Computational and Biological Mechanisms of Visual Coding

Edited by
M. Jenkin and L. Harris

Cambridge University Press, 1996

1

Modelling Binocular Neurons in the Primary Visual Cortex

David J. Fleet

David J. Heeger

Hermann Wagner

1.1 Introduction

Neurons sensitive to binocular disparity have been found in the visual cortex of many mammals and in the visual wulst of the owl, and are thought to play a significant role in stereopsis [Barlow et al., 1967, Nikara et al., 1968, Hubel and Wiesel, 1970, Clarke et al., 1976, Pettigrew and Konishi, 1976, Poggio and Fischer, 1977, Fischer and Kruger, 1979, Ferster, 1981, Poggio and Talbot, 1981, Ohzawa and Freeman, 986a, Ohzawa and Freeman, 986b, LeVay and Voigt, 1988, Ohzawa et al., 1990, DeAngelis et al., 1991, Wagner and Frost, 1993]. A number of physiologists have suggested that disparity might be encoded by a shift of receptive-field position [Hubel and Wiesel, 1962, Pettigrew et al., 1968, Pettigrew, 1972, Maske et al., 1984, Poggio et al., 1985, Wagner and Frost, 1993]. According to this *position-shift model*, disparity selective cells combine the outputs of similarly shaped, monocular receptive fields from different retinal positions in the left and right eyes. More recently, Ohzawa et al. (1990) and DeAngelis et al. (1991, 1995) have suggested that disparity sensitivity might instead be a result of interocular phase shifts. In this *phase-shift model*, the centers of the left- and right-eye receptive fields coincide, but the arrangement of receptive field subregions is different.

This chapter presents a formal description and analysis of a binocular energy model of disparity selectivity. According to this model, disparity selectivity results from a combination of position-shifts and/or phase-shifts. Our theoretical analysis suggests how one might perform an experiment to estimate the relative contributions of phase and position shifts to the disparity selectivity of binocular neurons, based on their responses to drifting sinusoidal grating stimuli of different spatial frequencies and disparities.

We also show that for drifting grating stimuli, the binocular energy response (with phase and/or position shifts) is a sinusoidal function of disparity, consistent with the physiology of neurons in primary visual cortex (area

17) of the cat [Freeman and Ohzawa, 1990]. However, Freeman and Ohzawa (1990) also found that the depth of modulation in the sinusoidal disparity tuning curves was remarkably invariant to interocular contrast differences. This is inconsistent with the binocular energy model.

As a consequence we propose a modified binocular energy model that incorporates two stages of divisive normalization. The first normalization stage is monocular, preceding the combination of signals from the two eyes. The second normalization stage is binocular. Our simulation results demonstrate that the normalized binocular energy model provides the required stability of the depth of response modulation. Simulations also demonstrate that the model's monocular and binocular contrast response curves are consistent with those of neurons in primary visual cortex.

1.2 Binocular Interaction and Disparity Selectivity

There are two major classes of neurons in primary visual cortex (V1 of the monkey or A17 of the cat), namely, simple cells and complex cells [Hubel and Wiesel, 1962]. Both types are selective for stimulus position and orientation. They respond vigorously to stimuli of a preferred orientation, but less well or not at all to stimuli of other orientations. Many neurons are also disparity selective.

Disparity-sensitive cells are often divided into four types: tuned-excitatory, tuned-inhibitory, near and far [Poggio and Fischer, 1977]. Disparity selectivity in these different types might arise from different mechanisms [Poggio and Fischer, 1977, Ferster, 1981] (but see [Nomura et al., 1990] for the opposite point of view). Tuned-inhibitory, near and far cells usually receive a strong excitatory input from one eye and an inhibitory input from the other eye (i.e., the monocular inputs are unbalanced), and most of them do not show binocular facilitation. Tuned-excitatory cells show a sharp response peak due to binocular facilitation, the responses at disparities flanking the peak are often inhibited, and they have balanced monocular inputs. This chapter is primarily concerned with tuned-excitatory complex cells, although all of the analysis and conclusions could be applied to tuned-excitatory simple cells as well.

1.2.1 Linear Neurons and Energy Neurons

There is a long tradition of modeling simple cells as *linear neurons* [Hubel and Wiesel, 1962, Campbell et al., 1968, Campbell et al., 1969, Movshon et al., 978a, Ohzawa and Freeman, 986a, Hamilton et al., 1989]. This model is attractive because a linear neuron can be characterized with a relatively small number of measurements.

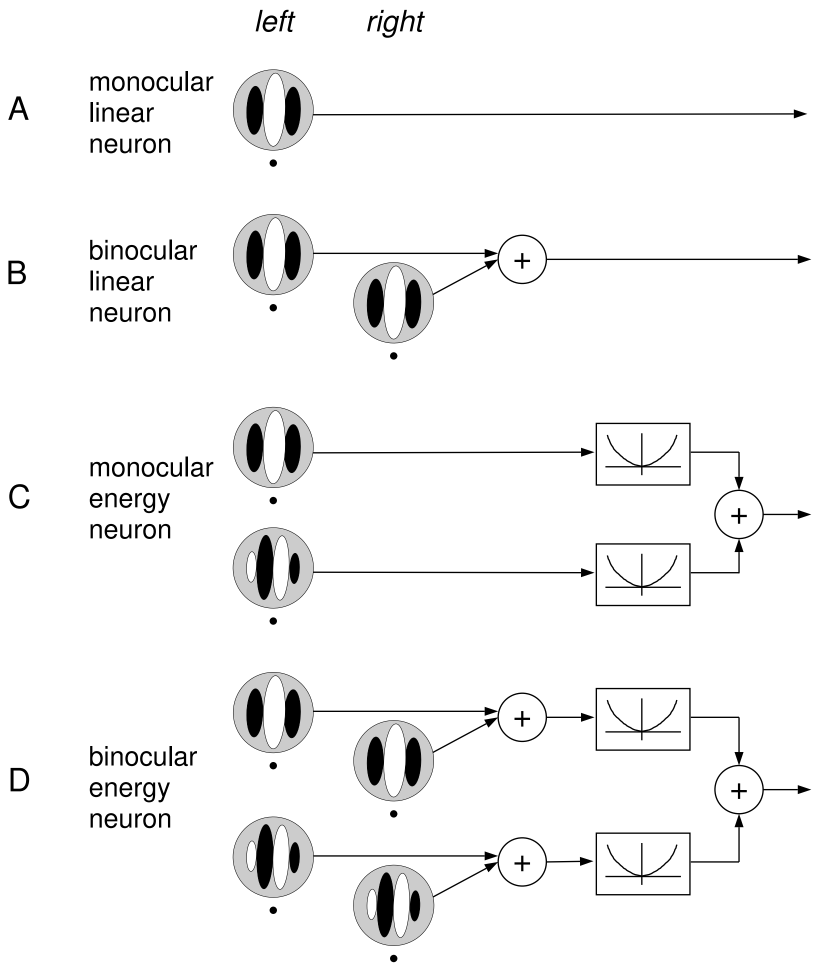

Fig. 1.1. Receptive fields of model neurons. (A) A monocular, orientation selective, linear neuron. Its response depends on a weighted sum of the stimulus intensities within its receptive field. Shaded ellipses depict inhibitory subregions of the weighting function, and the unshaded ellipse depicts an excitatory subregion. (B) A binocular linear neuron's response depends on a weighted sum of the stimulus intensities presented to both eyes. The reference points (black dots) below the weighting functions indicate that the two weighting functions are in exact binocular correspondence. (C) A monocular energy neuron sums the squared responses of two monocular, linear neurons. The weighting functions of the two linear neurons are identical except for a 90° phase shift. (D) A binocular energy neuron sums the squared responses of two binocular linear neurons. All four linear weighting functions are centered in exact (monocular and binocular) retinal correspondence.

Figure 1.1(A) shows a schematic diagram of a monocular linear neuron. A linear neuron's response is a weighted sum of stimulus intensities within a small region of the entire visual field, called the neuron's *receptive field.* In the illustration, the three ellipses depict subregions of the receptive field, one with positive weights (the unshaded ellipse), and two with negative weights (the shaded ellipses). The neuron is excited when a bright light is flashed in the positive subregion, and inhibited when a bright light is flashed in a negative subregion. Bright lights flashed simultaneously in both positive and negative subregions tend to cancel. The positive and negative weights are balanced so the neuron does not respond to blank stimuli. Rather, its response is proportional to stimulus contrast, for patterned stimuli that vary in intensity over space and time.

Figure 1.1(B) depicts a binocular linear neuron. This neuron's response depends on a weighted sum of the stimulus intensities presented to both eyes. The left- and right-eye receptive fields are identical for the neuron depicted in the figure, but this need not be the case in general. Also, the left- and right-eye receptive fields of this particular linear neuron are in exact binocular correspondence as indicated by the small reference points below the weighting functions.

One problem with the linear model of simple cells is that linear neurons can have negative responses because they sum input intensities using both positive and negative weights. However, extracellular responses (firing rates) of real neurons are, by definition, positive. Neurons with a high maintained firing rate could encode positive and negative values by responding either more or less than the maintained rate. But simple cells have very little maintained discharge. Instead, positive and negative values may be encoded by two neurons, one responsible for the positive part and one for the negative part. The two neurons are complements of one another; an excitatory subregion of one neuron's receptive field is aligned with an inhibitory subregion of the other neuron's receptive field. The response of each neuron is halfwave-rectified so that only one of the two neurons has a non-zero response at any given time. Simple cells are often characterized as halfwave-rectified linear neurons (e.g., [Movshon et al., 978a, Heeger, 992b]).

Complex cells do not have discrete ON and OFF receptive field subregions, and have been modeled as *energy neurons* [Adelson and Bergen, 1985, Emerson et al., 1992, Heeger, 992b, Pollen and Ronner, 1983]. An energy neuron sums the squared responses of a quadrature pair of linear neurons that are 90° out of phase, but with otherwise identical tuning properties (Fig. 1.1C). Equivalently, an energy neuron could sum the squared responses of four halfwave-rectified, linear neurons.

The monocular energy neuron depicted in Fig. 1.1(C) has one linear subunit that is even-symmetric (even phase) and another that is odd-symmetric (odd phase), but this is not necessary. The critical property is that the two subunits must be in quadrature phase (90° phase shift). Although simple cell weighting functions are not necessarily even- or odd-symmetric [Field and Tolhurst, 1986, Heggelund, 1986, Jones and Palmer, 987a], the receptive fields of adjacent simple cells tend to exhibit 90° or 180° phase relationships [Foster et al., 1983, Liu et al., 1992, Palmer and Davis, 1981, Pollen and Ronner, 1981]. A local pool of simple cells thus provides the right combination of signals for an ideal energy neuron. Approximately the same behavior may be obtained by summing the squared responses of many linear neurons (or halfwave-rectified, linear neurons), regardless of their phase, but with receptive fields distributed over a local spatial region.

A binocular energy neuron [Ohzawa et al., 1990] is depicted in Fig. 1.1(D). This neuron sums the squared responses of a quadrature pair of binocular linear neurons. This chapter is primarily concerned with the behavior of binocular energy neurons.

1.2.2 Disparity Selectivity: Position Shifts and Phase Shifts

Figure 1.2 depicts two ways that non-zero disparity preferences have been introduced in models of disparity selectivity. The neuron depicted in Fig. 1.2(A) is tuned for zero disparity because the locations of the monocular receptive fields are in exact binocular correspondence (indicated relative to the reference points) and the two pairs of weighting functions are identical. A non-zero disparity preference is introduced either by shifting the receptive field positions (Fig. 1.2B) or the receptive field phases (Fig. 1.2C). Both of the neurons in Fig. 1.2(B,C) are constructed to prefer uncrossed disparities; to evoke a maximal response, a visual feature (line, edge, grating) should be presented to the right eye in a position that is slightly shifted to the right.

1.3 Formalizing the Model

In order to examine the behavior of the model in detail, we derive formulas for the responses of these model neurons. A table of symbols (table 1.1) is provided to help the reader keep track of mathematical notation. In this chapter, we concentrate on the special case of drifting sinusoidal grating stimuli. The general case is discussed in detail in [Fleet et al., 1996]. We begin with linear neurons and with binocular energy neurons tuned for zero disparity. Then we introduce position-shifts and phase-shifts.

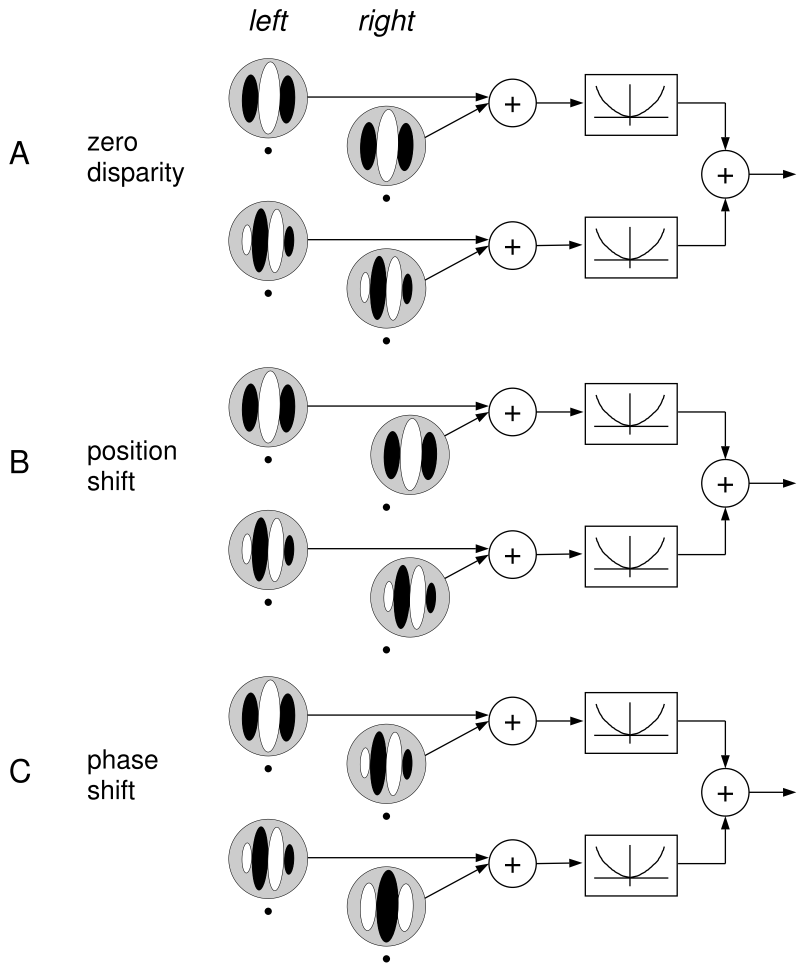

Fig. 1.2. Disparity preferences of binocular energy neurons. (A) Zero disparity preference. (B) Non-zero disparity preference is introduced by shifting the positions of both right-eye receptive fields by the same amount (relative to the reference points). (C) Non-zero disparity preference is introduced by shifting the phases of both right-eye weighting functions, in this case by 90°.

1.3.1 Monocular Linear Neurons

A visual neuron is linear (obeying the superposition property) if and only if its response is a weighted sum of the stimulus intensities. Mathematically, the response of a monocular linear neuron is the inner product in space and the convolution in time of a stimulus with the neuron's spatiotemporal weighting function. For a left-eye neuron with a weighting function

Terms	Definitions
$c_l\,, c_r$	left- and right-eye contrasts of a sinusoidal grating stimulus
$\vec{\omega} = (\omega_x, \omega_y, \omega_t)$	spatiotemporal frequency of a sinusoidal grating stimulus
$L_0\,,\; L_{90}$	responses of quadrature, monocular (left-eye) linear neurons
$R_0\,,\; R_{90}$	responses of monocular (right-eye) linear neurons
$A \equiv A(\vec{\omega})$	amplitude of linear neuron's transfer function
$\rho_l = c_l A$	response amplitude of a monocular (left-eye) linear neuron
θ	response phase of a linear neuron
$B_0\,,\; B_{90}$	responses of binocular linear neurons
s	receptive field position-shift
ψ	receptive field phase-shift
d	stimulus disparity
$E(d)$	binocular energy response at retinal position x
$E(d;\, s)$	response of binocular energy neuron with RF position-shift s
$E(d;\, \psi)$	response of binocular energy neuron with RF phase-shift $\Delta\psi$
$E(d;\, s,\, \psi)$	response of binocular hybrid energy neuron with position-shift s and phase-shift ψ

Table 1.1. *Symbol Table.*

$f_l(x,y,t)$, with stimulus $I_l(x,y,t)$, the response is given by

$$L(t) \;=\; \int\!\!\int\!\!\int_{-\infty}^{\infty} f_l(x,\, y,\, \tau)\, I_l(x,\, y,\, t-\tau)\; \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}\tau\ . \tag{1.1}$$

This triple integral is simply a weighted sum of the stimulus intensities over a small spatial neighbourhood and recently past time. The output response waveform, $L(t)$, is the model equivalent of a post-stimulus time histogram (PSTH), a measure of a cell's average response per unit time.

A linear neuron can be characterized by its transfer function (i.e., the Fourier transform of its weighting function). The transfer function has two components, amplitude $A(\vec{\omega})$ and phase $\theta(\vec{\omega})$, for each stimulus frequency. Here, $\vec{\omega} = (\omega_x, \omega_y, \omega_t)$ denotes the spatiotemporal frequency variables, with spatial frequency in cycles/degree and temporal frequency in Hz.

Consider a drifting sinusoidal grating stimulus,

$$I_l(x,\, y,\, t) \;=\; c_l\, \sin(2\pi\omega_x x \,+\, 2\pi\omega_y y \,+\, 2\pi\omega_t t)\ , \tag{1.2}$$

where c_l is the contrast of the grating. A linear neuron's response to a sinusoidal grating modulates sinusoidally over time with the same temporal frequency as the stimulus. This can be written as

$$L(t) = \rho_l(\vec{\omega})\, \sin[2\pi\omega_t t + \theta(\vec{\omega})], \tag{1.3}$$

where $\rho_l(\vec{\omega})$ is the response amplitude (peak height), and $\theta(\vec{\omega})$ is the response phase (relative peak latency).

Response amplitude depends on the amplitude component of the transfer function, $\rho_l(\vec{\omega}) = c_l A(\vec{\omega})$. Response phase depends on the phase component of the transfer function, $\theta(\vec{\omega})$, and on the receptive field position relative to the starting position of the drifting grating. A pair of monocular linear neurons with identical weighting functions, but at different receptive field positions, will respond with different phases. For example, let θ be the response phase of a neuron for a sine-grating stimulus with spatial frequency $\omega = \sqrt{\omega_x^2 + \omega_y^2}$. The response phase of a similar neuron that is a distance of s away (in the direction orthogonal to the stimulus orientation), would then be $\theta - 2\pi\omega s$. This phase behaviour is important below where position-shifts are used to introduce non-zero disparity tuning in binocular neurons.

To simplify notation we use ρ, A and θ in the equations below, dropping the explicit dependence on $\vec{\omega}$. It is important to remember that response amplitude and phase depend on the stimulus spatiotemporal frequency.

1.3.2 Monocular Energy Neurons

The monocular energy model depicted in Figure 1.1C consists of quadrature pairs of monocular linear neurons. These pairs of linear neurons have identical response amplitudes, but their response phases differ by 90 degrees. For example, the responses of a quadrature pair of monocular (left-eye) linear neurons can be expressed as:

$$\begin{aligned} L_0 &= \rho_l \sin(2\pi\omega_t t + \theta) \\ L_{90} &= \rho_l \cos(2\pi\omega_t t + \theta). \end{aligned}$$

The response amplitude, $\rho_l = c_l A$, is the same for both neurons, but the response phases differ by 90 degrees. Note that we have ignored the dependence of the response on time for notational simplicity.

1.3.3 Binocular Linear Neurons

A binocular linear neuron computes a weighted sum of the stimulus intensities presented to both eyes. It has two weighting functions, one for each eye, as depicted in Figure 1.1B. For now, we will assume that the two monocular weighting functions are identical and in exact binocular correspondence. We can express the responses of a pair of binocular linear neurons as follows:

$$\begin{aligned} B_0 &= L_0 + R_0 \\ B_{90} &= L_{90} + R_{90} \end{aligned}$$

where L_0, R_0, L_{90}, and R_{90} are responses of the monocular linear subunits.

1.3.4 Binocular Energy Neurons

A binocular energy neuron (Figure 1.1D) sums the squared responses of a quadrature pair of binocular linear neurons:

$$\begin{aligned} E &= B_0^2 + B_{90}^2 \\ &= (L_0 + R_0)^2 + (L_{90} + R_{90})^2 \\ &= L_0^2 + L_{90}^2 + R_0^2 + R_{90}^2 + 2R_0L_0 + 2R_{90}L_{90} \ . \end{aligned} \tag{1.4}$$

For now, let us assume that the drifting sinusoidal grating has a disparity of zero, so that it drifts over the same retinal positions in both eyes simultaneously. Let us further assume that the left and right monocular weighting functions are identical and in exact binocular correspondence (as depicted in Figures 1.1D and 1.2A). Then the left and right monocular responses are equal to one another, and equation (1.4) reduces to

$$E = \rho_l^2 + \rho_r^2 + 2\rho_l\rho_r \ , \tag{1.5}$$

where $\rho_l^2 = L_0^2 + L_{90}^2 = c_l^2A^2$ is the monocular left-eye energy, and $\rho_r^2 = R_0^2 + R_{90}^2 = c_r^2A^2$ is the right-eye energy. The binocular energy response clearly depends on stimulus contrast because ρ_r and ρ_l depend on contrast.

The binocular energy response also depends on stimulus disparity. Imagine that we introduce a small stimulus disparity d by shifting the spatial position of the sinusoidal grating in the right eye, in a direction perpendicular to the stimulus orientation. This will change the response phases of the right-eye monocular responses, R_0 and R_{90}, by an amount equal to $2\pi\omega d$, where $\omega = \sqrt{\omega_x^2 + \omega_y^2}$ is the stimulus spatial frequency. With this additional phase-offset, the the right-eye monocular responses become

$$\begin{aligned} R_0 &= \rho_r \sin(2\pi\omega_t t + \theta + 2\pi\omega d) \\ R_{90} &= \rho_r \cos(2\pi\omega_t t + \theta + 2\pi\omega d) \ . \end{aligned}$$

One can then show that the binocular energy neuron response becomes

$$E(d) = \rho_l^2 + \rho_r^2 + 2\rho_l\rho_r \cos(2\pi\omega d) \ . \tag{1.6}$$

To derive this equation from equation (1.4), use the trigonometry identity,

$$\cos(\alpha - \beta) = \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta) \ ,$$

with $\alpha = 2\pi\omega_t t + \theta + 2\pi\omega d$ and $\beta = 2\pi\omega_t t + \theta$.

The binocular energy response in equation (1.6) is a sum of three terms, namely, the two monocular energies, ρ_l^2 and ρ_r^2, and a term that is a cosinusoidal function of the interocular phase difference. The response therefore has a cosinusoidal dependence on binocular disparity. When $d = 0$, the cosine term is maximal, $\cos(0) = 1$, so the response is greatest when the disparity is zero. When $d = 1/(2\omega)$, the cosine term is minimized, $\cos(\pi) = -1$, so the response is smallest when the disparity is one-half the grating period. For disparities in between these two extremes, the binocular energy response varies as the cosine of the disparity times the grating frequency. We say that this model energy neuron has a preferred disparity of zero.

Equation (1.6) simplifies further when the drifting sinusoidal gratings presented to the two eyes have the same contrast:

$$E(d) \;=\; 2\,c^2 A^2\;[1 + \cos(2\pi\omega d)]. \tag{1.7}$$

From equations (1.6) and (1.7) one can see that the binocular energy response depends on the monocular energies and the binocular phase difference. The energy response does not depend directly on the individual monocular phases θ_l and θ_r. Also note that when the stimulus is turned off in one eye (e.g., if $c_l = 0$) then the binocular energy response reduces to the monocular energy in the other eye (e.g., ρ_r^2).

The cosinusoidal dependence on disparity is consistent with physiological data. Several studies have demonstrated that binocular simple and complex cell responses exhibit a sinusoidal dependence on stimulus disparity for drifting sinusoidal grating stimuli [Ohzawa and Freeman, 986a, Ohzawa and Freeman, 986b, Hammond, 1991, Wagner and Frost, 1994].

However, as discussed below in Section 1.5, the contrast dependence of real binocular neurons is not consistent with predictions of the binocular energy model. Therefore we will have to modify the binocular energy model. But first, we introduce position- and phase-shifts to construct binocular energy neurons tuned for non-zero disparities.

1.3.5 Position-Shift Model

A non-zero disparity preference can be introduced by shifting the receptive field position in one eye (Fig. 1.2B). As noted above in Section 1.3.1, the response phase of a monocular linear neuron depends on receptive field position. If θ is the response phase for a sine-grating stimulus with spatial frequency $\omega = \sqrt{\omega_x^2 + \omega_y^2}$, then after introducing a position shift (s) of the receptive field in the direction orthogonal to the stimulus orientation, the new response phase will be $\theta - 2\pi\omega s$. If the right-eye stimulus is also shifted by the disparity d, then the right-eye monocular responses become

$$\begin{aligned} R_0 &= \rho_r \sin(2\pi\omega_t t + \theta + 2\pi\omega(d - s)) \\ R_{90} &= \rho_r \cos(2\pi\omega_t t + \theta + 2\pi\omega(d - s))\;, \end{aligned}$$

where s is the position shift of the right-eye monocular weighting functions. Then, following the derivation above, the binocular energy response becomes

$$E(d;\, s) \;=\; \rho_l^2 \,+\, \rho_r^2 \,+\, 2\rho_l\rho_r \cos(2\pi\omega(d-s)) \;.$$

For equal contrast in both eyes, this simplifies to

$$E(d;\, s) \;=\; 2\,c^2 A^2 \;[1+\cos(2\pi\omega(d-s))] \;. \tag{1.8}$$

Peaks in the energy response occur whenever the cosine term is equal to one. This happens when the disparity satisfies $2\pi\omega(s-d) = 0$, i.e., when $s = d$. The cosine term is also one when the disparity is increased or decreased by multiples of the stimulus wavelength, that is, when

$$d \;=\; s + \frac{n}{\omega} \;, \tag{1.9}$$

for integer values of n. Because n can be any integer, response peaks occur periodically as a function of stimulus disparity, spaced by the stimulus wavelength.

Now consider what happens when you fix the stimulus disparity, but vary the stimulus spatial frequency. One of the energy response peaks ($n = 0$) always occurs when the disparity equals the position-shift $d = s$, independent of the frequency of the input (Fig. 1.3A). In fact, this is a key property of the position-shift model; the location of the primary response peak does not depend on stimulus spatial frequency [Wagner and Frost, 1993, Wagner and Frost, 1994]. The disparities of the other response peaks ($n \neq 0$), however, depend on stimulus frequency (see Fig. 1.3A).

1.3.6 Phase-Shift Model

Non-zero disparity preference can also be introduced by shifting the phase of the monocular subfields (Fig. 1.2C). Formally, let ψ denote a phase shift of the right-eye weighting functions. The right monocular responses are then

$$\begin{aligned} R_0 &= \rho_r \,\sin(\omega_t t + \theta - \psi + 2\pi\omega d) \\ R_{90} &= \rho_l \,\cos(\omega_t t + \theta - \psi + 2\pi\omega d) \;. \end{aligned}$$

Following the derivation used above, the binocular energy response becomes

$$E(d;\, \psi) \;=\; \rho_l^2 \,+\, \rho_r^2 \,+\, 2\rho_l\rho_r \cos(2\pi\omega d - \psi) \;.$$

For equal contrast in both eyes, this simplifies to

$$E(d;\, \psi) \;=\; 2\,c^2 A^2 \;[1+\cos(2\pi\omega d - \psi)] \;. \tag{1.10}$$

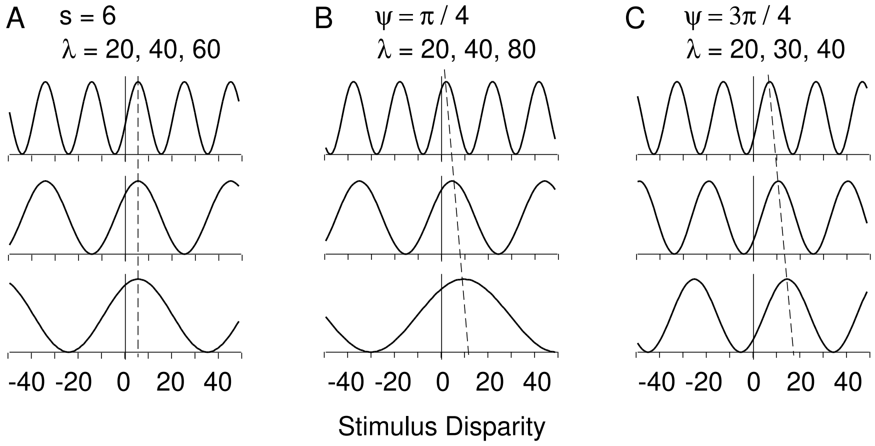

Fig. 1.3. Simulated disparity tuning curves of binocular energy neurons for drifting sine-grating stimuli. (A) A position-shift neuron with a preferred disparity of 6, and stimulus wavelengths spanning 1.5 octaves. The dashed line shows the alignment of primary response peaks. (B and C) Energy neurons with different phase-shifts, and different stimulus frequency ranges. A systematic shift in peak responses is evident for wide frequency ranges (in B) and for large phase-shifts (in C).

Peaks in the energy response now occur when $2\pi\omega d - \psi = n2\pi$, i.e., when

$$d \;=\; \frac{\psi}{2\pi\omega} + \frac{n}{\omega} \;. \tag{1.11}$$

As illustrated in Figure 1.3(B,C), this means that the neuron's "disparity preference" depends in a systematic way on the stimulus frequency. Even the primary response peak ($n = 0$) depends on the stimulus spatial frequency, since $d = \psi/(2\pi\omega)$ in this case. The range of possible "disparity preferences" depends on the range of spatial frequencies to which the neuron responds (i.e., the spatial frequency bandwidth), and on the neuron's interocular phase shift. It would be inaccurate to say that a phase-shifted binocular energy neuron has a unique preferred disparity.

Position-shifts and phase-shifts (s and ψ) have different effects on disparity tuning. One way to discriminate phase-shift neurons from position-shift neurons is to measure disparity tuning curves for sine-grating stimuli with different spatial frequencies. For a position-shift neuron, the primary response peak occurs at a single preferred disparity (i.e., the position shift) for all frequencies. For a phase-shift neuron, peaks in the tuning curves will occur at different disparities for different frequencies. Data of this sort have been obtained for the owl (Wagner & Frost, 1993, 1994), and are more consistent with the position-shift model.

1.3.7 Hybrid Model

As discussed below, the disparity selectivity of binocular neurons in cat and monkey primary visual cortices is probably due to a combination of position shifts *and* phase shifts. It is therefore natural to consider a hybrid model that incorporates both. Using the same analysis as above, with equal contrasts in the two eyes, one can show that the response of a hybrid binocular energy model, with a position shift s and a phase shift ψ, is given by

$$E(d; s, \psi) \; = \; 2\, c^2 A^2 \; [1 + \cos(2\pi\omega(d - s) - \psi)]. \tag{1.12}$$

The response, $E(d; s, \psi)$, now depends cosinusoidally on both the position shift s and the phase shift ψ.

1.4 Experimental Support for Position-Shifts and Phase-Shifts

The position-shift model involves binocular combinations of monocular receptive fields of similar shape at different retinal positions, while the phase-shift model combines monocular receptive fields with different shapes from the corresponding retinal locations. Only when both are tuned to a disparity of zero are they strictly equivalent. The next sections review neurophysiological evidence for position shifts and phase shifts. Then we propose an experiment for estimating both the position-shifts and phase-shifts of both simple and complex cells, based on their responses to drifting sinusoidal grating stimuli of different spatial frequencies and disparities.

1.4.1 Distribution of Preferred Disparities

One restriction on phase-shifted energy neurons stems from the fact that phase shifts are unique only between $-\pi$ and π. When combined with a restricted spatial frequency bandwidth, this means that for any one spatial frequency band, there is a limited range of disparities that one could hope to detect. The upper limits are reached as the phase-shift approaches $\pm\pi$ (i.e., half a wavelength) and the stimulus frequency approaches the lowest spatial frequencies to which the neurons are responsive. This limitation of the phase-shift model is particularly restrictive for neurons tuned to high spatial frequencies. Thus if a broad distribution of preferred disparities is found in a sample of neurons, relative to their preferred spatial frequencies, then one can infer that position shifts occur.

In attempting to measure the range of preferred disparities caution must be taken because eyes tend to drift and rotate under anaesthesia. To control for this, Hubel & Wiesel (1970) introduced the reference-cell method, in which a binocular cell is recorded for an extended period to find the disparity that elicits a maximal response. A second electrode is used to record from other neurons. By adjusting disparity settings to maintain the maximal response from the reference cell, one can track eye movements. Interestingly, it is not necessary to track eye drift in the owl, as their eye movements are negligible [Steinbach and Money, 1973].

In the cat, early reports gave a range of $\pm 3^\circ$ for the distribution of preferred disparities [Barlow et al., 1967]. Later studies using a reference-cell method found that the range of preferred disparities of tuned-excitatory cells in area 17 is less than 1° for eccentricities up to 8° [Ferster, 1981, LeVay and Voigt, 1988]. In the owl, the range of preferred disparities was found to be $\pm 2.5^\circ$ [Pettigrew, 1979]. In anesthetized monkeys, cells with preferred disparities up to $30'$ were documented in V2 [Hubel and Wiesel, 1970]. Studies on awake, behaving monkeys seldomly found preferred disparities greater than $12'$ (crossed or uncrossed) for eccentricities within 2 degrees of the fovea [Poggio and Fischer, 1977]. One would expect that cells in the parafoveal region might have larger preferred disparities, but we are aware of no quantitative data regarding this issue. In the monkey, near and far cells often respond maximally at the largest disparities that have been tested (up to 1°) [Poggio and Fischer, 1977]. Near and far cells of cats cover a range of at least $\pm 5^\circ$ of disparity [Ferster, 1981, LeVay and Voigt, 1988].

Unfortunately, spatial frequency tuning has usually not been measured along with disparity tuning. However, data from Ohzawa and Freeman (1986a,b) suggest that the range of preferred spatial frequencies in disparity-sensitive cells is similar to the overall range of preferred frequencies in cat area 17. Assuming the same in the monkey, with foveal simple and complex cells having preferred spatial frequencies between 1 and 10 cpd [DeValois et al., 982b], one can indirectly conclude that in monkeys, cats, and owls the preferred disparities cover a range that is larger than one period of the typical spatial frequency preference. This suggests that position-shifts occur, but it does not rule out the existence of additional phase-shifts.

1.4.2 Monocular Receptive-Field Shape

To determine whether there are phase-shifts, a more elaborate method is required. One method is to directly examine the shapes of the monocular receptive fields using white noise stimuli and reverse-correlation procedures. Ohzawa et al. (1991) and DeAngelis et al. (1991, 1995) applied this method to simple cells in cat area 17. They then fitted Gabor functions to the monocular receptive fields and used the phase of the fitted Gabor functions as a measure of receptive field shape. They found that the monocular receptive field shapes of binocular cells are often different. Moreover, the differences depend on orientation; cells tuned to horizontal orientations have similar receptive field shapes, while cells tuned to near vertical orientations exhibit a wide range of phase shifts (from 0° to 180°). While these data show that phase shifts exist, the existence of additional positional-shifts cannot be excluded because eye movements were not strictly controlled.

This reverse-correlation procedure works well for simple cells as their monocular responses depend strongly on the stimulus position within the receptive field. More sophisticated procedures, analyzing higher-order kernels of the white noise responses, would be needed to determine the monocular receptive field properties underlying disparity selectivity of complex cells.

1.4.3 Position- and Phase-Shifts of Simple Cells

Anzai et al. (1995) used white noise stimuli to estimate the best fitting Gabor functions to the monocular subfields of cat simple cells . From the phases of the Gabor functions they estimate receptive field phase shifts. From the locations of the centers of the fitted Gabor functions, with respect to a reference cell, they estimate the receptive field position shift.

Their results suggest that both position- and phase-shifts contribute to the disparity selectivity of binocular simple cells. Phase-shifts appear more significant at lower spatial frequencies, while position shifts contribute more to the disparity selectivity in cells tuned to higher spatial frequencies.

1.4.4 Proposed Experiment for Estimating Position- and Phase-Shifts of Simple and Complex Cells

The method used by Anzai *et al.* (1995) applies only to simple cells. In addition, their analysis depends on the validity of the Gabor function receptive field model. Here, we propose an experiment for estimating the phase-shift and the position-shift of all binocular (both simple and complex) neurons. Our method relies only on the cosinusoidal nature of the disparity tuning curves (see Figure 1.4).

The procedure is as follows: *1)* Measure disparity tuning curves with drifting sinusoidal gratings of various spatial frequencies ω_j. As expressed in equation (1.12) the response of a binocular energy neuron depends cosinusoidally on stimulus disparity, where the frequency of the cosinusoid is ω_j. *2)* Fit a cosinusoidal function with frequency ω_j to each tuning curve, from which a phase, denoted by $\Phi(\omega_j)$, is obtained (see Figure 1.4A). *3)* Plot these

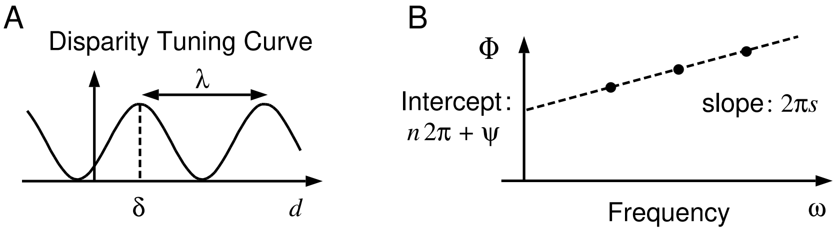

Fig. 1.4. (A) Cosinusoidal disparity tuning curve of a hybrid energy neuron. With the peak closest to the origin (zero disparity) at δ and an input spatial frequency of $\omega = 1/\lambda$, the phase of the cosinusoidal tuning curve is $\Phi = 2\pi\delta\omega$. (B) This phase Φ is a linear function of stimulus spatial frequency, the slope and intercept of which give the position-shift and the phase-shift of the monocular receptive fields.

fitted phase values $\Phi(\omega_j)$ as a function of spatial frequency. From equation (1.12), these fitted phase values should depend linearly on stimulus spatial frequency, $\Phi(\omega_j) = 2\pi\omega_j s + \psi$, as depicted in Figure 1.4(B). *4)* The slope of the line gives the receptive field position shift s, and the y-intercept of the line gives the receptive field phase-shift ψ. In particular, the slope equals $2\pi s$ and the y-intercept equals $\psi + n2\pi$. Because ψ is unique only within $-\pi$ and π, one can find the phase-shift from the intercept by adding whatever multiple of 2π is required to bring the result into the range $(-\pi, \pi]$.

Similar methods have been used to explore the encoding of interaural time differences in the auditory system of cats (Yin & Kuwada, 1983). In the visual system this method can be used to measure position- and phase-shifts in simple and complex cells, without requiring that the monocular receptive fields shapes be accurately localized or described. It is necessary however to stabilize the eyes for the duration of the matrix of spatial frequency and disparity conditions. If stabilization can not be guaranteed, then one could define disparity with respect to a reference cell.

1.5 Response Normalization

Many aspects of simple and complex cell responses are consistent with the linear and energy models. However, the linear/energy model falls short of a complete explanation of cell responses in primary visual cortex. One major fault with the model is the fact that cell responses saturate (level off) at high contrasts. A second fault with the linear model is revealed by testing superposition. A typical simple cell responds vigorously to its preferred orientation but not at all to the perpendicular orientation. According to the linear model, the response to the superimposed pair of stimuli (preferred plus perpendicular) should equal the response to the preferred stimulus alone. In fact, the response to the superimposed pair is about half that predicted (e.g., [Bonds, 1989]), a phenomenon known as cross-orientation suppression.

To explain response saturation, cross-orientation suppression, and other violations of the linear/energy models, we and others have recently proposed a new model of V1 cell responses called the *normalization* model [Robson et al., 1991, Albrecht and Geisler, 1991, Heeger, 1991, Heeger, 992a, Heeger, 1993, Carandini and Heeger, 1994, Fleet et al., 1995]. The normalization model is based on an underlying linear stage. The linear stage is followed by a normalization stage, where each neuron's linear response to the stimulus is divided by a quantity proportional to the pooled activity of a large number of other neurons. Thus, the activity of a large pool of neurons partially suppresses the response of each individual neuron. Normalization is a nonlinear operation: one input (a neuron's underlying linear response) is divided by another input (the pooled activity of a large number of neurons). The effect of this divisive suppression is that the response of each neuron is normalized (rescaled) with respect to stimulus contrast.

The normalization model explains a large body of otherwise unexplained physiological phenomena [Heeger, 992a]. According to the model, a cell's selectivity is attributed to summation (the linear stage) and its nonlinear behavior is attributed to division (the normalization stage). The model explains response saturation because the divisive suppression increases with stimulus contrast. The model explains cross-orientation suppression because a given cell is suppressed by many other cells including those with perpendicular orientation tunings.

Response normalization and gain control appear to be significant in binocular neurons as well. With respect to the site of the normalization, there is evidence for response normalization both before and after the combination of signals from the two eyes.

In support of binocular normalization, Anzai et al. (1995) found that binocular neurons exhibit sub-linear binocular interactions. Binocular contrast response saturates at a higher firing rate than the monocular curves, but only by a factor of about $\sqrt{2}$. In other words, the response to a binocular stimulus is less than the sum of the two responses to the component monocular stimuli. In addition, Sclar et al. (1985) reported that there is a significant interocular transfer of contrast adaptation.

In support of monocular normalization, the preponderance of evidence implies that cross-orientation suppression is monocular. There is little or no suppression when the preferred and perpendicular gratings are presented dichoptically (one in one eye and one in the other) [DeAngelis et al., 1992, Ohzawa and Freeman, 1994]. However, Bonds (1989) found no evidence for cross-orientation suppression in the LGN, suggesting that it occurs in the cortex.

A second source of evidence for monocular normalization stems from a surprising result reported by Freeman and Ohzawa (1990) and by Ohzawa and Freeman (1993). As mentioned above, simple and complex cells respond with periodic disparity-tuning curves when stimulated with drifting sinusoidal gratings. Freeman and Ohzawa measured the depth of modulation of the periodic disparity-tuning curves while varying grating contrast in one eye. As shown in Figure 1.5 they found that the depth of modulation depends surprisingly little on interocular contrast differences.

The energy model predicts a strong dependence on interocular contrast differences. As shown in Figure 1.6, when the left and right contrasts are equal, then the depth of modulation is expected to be one. But when there is a 10-fold difference in the left and right contrasts, the depth of modulation predicted by the un-normalized energy model is below 0.2. This does not agree with the data shown in Figure 1.5. The predictions in Figure 1.6 were computed from equation (1.6); a cosinusoid modulating about a mean value given by the sum of the monocular energies, and with an amplitude given by twice the product of the monocular energies.

1.5.1 Binocular Normalization Model

There are two simple ways that the depth of modulation can be kept nearly constant as the contrast of one eye's input is changed. Both involve monocular normalization preceding the binocular summation in the energy model. The first way, which we call the *monocular gain model*, involves independent monocular normalization of the left- and right-eye responses. This is shown schematically in Figure 1.7. The second way is an *interocular model*, where the amplitude of the right eye's response (plus an additive constant) acts as a multiplicative gain for the left signal, and vice versa. In what follows we will concentrate on the monocular gain model as this is more consistent with the lack of interocular cross-orientation suppression.

After binocular summation of the (normalized) monocular signals, we propose that there is a second stage of binocular normalization, as depicted in Figure 1.8. The binocular normalization stage is a straightforward extension of the monocular normalization model proposed by Heeger (1992b).

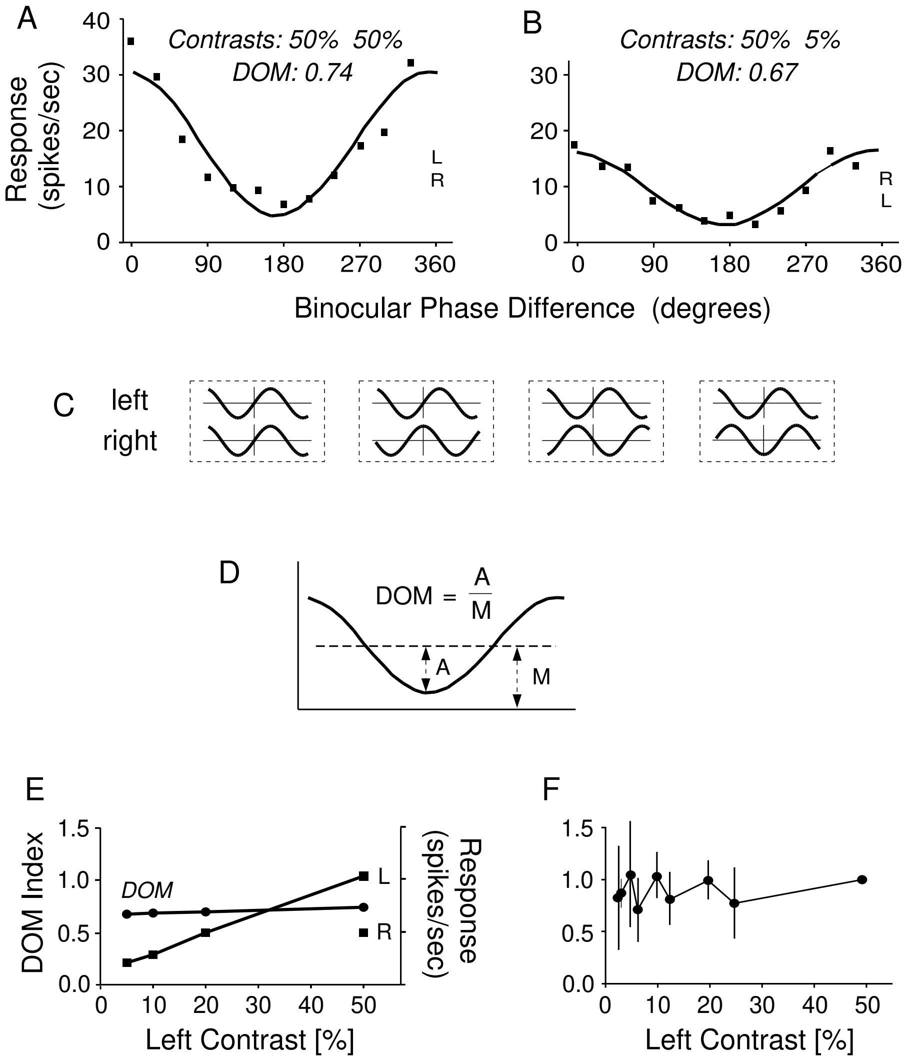

Fig. 1.5. (A) Response as a function of disparity for drifting grating stimuli with equal (50%) contrasts in both eyes. The disparity tuning curve is nearly sinusoidal. (B) Response as a function of disparity when the left eye contrast is reduced to 5%. The mean response drops somewhat but the depth of modulation is virtually unaffected. (C) Left and right stimulus intensity profiles for binocular phase differences of 0, 90, 180, and 270 degrees. (D) Depth of modulation is defined as the amplitude of the first harmonic divided by the mean response. (E) Depth of modulation of the binocular response as a function of left eye contrast (with right eye contrast fixed at 50%), and response as a function of contrast for monocular (left eye) stimulation. The monocular (left eye) contrast-response curve increases monotonically with contrast, but the depth of modulation of the binocular response is virtually unaffected by left eye contrast. (F) Mean and standard deviation of depth of modulation for 21 cells as left eye contrast was varied. (Redrawn from Freeman and Ohzawa, 1990 and Ohzawa and Freeman, 1993).

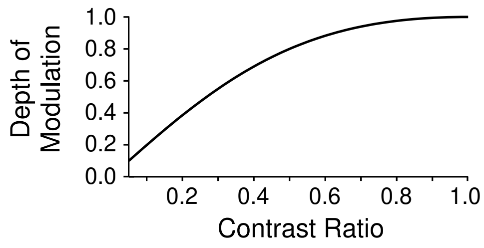

Fig. 1.6. Depth of modulation prediction of unnormalized energy model, as a function of the ratio of contrasts in the right and left stimuli.

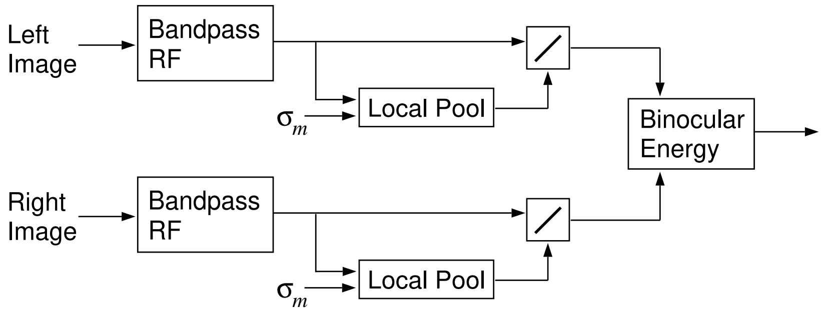

Fig. 1.7. Monocular component of response normalization. The left and right monocular neurons are normalized separately. The normalization pool includes a range of spatial positions, a range of spatial frequency preferences, and all orientation preferences.

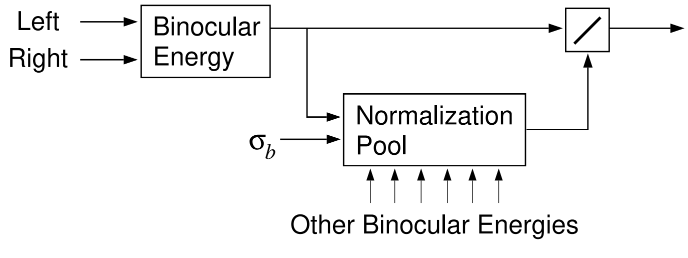

Fig. 1.8. Binocular component of response normalization. The binocular normalization pool includes a range of spatial positions, and all disparity preferences.

Monocular Normalization

The monocular normalization signal is the sum of a pool of rectified linear responses plus a small constant σ_m, as shown in Figure 1.7. Since we use half-squaring for the rectifier, the monocular gain can be computed either as an average of the monocular half-squared (model simple cell) responses or as an average of the monocular energy (complex cell) responses [Heeger, 992b]. In either case the normalizing signal is proportional to the local Fourier energy of the stimulus plus a small constant.

Formally, let $\lfloor L_{90}(t) \rfloor^2$ denote the half-squared response of a monocular, left-eye, linear neuron, where $\lfloor L_{90} \rfloor^2 = L_{90} \, |L_{90}|$. Let $E_l(t)$ be the local monocular energy, i.e., a local average of half-squared neurons with different orientation and spatial frequency preferences, and different receptive field phases. Then, the normalized, monocular, left-eye, response is:

$$\bar{L}_{90}(t) \;=\; \frac{\lfloor L_{90}(t) \rfloor^2}{E_l(t) + \sigma_m} \,. \tag{1.13}$$

Since the energy in the denominator includes the half-squared response in the numerator, the normalized response will saturate at high contrasts. The σ_m parameter is called the monocular semi-saturation constant because it determines the contrast that evokes half the maximum attainable response.

The right-eye responses are also normalized:

$$\bar{R}_{90}(t) \;=\; \frac{\lfloor R_{90}(t) \rfloor^2}{E_r(t) + \sigma_m} \,, \tag{1.14}$$

The inputs $L_{90}(t)$ and $R_{90}(t)$ to the energy model described by equation (1.4) are then replaced by $\bar{L}_{90}(t)$ and $\bar{R}_{90}(t)$, as depicted in Figure 1.7.

Note that there are now two stages of squaring in this model; half-squaring before the monocular normalization, and then full-squaring in the binocular energy computation. Even with this "extra" squaring step, the disparity tuning of the binocular energy responses are still (very nearly) sinusoidal.

For the simulations results reported here, we used even and odd-symmetric Gabor functions with a bandwidth of 1.5 octaves for the underlying linear weighting functions. The monocular normalization signal was pooled over a small, Gaussian weighted, spatial neighbourhood; the spatial extent of the Gaussian window was equal to that of the Gabor weighting functions.

Binocular Normalization

The binocular normalization signal is the sum of a pool of binocular energy responses. The binocular pool includes the complete range of preferred binocular disparities (arising from position- and/or phase-shifts), from within a small local spatial neighbourhood. But unlike the monocular normalization, only neurons with the same spatial frequency and orientation preference are included. Formally, we write the normalized binocular energy as follows

$$\bar{E}(t) \;=\; \frac{E(t;\, s,\, \psi)}{S(t) + \sigma_b} \;, \tag{1.15}$$

where

$$S(t) \;=\; \sum_{x,\, s,\, \psi} E(t;\, s,\, \psi) \;, \tag{1.16}$$

where $E(t;\, s,\, \psi)$ is the binocular energy response computed from the position- and phase-shifted, monocularly normalized responses.

For the simulations reported here, we used position-shift energy neurons with preferred disparities that span three wavelengths. The binocular normalization signal was pooled over a Gaussian weighted spatial neighbourhood with standard deviation equal to that of the monocular spatial pooling.

1.5.2 Contrast Response of Normalized Energy Neurons

The full model includes some monocular neurons and some binocular neurons. The monocular neurons in the model are identical to those in Heeger's (1992a) normalization model. The contrast response function of these monocular model neurons is given by

$$R(c) = K\,\frac{c^2}{c^2 + \sigma_m^2} \;,$$

that is, a hyperbolic ratio function with an exponent of 2.

The contrast response curves of the model's binocular neurons are similar: Figure 1.9(A) shows monocular contrast response curves from a model binocular energy neuron. The curves look like hyperbolic ratio functions with an exponent slightly greater than 2. Although the binocular energy neurons have gone through two squaring nonlinearities, the cumulative effect is an exponent between 2 and 3 (see *Appendix*).

The contrast-response curves also depend on the semi-saturation constants, σ_m and σ_b. In particular, changing σ_m shifts the contrast response curve laterally (Figure 1.9A). As σ_b is increased, its effect eventually diminishes and the contrast response curve depends primarily on σ_m. If σ_b is very small, however, then the contrast response curve shifts to the left and becomes somewhat steeper (Figure 1.9B). The slope on the log-contrast

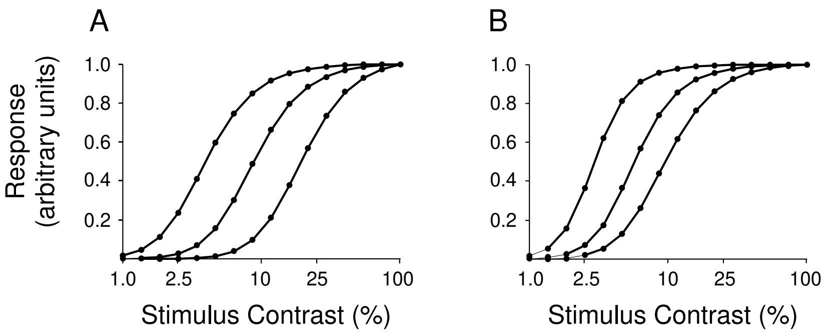

Fig. 1.9. Monocular contrast response curves for a model binocular energy neuron. (A) The three curves correspond to three different values of σ_m, with σ_b held constant. (B) The three curves correspond to three different values of σ_b, with σ_m held constant.

axis corresponds to roughly to the exponent of the hyperbolic ratio function. Therefore, Figure 1.9(B) shows that the effective exponent depends implicitly on σ_b (see *Appendix*).

Figure 1.10(A) shows the monocular and the binocular contrast response curves recorded extracellularly from a simple cell and a complex cell in the cat A17 [Anzai et al., 1995a]. The binocular contrast response curves have a lower threshold and rise somewhat more steeply than the monocular curves. Moreover they saturate at a high firing rate. Figure 1.10(B) shows that a normalized, energy neuron behaves similarly.

1.5.3 Stability of Depth of Modulation and Mean

As discussed above, a remarkable property of binocular neurons in A17 of the cat is that the depth of modulation in their sinusoidal disparity tuning curves is invariant to interocular contrast differences. In many cells, the depth of modulation was largely unaffected even with 10-fold differences in contrast between the right-eye and left-eye (Figure 1.5). The unnormalized energy model predicts a 5-fold decrease in the depth of modulation for a 10-fold contrast difference (Figure 1.6). When normalization is included in the model, however, the depth of modulation is significantly less sensitive to interocular contrast differences, as shown in Figure 1.11. The simulations of the binocular energy neuron in Figures 1.11(A,B) can be compared with Freeman and Ohzawa's data which is replotted here in Figures 1.5(A,B). Finally, Figure 1.12 shows both the depth of modulation as a function of left eye contrast, and the contrast-response curve for monocular (left eye)

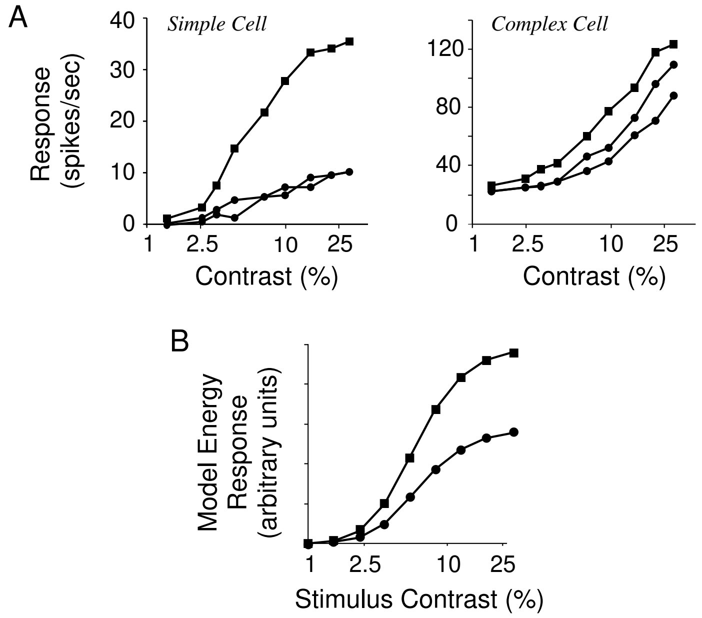

Fig. 1.10. (A) Monocular (left and right, circles) and binocular (squares) contrast response curves for a simple cell and a complex cell in A17 of the cat (redrawn from Anzai et al., 1995). (B) Monocular and binocular contrast responses curves for a normalized, binocular, energy neuron.

stimulation. These simulation results can be compared with Freeman and Ohzawa's data, replotted here in Figure 1.5(E).

As explained in the appendix, the depth of modulation is controlled mainly by the monocular semi-saturation constant σ_m. In these simulations we have therefore set σ_m in order to keep the depth of modulation above 0.95.

1.6 Discussion

To understand the neural basis for stereoscopic vision one must address several issues including the form of binocular interaction in simple and complex cells, the basis for their disparity selectivity, and the way in which they encode disparity. This article examines an energy model of binocular interaction with monocular and binocular response normalization. Disparity selectivity of the model neurons arises from a combination of position-shifts and phase-shifts between the monocular subfields of binocular receptive fields. Position- and phase-shifts have different quantitative properties, and it is argued that both likely contribute to the disparity selectivity of cells in V1.

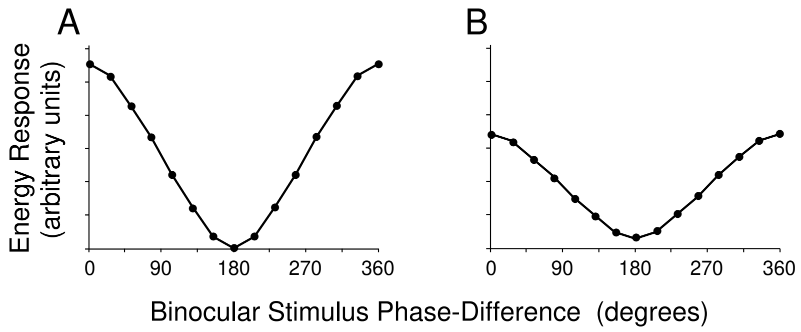

Fig. 1.11. Disparity tuning curves, using drifting sinusoidal grating stimuli, of a normalized, energy neuron. The contrast of the right-eye stimulus was fixed at 50%. (A) When left-eye contrast is 50% the depth of modulation is 1. (B) When the left-eye contrast is 5% the depth of modulation decreases only slight to 0.95. But the mean response drops by a factor of 0.75.

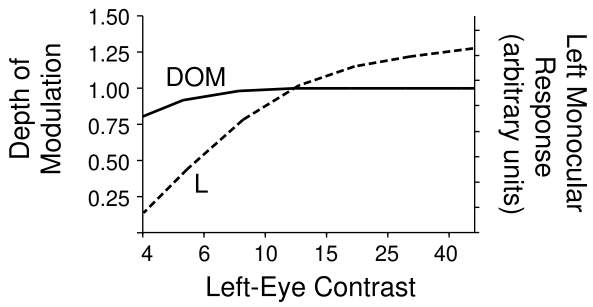

Fig. 1.12. Depth of modulation (solid curve) as a function of left-eye contrast (with right-eye contrast fixed at 50%), and monocular (left-eye) contrast-response curve (dashed curve), for a normalized, binocular energy neuron. The monocular (left eye) contrast-response curve increases monotonically with contrast, but the depth of modulation of the binocular response is largely unaffected by left-eye contrast.

The relative contribution of position- and phase-shift can be inferred by measuring disparity-tuning curves using drifting sinusoidal grating stimuli with several different spatial frequencies.

The model also involves two stages of response normalization. There is a monocular form of normalization that occurs before binocular interaction, followed by normalization of binocular responses. These stages of normalization help to account for observed properties of the monocular and binocular contrast response curves. Normalization also accounts for the invariance of disparity tuning with response to interocular contrast differences.

Acknowledgements: This work was supported by grants from NSERC Canada and Queen's University to DJF, by a DFG grant to HW and DJF, and by an NIMH grant (MH50228) and an Alfred P. Sloan Research Fellowship to DJH.

Appendix: Mathematical Notes on Binocular Normalization

According to equation 1.6 in Section 1.3.4, the response of an unnormalized binocular energy neuron has the form

$$E(d) \; = \; \rho_l^2 \, + \, \rho_r^2 \, + \, 2\rho_l\rho_r \cos(2\pi\omega d) \; .$$

The response mean is $\rho_l^2 + \rho_r^2$, and the amplitude of modulation is $2\rho_l\rho_r$. The depth of modulation, shown in Figure 1.6, is equal to the amplitude divided by the mean. In the normalized model, ρ_l^2 and ρ_r^2 are replaced by the normalized monocular energies

$$\begin{aligned} \bar{\rho}_l^2 &= \bar{L}_0^2 \, + \, \bar{L}_{90}^2 \\ \bar{\rho}_r^2 &= \bar{R}_0^2 \, + \, \bar{R}_{90}^2 \end{aligned}$$

where $\bar{L}_0$ and the other normalized responses are defined by equation 1.13.

The normalized binocular response is then approximately equal to:

$$\bar{E}(d) \; \approx \; \frac{\bar{\rho}_l^2 \, + \, \bar{\rho}_r^2 \, + \, 2\bar{\rho}_l\bar{\rho}_r \cos(2\pi\omega d)}{\bar{\rho}_l^2 \, + \, \bar{\rho}_r^2 \, + \, \sigma_b}$$

Both the mean response and the modulation amplitude of the response have the same denominator. Thus the ratio of the modulation amplitude and the mean response is given by

$$\frac{2\bar{\rho}_l\bar{\rho}_r}{\bar{\rho}_l^2 \, + \, \bar{\rho}_r^2} \; .$$

Note that the depth of modulation does not depend on the binocular gain parameter σ_b. Rather, it depends only on the monocular energies, which depend on the input contrasts and the monocular gain parameter σ_m. Given the contrasts of the left and right stimuli, and a desired depth of modulation, one can solve for σ_m in closed form. In the simulations reported here, unless otherwise stated, we set σ_m to 0.0005.

The normalized binocular energy model has two squaring nonlinearities.

The first, in equation (1.13), is monocular. The second occurs after binocular interaction, in equation (1.4). The monocular contrast response curves shown in Figure 1.9 are well approximated by equation (1.5.2) with an exponent between 2 and 3, rather than an exponent of 4 as one might expect.

To explain this, consider the binocular energy response when the right stimulus contrast is zero, that is

$$\bar{E} \;=\; \frac{\bar{L}_0^2 \;+\; \bar{L}_{90}^2}{\bar{L}_0^2 \;+\; \bar{L}_{90}^2 \;+\; \sigma_b} \tag{1.17}$$

One can derive this from equation 1.15 by setting $\bar{R}_0 \;=\; \bar{R}_{90} \;=\; 0$. If we substitute the normalized monocular responses defined by equation (1.13) into equation (1.17), then we obtain

$$\bar{E} \;\approx\; \frac{L_0^4 \;+\; L_{90}^4}{L_0^4 \;+\; L_{90}^4 \;+\; \sigma_b(E_l + \sigma_m)^2}$$

Here, E_l is the average broad-band monocular energy in the left stimulus. For sine-grating stimuli this is constant and equal to c_l^2.

To simplify this expression further one can approximate $L_0^4 + L_{90}^4$ by $\rho_l^4 = c_l^4\, A^4$. If we ignore the attenuation of the response due to the sensitivity of the neuron's transfer function to the stimulus frequency (which introduces a multiplicative constant), then one can approximate ${L_0}^4 \;+\; {L_{90}}^4$ by c_l^4. This allows us to write the energy response as

$$\begin{aligned} \bar{E} \;&=\; \frac{c_l^4}{c_l^4 \;+\; \sigma_b(c_l^2 + \sigma_m)^2} \\ &=\; \frac{c_l^4}{c_l^4 \;+\; \sigma_b(c_l^4 + \sigma_m^2 + 2c_l^2\sigma_m)} \\ &=\; \frac{c_l^4}{c_l^4\,(1+\sigma_b) + 2c_l^2\sigma_b\sigma_m + \sigma_b\sigma_m^2} \end{aligned} \tag{1.18}$$

The third term in the denominator will usually be extremely small, because σ_m is small in order to keep the depth of modulation stable. Thus, when c_l is relatively large, this third term can be ignored and the monocular contrast response then simplifies to

$$\bar{E} \;=\; \frac{c_l^2}{c_l^2\,(1+\sigma_b) + 2\sigma_b\sigma_m}\;. \tag{1.19}$$

Thus for larger contrasts, the contrast response resembles a hyperbolic ratio function with an exponent of about 2. For smaller values of c_l the third term in the denominator of equation (1.18) remains significant and the exponent of the effective hyperbolic ratio function is between 2 and 4. When σ_b is small this effect is more evident.

Bibliography

[Adelson and Bergen, 1985] Adelson, E. H. and Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, 2:284–299.

[Albrecht and Geisler, 1991] Albrecht, D. G. and Geisler, W. S. (1991). Motion sensitivity and the contrast-response funtion of simple cells in the visual cortex. *Visual Neuroscience*, 7:531–546.

[Anzai et al., 1995a] Anzai, A., Bearse, M. A., Freeman, R. D., and Cai, D. (1995a). Contrast coding in cells in the cat's striate cortex: Monocular vs binocular detection. *Visual Neuroscience*, page (in press).

[Anzai et al., 1995b] Anzai, A., Ohzawa, I., Freeman, R. D., and Cohn, T. E. (1995b). Do simple cells in the cat's striate cortex encode binocular disparity through position and phase incongruities? *Society for Neuroscience Abstracts*, 21:1648.

[Barlow et al., 1967] Barlow, H. B., Blakemore, C., and Pettigrew, J. D. (1967). The neural mechanism of binocular depth discrimination. *Journal of Physiology (London)*, 193:327–342.

[Bonds, 1989] Bonds, A. B. (1989). Role of inhibition in the specification of orientation selectivity of cells in the cat striate cortex. *Visual Neuroscience*, 2:41–55.

[Campbell et al., 1968] Campbell, F. W., Cleland, B. G., Cooper, G., and Enroth-Cugell, C. (1968). The angular selectivity of visual cortical cells to moving gratings. *Journal of Physiology (London)*, 198:237–250.

[Campbell et al., 1969] Campbell, F. W., Cooper, G., and Enroth-Cugell, C. (1969). The spatial selectivity of visual cells of the cat. *Journal of Physiology (London)*, 203:223–235.

[Carandini and Heeger, 1994] Carandini, M. and Heeger, D. J. (1994). Summation and division by neurons in primate visual cortex. *Science*, 264:1333–1336.

[Clarke et al., 1976] Clarke, P. G. H., Donaldson, I. M. L., and Witteridge, D. (1976). Binocular visual mechanisms in cortical areas i and ii of the sheep. *Journal of Physiology (London)*, 256:509–526.

[DeAngelis et al., 1991] DeAngelis, G. C., Ohzawa, I., and Freeman, R. D. (1991). Depth is encoded in the visual system by a specialized receptive field structure. *Nature*, 352:156–159.

[DeAngelis et al., 1995] DeAngelis, G. C., Ohzawa, I., and Freeman, R. D. (1995). Neuronal mechanisms underlying stereopsis: how do simple cells in the visual cortex encode binocular disparity? *Perception*, 24:3–32.

[DeAngelis et al., 1992] DeAngelis, G. C., Robson, J. G., Ohzawa, I., and Freeman, R. D. (1992). The organization of suppression in receptive fields of neurons in the cat's visual cortex. *Journal of Neurophysiology*, 68:144–163.

[DeValois et al., 982b] DeValois, R. L., Albrecht, D. G., and Thorell, L. G. (1982b). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, 22:545–559.

[Emerson et al., 1992] Emerson, R. C., Bergen, J. R., and Adelson, E. H. (1992). Directionally selective complex cells and the computation of motion energy in cat visual cortex. *Vision Research*, 32:203–218.

[Ferster, 1981] Ferster, D. (1981). A comparison of binocular depth mechanisms in areas 17 and 18 of the cat visual cortex. *Journal of Physiology (London)*, 311:623–655.

[Field and Tolhurst, 1986] Field, D. J. and Tolhurst, D. J. (1986). The structure and symmetry of simple-cell receptive field profiles in the cat's visual cortex. *Proceedings of the Royal Society of London, B*, 228:379–400.

[Fischer and Kruger, 1979] Fischer, B. and Kruger, J. (1979). Disparity tuning and binocularity of single neurons in cat visual cortex. *Experimental Brain Research*, 35:1–8.

[Fleet, 1994] Fleet, D. J. (1994). Disparity from local weighted phase-correlation. In *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, pages 48–56, San Antonio.

[Fleet et al., 1995] Fleet, D. J., Heeger, D. J., and Wagner, H. (1995). Computational model of binocular disparity. *Investigative Opthalmology and Visual Science Supplement*, 36:365.

[Fleet et al., 1996] Fleet, D. J., Wagner, H., and Heeger, D. J. (1996). Neural encoding of binocular disparity: Energy models, position shifts and phase shifts. *Vision Research*, (in press).

[Foster et al., 1983] Foster, K. H., Gaska, J. P., Marcelja, S., and Pollen, D. A. (1983). Phase relationships between adjacent simple cells in the feline visual cortex. *Journal of Physiology (London)*, 345:22P.

[Freeman and Ohzawa, 1990] Freeman, R. D. and Ohzawa, I. (1990). On the neurophysiological organization of binocular vision. *Vision Research*, 30:1661–1676.

[Hamilton et al., 1989] Hamilton, D. B., Albrecht, D. G., and Geisler, W. S. (1989). Visual cortical receptive fields in monkey and cat: spatial and temporal phase transfer function. *Vision Research*, 29:1285–1308.

[Hammond, 1991] Hammond, P. (1991). Binocular phase specificity of striate cortical neurons. *Experimental Brain Research*, 87:615–623.

[Heeger, 1991] Heeger, D. J. (1991). Nonlinear model of neural responses in cat visual cortex. In Landy, M. and Movshon, J. A., editors, *Computational Models of Visual Processing*, pages 119–133. MIT Press, Cambridge, MA.

[Heeger, 992a] Heeger, D. J. (1992a). Normalization of cell responses in cat striate cortex. *Visual Neuroscience*, 9:181–198.

[Heeger, 992b] Heeger, D. J. (1992b). Half-squaring in responses of cat simple cells. *Visual Neuroscience*, 9:427–443.

[Heeger, 1993] Heeger, D. J. (1993). Modeling simple cell direction selectivity with normalized, half-squared, linear operators. *Journal of Neurophysiology*, 70:1885–1898.

[Heggelund, 1986] Heggelund, P. (1986). Quantative studies of the discharge fields of single cells in cat striate cortex. *Journal of Physiology (London)*, 373:277–292.

[Hubel and Wiesel, 1970] Hubel, D. and Wiesel, T. (1970). Stereoscopic vision in macaque monkey. *Nature*, 225:41–42.

[Hubel and Wiesel, 1962] Hubel, D. H. and Wiesel, T. N. (1962). Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. *Journal of Physiology (London)*, 160:106–154.

[Jones and Palmer, 987a] Jones, J. P. and Palmer, L. A. (1987a). The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, 58:1187–1211.

[LeVay and Voigt, 1988] LeVay, S. and Voigt, T. (1988). Ocular dominance and disparity coding in cat visual cortex. *Visual Neuroscience*, 1:395–413.

[Liu et al., 1992] Liu, A., Gaska, J. P., Jacobson, L. D., and Pollen, D. A. (1992). Interneuronal interaction between members of quadrature phase and anti-phase pairs in the cat's visual cortex. *Vision Research*, 32:1193–1198.

[Maske et al., 1984] Maske, R., Yamane, S., and Bishop, P. O. (1984). Binocular simple cells for local stereopsis: comparison of receptive field organizations for the two eyes. *Vision Research*, 24:1921–1929.

[Movshon et al., 978a] Movshon, J. A., Thompson, I. D., and Tolhurst, D. J. (1978a). Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *Journal of Physiology (London)*, 283:53–77.

[Nikara et al., 1968] Nikara, T., Bishop, P. O., and Pettigrew, J. D. (1968). Analysis of retinal correspondence by studying receptive fields of binocular single units in cat striate cortex. *Experimental Brain Research*, 6:353–372.

[Nomura et al., 1990] Nomura, M., Matsumoto, G., and Fujiwara, S. (1990). A binocular model for the simple cell. *Biological Cybernetics*, 63:237–242.

[Ohzawa et al., 1990] Ohzawa, I., DeAngelis, G. C., and Freeman, R. D. (1990). Stereoscopic depth discrimination in the visual cortex: neurons ideally suited as disparity detectors. *Science*, 249:1037–1041.

[Ohzawa and Freeman, 986a] Ohzawa, I. and Freeman, R. D. (1986a). The binocular organization of simple cells in the cat's visual cortex. *Journal of Neurophysiology*, 56:221–242.

[Ohzawa and Freeman, 986b] Ohzawa, I. and Freeman, R. D. (1986b). The binocular organization of complex cells in the cat's visual cortex. *Journal of Neurophysiology*, 56:243–259.

[Ohzawa and Freeman, 1994] Ohzawa, I. and Freeman, R. D. (1994). Monocular and binocular mechanisms of contrast gain control. In Lawton, T. A., editor, *Computational Vision Based on Neurobiology, SPIE Proceedings, V. 2054*.

[Palmer and Davis, 1981] Palmer, L. A. and Davis, T. L. (1981). Receptive-field structure in cat striate cortex. *Journal of Neurophysiology*, 46:260–276.

[Pettigrew, 1972] Pettigrew, J. D. (1972). The neurophysiology of binocular vision. *Scientific American*, August:84–95.

[Pettigrew, 1979] Pettigrew, J. D. (1979). Binocular visual processing in the owl's telencephalon. *Proceedings of the Royal Society of London B*, 204:435–454.

[Pettigrew and Konishi, 1976] Pettigrew, J. D. and Konishi, M. (1976). Neurons selective for orientation and binocular disparity in the visual wulst of the barn owl (tyto alba). *Science*, 193:675–678.

[Pettigrew et al., 1968] Pettigrew, J. D., Nikara, T., and Bishop, P. O. (1968). Responses to moving slits by single units in cat striate cortex. *Experimental Brain Research*, 6:373–390.

[Poggio and Fischer, 1977] Poggio, G. F. and Fischer, B. (1977). Binocular interaction and depth sensitivity in the striate and prestriate cortex of the behaving monkey. *Journal of Neurophysiology*, 40:1392–1405.

[Poggio et al., 1985] Poggio, G. F., Motter, B. C., Squatrito, S., and Trotter, Y. (1985). Responses of neurons in visual cortex (v1 and v2) of the alert macaque to dynamic random-dot stereograms. *Vision Research*, 25:397–406.

[Poggio and Talbot, 1981] Poggio, G. F. and Talbot, W. H. (1981). Mechanisms of static and cynamic stereopsis in foveal cortex of the rhesus monkey. *Journal of Physiology (London)*, 315:469–492.

[Pollen and Ronner, 1981] Pollen, D. A. and Ronner, S. (1981). Phase relationships between adjacent simple cells in the visual cortex. *Science*, 212:1409–1411.

[Pollen and Ronner, 1983] Pollen, D. A. and Ronner, S. (1983). Visual cortical neurons as localized spatial frequency filters. *IEEE Transactions on Systems, Man, and Cybernetics*, 13:907–916.

[Robson et al., 1991] Robson, J. G., DeAngelis, G. C., Ohzawa, I., and Freeman, R. D. (1991). Cross-orientation inhibition in cat cortical cells originates from within the receptive field. *Investigative Opthalmology and Visual Science Supplement*, 32:429.

[Sclar et al., 1985] Sclar, G., Ohzawa, I., and Freeman, R. D. (1985). Contrast gain control in the kitten's visual system. *Journal of Neurophysiology*, 54:666–673.

[Steinbach and Money, 1973] Steinbach, M. J. and Money, K. E. (1973). Eye movements of the owl. *Vision Research*, 13:889–891.

[Wagner and Frost, 1993] Wagner, H. and Frost, B. (1993). Disparity-sensitive cells in the owl have a characteristic disparity. *Nature*, 364:796–798.

[Wagner and Frost, 1994] Wagner, H. and Frost, B. (1994). Binocular responses of neurons in the barn owl's visual wulst. *Journal of Comparative Physiology A*, 174:661–670.