nature of visual response properties, and provides a computational rationale for these changes.

Acknowledgements

Financial support was provided by a 1967 Scholarship from the Natural Sciences and Engineering Research Council of Canada to MJW, and NIH EY09258 to PC. The author would like to thank Patrick Cavanagh, Ken Nakayama, and the Vision Sciences Laboratory at Harvard University for their support and guidance during completion of this work; Peter Dayan, Richard Kronauer, Markus Meister, Ikuya Murakami, Eero Simoncelli, and Charles Stromeyer for helpful comments on earlier drafts of this paper; and the anonymous reviewers for their constructive criticism and insight.

Appendix A

For the sake of completeness. we provide a rough outline of the derivation of Eqs. (1) and (2) for the information rate and response variance respectively, which together constitute the constrained optimization problem. As described in Section 3, our input consists of an ensemble of images $\{\mathscr{S}\}$, with an associated power spectral density $S(f_x, f_y)$. We can consider this power spectrum either in Cartesian coordinates (f_x, f_y), or polar coordinates (f, θ). A single image $S(x, y)$ can also be considered in this Fourier space, via its associated two-dimensional transform $\tilde{S}(f_x, f_y)$ or $\tilde{S}(f, \theta)$.

Since the system is assumed to act linearly, its operation on a single image can be characterized by its transfer function $F(f_x, f_y)$ as follows:

$$R(f_x, f_y) = F(f_x, f_y)\{\tilde{S}(f_x, f_y) + N_1(f_x, f_y)\} + N_2(f_x, f_y) \quad (5)$$

That is, frequency component by component, the response is as schematized in Fig. 1. The signal $\tilde{S}(f_x, f_y)$ contaminated by the first noise source N_1, is multiplied by the transfer function F, after which a second noise source N_2 is added.

Given the signal ensemble $\{\mathscr{S}\}$ and the noise sources, each choice of the transfer function F determines a response ensemble $\{\mathscr{R}\}$. Each such response ensemble conveys a certain amount of information about the signal. Specifically, the mutual information between response and signal is given by:

$$\mathscr{I}(\mathscr{S},\mathscr{R}) = \mathscr{H}(\mathscr{S}) - \mathscr{H}(\mathscr{S}|\mathscr{R}) = \mathscr{H}(\mathscr{R}) - \mathscr{H}(\mathscr{R}|\mathscr{S}) \quad (6)$$

where $\mathscr{H}$ is the entropy. This quantity captures the reduction in uncertainty about the signal provided by the response ensemble. We wish to choose our transfer function F (and thereby our response ensemble) so as to maximize this reduction in uncertainty.

As previously noted, this problem is not well-defined without a constraint on F; otherwise allowing the magnitude of F to become arbitrarily large would render the second noise source N_2 irrelevant. Two possible choices of this constraint are a 'hard' maximum on the amplitude of the response (i.e. $|R(f_x, f_y)| \leq M^2 \ \forall f_x, f_y$) or, alternatively, an upper bound on the response variance (i.e. $\mathbb{E}[|R|^2] \leq M^2$). Neurons rarely (if ever) operate at firing frequencies near the limits imposed by biophysical constraints, which are on the order of 1000 Hz. For this reason, the 'hard' maximum does not seem to be a reasonable choice. Rather, the second choice—that is, a constraint on response variance—seems most natural. This choice can be seen as a bound on the expected power of the neural response, which could be interpreted in terms of metabolic cost. An alternative interpretation follows from Chebyshev's inequality: a constraint on the response variance places a bound on the probability of large deviations in the response from its mean.

Therefore, we need to derive two equations: one for the mutual information between signal and response, and a second for the response variance. Both equations will be in terms of the amplitude of the transfer function $|F|$, and the power spectra of the signal and noise sources.

We begin by noting that Eq. (5) can be rewritten as

$$R(f_x, f_y) = \{F(f_x, f_y)\tilde{S}(f_x, f_y)\} + \{F(f_x, f_y)N_1(f_x, f_y) + N_2(f_x, f_y)\} \quad (7)$$

where the first term in curly braces corresponds to signal, and the second term corresponds to noise. Thus, for each frequency component the response is the sum of a signal term and a noise term, which are assumed to be uncorrelated. Therefore, the entropy of the response given the signal (i.e. $\mathscr{H}(\mathscr{R}|\mathscr{S})$) is equal to the entropy of the noise. Overall then, the mutual information is equal to the entropy of the response minus the entropy of the noise. For a Gaussian signal with variance σ^2, the entropy is given by (ignoring irrelevant constant terms) $\log(\sigma^2)$. Therefore, in our case, for a given frequency component, the mutual information is given by (after simplification):

$$\mathscr{I}(f_x, f_y) = \log\{\mathbb{E}[|R|^2]\} - \log\{\mathbb{E}[|F(f_x, f_y)N_1(f_x, f_y) + N_2(f_x, f_y)|^2]\} \quad (8)$$

where the argument of the second log corresponds to the variance of the noise term in Eq. (7). Expanding and simplifying, we find that the noise variance takes the form

$$|F(f_x, f_y)|^2\mathbb{E}[|N_1(f_x, f_y)|^2] + \mathbb{E}[|N_2(f_x, f_y)|^2]$$

whereas the response variance is equal to

$$\mathbb{E}[|R|^2] = |F(f_x, f_y)|^2\mathbb{E}[|\tilde{S}(f_x, f_y)|^2] + |F(f_x, f_y)|^2\mathbb{E}[|N_1(f_x, f_y)|^2] + \mathbb{E}[|N_2(f_x, f_y)|^2]$$

In both calculations, we have used the fact that signal and noise are assumed to be uncorrelated.

Normalizing the variances by the sampling areas (and eventually taking limits as the sampling size tends to infinity) will yield power spectral densities $\mathrm{S}(f_x, f_y)$, N_1, and N_2. Frequency dependence for the two noise power spectra are omitted, because they are assumed to be constant. Consequently, for a given frequency component, our expression for the mutual information between signal and response takes the form:

$$\mathscr{I}(f_x, f_y) = \log\left\{\frac{|F(f_x, f_y)|^2\mathrm{S}(f_x, f_y) + |F(f_x, f_y)|^2\mathrm{N}_1 + \mathrm{N}_2}{|F(f_x, f_y)|^2\mathrm{N}_1 + \mathrm{N}_2}\right\} \quad (9)$$

Now we have applied a linear shift-invariant operation to signal and noise sources that are stationary. Therefore, by the spectral representation theorem (Doob, 1953), frequency components in the response are uncorrelated. Therefore, since we are characterizing our signal and noise only by their power spectra, we can obtain the information between signal and response in a collection of frequency components by summing across them. That is, for some collection $\mathscr{C}$ of frequencies:

$$\mathscr{I}_{\mathscr{C}} = \sum_{\mathscr{C}} \mathscr{I}(f_x, f_y) = \frac{1}{\Delta f_x \Delta f_y} \sum_{\mathscr{C}} \mathscr{I}(f_x, f_y)\Delta f_x \Delta f_y \quad (10)$$

Recall that the spacing of frequency samples Δf *is* inversely proportional to the spatial distance over which the signal is sampled. So normalizing by the sampling area and taking limits allows us to pass to an integral, and we obtain an expression for the information rate in bits per degrees of visual angle squared:

$$\mathscr{T} = \int\int \log \left\{ \frac{|F(f_x,f_y)|^2 \mathrm{S}(f_x,f_y) + |F(f_x,f_y)|^2 \mathrm{N}_1 + \mathrm{N}_2}{|F(f_x,f_y)|^2 \mathrm{N}_1 + \mathrm{N}_2} \right\} \mathrm{d}f_x \, \mathrm{d}f_y \quad (11)$$

In similar fashion, working from steps above, we obtain that the overall response variance is given by:

$$\mathbb{E}[|R|^2] = \int\int \log \{|F(f_x,f_y)|^2 [\mathrm{S}(f_x,f_y) + \mathrm{N}_1] + \mathrm{N}_2\} \mathrm{d}f_x \, \mathrm{d}f_y \quad (12)$$

which is assumed to be bounded above by M^2.

These two equations constitute the constrained optimization problem. In order to solve it, we make use of the calculus of variations (Gel'fand & Fomin, 1963) with a Lagrange multiplier λ. In this way, the optimal response is found as a function of the Lagrange multiplier to be:

$$|F(f_x,f_y)|^2 = \frac{-\mathrm{N}_2[2\mathrm{N}_1 + \mathrm{S}(f_x,f_y)] + \sqrt{[\mathrm{N}_2 \mathrm{S}(f_x,f_y)]^2 + 4/\lambda [\mathrm{N}_1 \mathrm{N}_2 \mathrm{S}(f_x,f_y)]}}{2\mathrm{N}_1[\mathrm{N}_1 + \mathrm{S}(f_x,f_y)]} \quad (13)$$

The appropriate value of λ is obtained numerically so that the bound on the response variance in Eq. (12) is satisfied.

Note that we can change variables from Cartesian frequency coordinates (f_x, f_y) to polar frequency coordinates (f, θ), where f is the radial spatial frequency and θ is the orientation. In the section on natural scenes, we consider both radial spatial frequency and orientation, whereas in applying the theory to the tilt aftereffect, we consider only the orientation θ, and ignore the radial frequency. In the latter case, one should either think of pooling across a range of spatial frequencies at each orientation θ, or of taking a circular slice through the signal power spectrum at a fixed radial frequency.

References

Addams, R. (1834). An account of a peculiar optical phenomenon seen after having looked at a moving body. *London and Edinburgh Philosophical Magazine and Journal of Science*, *5*, 373–374.

Annis, R., & Frost, B. (1973). Human visual ecology and orientation anisotropies in acuity. *Science*, *182*, 729–731.

Appelle, S. (1972). Perception and discrimination as a function of orientation: the 'oblique effect' in man and animals. *Psychological Bulletin*, *78*, 266–278.

Atick, J., & Redlich, A. (1990). Towards a theory of early visual processing. *Neural Computation*, *2*, 308–320.

Atick, J., & Redlich, A. (1992). What does the retina know about natural scenes. *Neural Computation*, *4*, 196–210.

Attneave, F. (1954). Informational aspects of visual processing. *Psychological Review*, *61*, 183–193.

Banks, M., & Stolarz, S. (1975). The effects of head tilt on meridional differences in acuity: implications for orientation constancy. *Perception and Psychophysics*, *17*, 17–22.

Barlow, H. (1961a). Possible principles underlying the transformation of sensory messages. In W Rosenblith, *Sensory communication*. Cambridge MA: MIT Press.

Barlow, H. (1961b). Sensory mechanisms, the reduction of redundancy, and intelligence. In *Proceedings of the national physical laboratory symposium on the mechanisation of thought processes*. London, UK: H.M. Stationery Office.

Barlow, H. (1990). A theory about the functional role and synaptic mechanism of visual aftereffects. In C. Blakemore, *Vision: coding and efficiency*. Cambridge, UK: Cambridge University Press.

Barlow, H., & Foldiak, P. (1989). Adaptation and decorrelation in the cortex. In R. Durbin, C. Miall, & G. Mitchinson, *The computing neuron*. New York: Addison–Wesley.

Barlow, H., MacLeod, D., & van Meeteren, A. (1976). Adaptation to gratings: no compensatory advantages found. *Vision Research*, *16*, 1043–1045.

Berkeley, M., Kitterle, F., & Watkins, D. (1975). Grating visibility as a function of orientation and retinal eccentricity. *Vision Research*, *15*, 239–244.

Blakemore, C., & Campbell, F. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology*, *203*, 237–260.

Blakemore, C., & Nachmias, J. (1971). The orientation specificity of two visual after-effects. *Journal of Physiology*, *213*, 157–174.

Campbell, F., & Gubisch, R. (1966). Optical quality of the human eye. *Journal of Physiology*, *186*, 558–578.

Campbell, F., Kulikowski, J., & Levinson, J. (1966). The effect of orientation on the visual resolution of gratings. *Journal of Physiology* (*London*), *187*, 427–436.

Campbell, F., & Maffei, L. (1971). The tilt-aftereffect: a fresh look. *Vision Research*, *11*, 833–840.

Campbell, F., & Robson, J. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology* (*London*), *197*, 551–566.

de Valois, K. (1977). Spatial frequency adaptation can enhance contrast sensitivity. *Vision Research*, *17*, 1057–1065.

Dong, D., & Atick, J. (1996). Statistics of natural time-varying images. *Network*: *Computation in Neural Systems*, *6*, 345–358.

Doob, J. (1953). *Stochastic processes*. New York: Wiley.

Enroth-Cugell, C., & Robson, J. (1966). The contrast sensitivity of retinal ganglion cells of the cat. *Journal of Physiology*, *187*, 517–552.

Enroth-Cugell, C., & Shapley, R. (1973). Adaptation and dynamics of cat retinal ganglion cells. *Journal of Physiology*, *233*, 271–309.

Field, D. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A*, *4*, 2379–2394.

Freeman, R., & Thibos, L. (1975). Contrast sensitivity in humans with abnormal visual experience. *Journal of Physiology*, *247*, 687–710.

Gel'fand, I., & Fomin, S. (1963). *Calculus of variations*. Englewood Cliffs, NJ: Prentice-Hall.

Gibson, J., & Radner, M. (1937). Adaptation, after-effect and contrast in the perception of tilted lines: I. quantitative studies. *Journal of Experimental Psychology*, *20*, 453–467.

Gilinksy, A. (1968). Orientation-specific effects of patterns of adapting light on visual acuity. *Journal of the Optical Society of America.*, *58*, 13–18.

Harrison, C. (1952). Experiments with linear prediction in television. *Bell Systems Technology Journal*, *31*, 764–783.

Heeger, D. (1992). Normalization of cell responses in cat striate cortex. *Visual Neuroscience*, *9*, 181–197.

Heeley, D., & Timney, B. (1988). Meridional anisotropies of orientation discrimination for sine wave gratings. *Vision Research*, *28*, 337–344.

Kelly, D. (1979). Motion and vision II: Stabilized spatio-temporal threshold surface. *Journal of the Optical Society of America*, *69*, 1340–1349.

Laughlin, S. (1989). The role of sensory adaptation in the retina. *Journal of Experimental Biology*, *146*, 39–62.

Lillywhite, P., & Laughlin, S. (1979). Transducer noise in a photoreceptor. *Nature*, *277*, 569–572.

Linsker, R. (1992). Local synaptic learning rules suffice to maximize mutual information in a linear network. *Neural Computation*, *4*, 691–702.

Mitchell, D., Freeman, R., & Westheimer, G. (1967). Effect of orientation on the modulation sensitivity for interference fringes on the retina. *Journal of the Optical Society of America*, *57*, 246–249.

Movshon, J., & Blakemore, C. (1973). Orientation specificity and spatial selectivity in human vision. *Perception*, *2*, 53–60.

Pantle, A., & Sekuler, R. (1968). Size-detecting mechanisms in human vision. *Science*, *162*, 1146–1148.

Papoulis, A. (1965). *Probability*, *random variables and stochastic processes*. New York: McGraw–Hill.

Patel, A. (1966). Spatial resolution by the human visual system. *Journal of the Optical Society of America*, *56*, 689–694.

Regan, D., & Beverley, K. (1985). Postadaptation orientation discrimination. *Journal of the Optical Society of America*, *2*, 147–155.

Ross, J., & Speed, H. (1991). Contrast adaptation and contrast masking in human vision. *Proceedings of the Royal Society of London B*, *246*, 61–69.

Ruderman, D. (1994). The statistics of natural images. *Network*: *Computation in Neural Systems*, *5*, 517–548.

Shannon, C., & Weaver, W. (1949). *The mathematical theory of communication*. Urbana, IL: The University of Illinois Press.

Shapley, R., & Enroth-Cugell, C. (1984). Visual adaptation and retinal gain control. *Progress in Retinal Research*, *3*, 263–346.

Simoncelli, E., Adelson, E. (1996). Noise removal via Bayesian wavelet coring. In *Third International Conference on Image Processing*. Lausanne, Switzerland.

Srinivasan, M., Laughlin, S., & Dubs, A. (1982). Predictive coding: a fresh view of inhibition in the retina. *Proceedings of the Royal Society of London B*, *216*, 427–459.

Switkes, E., Mayer, M., & Sloan, J. (1978). Spatial frequency analysis of the visual environment: anisotropy and the carpentered environment hypothesis. *Vision Research*, *18*, 1393–1399.

Tolhurst, D., & Heeger, D. (1997). Cormparison of contrast-normalization and threshold models of the responses of simple cells in cat striate cortex. *Visual Neuroscience*, *14*, 293–309.

van der Schaaf, A., & van Hateren, J. (1996). Modelling the power spectra of natural images: statistics and information. *Vision Research*, *36*, 2759–2770.

van Hateren, J. (1992a). Real and optimal neural images in early vision. *Nature*, *360*, 68–69.

van Hateren, J. (1992b). Theoretical predictions of spatiotemporal receptive fields of fly LMCs, and experimental validation. *Journal of Comparative Physiology A*, *171*, 157–170.

van Hateren, J. (1993). Spatiotemporal contrast sensitivity of early vision. *Vision Research*, *33*, 257–267.

Vogels, R. (1990). Population coding of stimulus orientation by striate cortical cells. *Biological Cybernetics*, *64*, 25–31.

Wainwright, M. J., & Simoncelli, E. P. (1999). Explaining adaptation in V1 neurons with a statistically-optimized normalization model. *Investigative*, *Opthamology and Visual Science* (Suppl.), *40* (4), S573.

Wilson, H., & Humanski, R. (1993). Spatial frequency adaptation and contrast gain control. *Vision Research*, *33*, 1133–1149.

Wu, N. (1997). *The maximum entropy method*. New York: Springer.