9

Nonlinear Model of Neural Responses in Cat Visual Cortex

David J. Heeger

The first description of the physiological responses of neurons in the cat's primary (striate) visual cortex was provided by Hubel and Wiesel in 1962. They were able to divide striate cells into two groups, simple and complex. Simple cells have receptive fields that are, like those of the retina and lateral geniculate nucleus (LGN), divided into distinct excitatory and inhibitory regions. Hubel and Wiesel found that "summation occurred within either type of region," and that "when the two opposite regions were illuminated together their effects tended to cancel." This suggests that simple cells sum linearly over their receptive fields. Complex cells, on the other hand, do not have excitatory and inhibitory subregions and do not sum linearly over their receptive fields. Rather, complex cells generally respond to either an increase or a decrease in the intensity of a properly oriented stimulus placed anywhere within the receptive field. Hubel and Wiesel suggested that complex cell receptive fields might result from combinations of simple cell inputs.

Since the pioneering work of Hubel and Wiesel, vision scientists have used the tools of linear systems analysis to investigate the properties of visual neurons. In 1968, Campbell and Robson reported psychophysical evidence for linear mechanisms selectively sensitive to a limited range of spatial frequencies. At about the same time, physiological experiments (Campbell, Cooper & Enroth-Cugell, 1968, 1969) demonstrated that cells (both simple and complex) in cat striate cortex are selective for spatial frequency and for orientation.

The linear systems approach to cortical function is attractive because if correct, the response of a linear cell can be completely characterized with a relatively small number of measurements. A linear cell's response to any stimulus can be predicted from its responses to impulses (spots of light) flashed throughout its receptive field. Sine grating stimuli also play an important role in the linear systems approach because a linear cell transforms an input sinusoid into an output sinusoid of the same frequency; only the amplitude and phase may change.

A popular view is that simple cells act like halfwave-rectified linear operators, at least over a limited range of stimulus contrasts. The rectification is due to the fact that neurons can give only positive responses. A popular model for complex cells is that they act like energy mechanisms that compute the sum of the squared outputs of a quadrature pair of linear subunits (Adelson & Bergen, 1985).

The linear systems approach has found success in psychophysical modeling as well. It is now generally believed that the detection and identification of simple spatial patterns is mediated by linear mechanisms tuned for spatial frequency. More recently, energy mechanisms have been used to model psychophysical data on human texture discrimination (see chapters 17 and 18) and on the perception of Mach bands (Ross, Morrone & Burr, 1989).

Linear filters and energy mechanisms have also been useful in machine vision research. Quadrature pair linear filters have been used for optical flow measurement (e.g., Heeger, 1987; Watson & Ahumada, 1985; see chapter 16), texture discrimination (Bergen & Adelson, 1988; Malik & Perona, 1990; Turner, 1986), shape from shading (Pentland, 1989), and stereo disparity estimation (Sanger, 1988).

The linear systems approach has gone a long way toward explaining visual physiology, psychophysics, and machine vision. It is clear, however, that the linear/energy model falls short of a complete account of early vision. Even so, there is good reason to pursue the linear/energy paradigm. Rather than throw away the model, researchers have proposed modifying it in various ways.

One of several major objections to the linear/energy model comes from experiments that test for linearity in simple cells. Although some of the experimental data (discussed below) is consistent with linearity, some of it is not. This has led physiologists (for example, Tadmor & Tolhurst, 1989; Movshon, Thompson & Tolhurst, 1978a) to suggest replacing halfwave-rectification with over-rectification at the output of the model simple cells (halfwave-rectification clips responses less than zero whereas over-rectification clips responses less than some fixed positive threshold). In this chapter, I propose using half-squaring instead (half-squaring is halfwave-rectification followed by squaring).

A second objection to the linear/energy model comes from experiments that reveal nonspecific suppression in cortical cells. Excitation of cortical cells is highly stimulus specific, that is, cells are selective for stimulus orientation, spatial frequency, and direction of motion. The excitatory response to a preferred stimulus can be suppressed by superimposing an additional stimulus (e.g., Bonds. 1989). This suppression has been found to be largely nonspecific; it is independent of direction of motion, it is largely independent of orientation, it is broadly tuned for spatial frequency, and it is broadly tuned for spatial position.

A third objection to the linear/energy model is the fact that cell responses saturate at high contrasts. The responses of ideal linear operators and energy mechanisms, on the other hand, increase with increased stimulus contrast over the entire range of contrasts.

To explain nonspecific suppression and response saturation, several physiologists (for example, Bonds, 1989; Robson, 1988) have suggested that striate cells mutually inhibit one another, effectively normalizing their responses with respect to stimulus contrast.

In this chapter, I show that the linear/energy model taken together with half-squaring and contrast normalization yields a simple and mathematically elegant model that explains a large body of physiological data. In the next section, I explain the model: linear operators, energy mechanisms, half-squaring, and contrast normalization. In the following section, I compare model cell responses with responses of real cells for a variety of physiological measurements. Some of the experimental results can be explained simply with the linear/energy model. For these cases, I show that including half-squaring and contrast normalization does just as well. Some of the experimental results are inconsistent with the linear/energy model. I show that these results can be explained by including half-squaring and contrast normalization.

This chapter treats the visual system as a black box up to the level of striate cortex. I pay no attention to the responses of retinal or geniculate cells, but rather relate cortical cell responses directly to the time-varying stimulus intensities. In doing so I implicitly assume that the retina is at a fixed state of light adaptation.

The Model

Linear Operators and Energy Mechanisms

An operator is linear, by definition, if it obeys the property of superposition. That is, the response of the oper-

ator to a linear combination of two stimuli is equal to the linear combination of the responses to each of the component stimuli. If the response of the operator is given by $R(I)$ for stimulus I, then this property is expressed mathematically as:

$$R(aI_1 + bI_2) = aR(I_1) + bR(I_2),$$

for any constants a and b, and for any stimuli I_1 and I_2.

It is straightforward to show that a visual neuron is a linear operator (obeying the superposition property) if and only if its response is a weighted sum of the stimulus intensities. Let (x, y) denote position in the visual field, $f(x, y)$ be the receptive field of a linear operator, and $I(x, y)$ be a stimulus. The response, $R(I)$, of the cell is expressed as the inner product of the receptive field and the stimulus,

$$R(I) = f(x, y) \cdot I(x, y) = \int\int_{-\infty}^{\infty} f(x, y)I(x, y)\,dx\,dy.$$

The impulse response, $h(x, y)$, of the operator is simply related to its receptive field:

$$h(x, y) = f(-x, -y).$$

The response of the operator can be expressed as convolution with the impulse response:

$$R(I) = h(x, y) * I(x, y)|_{(x,y)=(x_0,y_0)}$$
$$= \int\int_{-\infty}^{\infty} h(\xi, \eta)I(x - \xi, y - \eta)\,d\xi\,d\eta \Big|_{(x,y)=(x_0,y_0)}$$

where $*$ denotes convolution, and where $f(x)|_{x_0} = f(x_0)$. If there were copies of the operator centered at each location in the visual field then convolving, $h(x, y) * I(x, y)$, would give their collective outputs. We consider the response of one operator by sampling the convolved output at (x_0, y_0), the center of the receptive field.

An example of a linear operator is the two-dimensional Gabor operator (see Daugman, 1980; Gabor, 1946). The receptive field of a Gabor operator is a sine grating multiplied by a two-dimensional Gaussian window, and as such it is made up of alternating excitatory and inhibitory subregions. A Gabor operator responds well to stimuli in the excitatory subregions that are brighter than the mean luminance and to stimuli in the inhibitory subregions that are darker than the mean luminance. The number of alternating subregions, the width and orientation of the subregions, and the symmetry (phase) of the operator can all be varied by changing the parameters of the sine grating and the Gaussian window that make up the Gabor operator.

The transfer function of a linear operator is defined as the Fourier transform of its impulse response, and it is made up of two parts, the amplitude response and the phase response. A linear operator is completely characterized by either its impulse response or its transfer function. For a Gabor operator, there is a peak in the amplitude response corresponding to the operator's preferred spatial frequency and orientation. The spatial frequency and orientation of a Gabor operator's underlying sine grating determine the peak in the amplitude response, and the width of the operator's Gaussian window is inversely proportional to the width (bandwidth) of the amplitude response. The phase response depends on the symmetry of the operator. For example, an operator with a central excitatory subregion flanked on either side by equal inhibitory subregions has even phase, and an operator with excitatory and inhibitory subregions to either side of center has odd phase.

Two linear operators with the same amplitude response but with phase responses that are shifted 90° in phase relative to one another are called a quadrature pair (or Hilbert transform pair). Intuitively, one often thinks of even- and odd-phase operators, like cosine- and sine-phase Gabor operators, as standard examples of quadrature pairs. Strictly speaking however, sine- and cosine-phase Gabor operators are not quadrature pairs because cosine phase Gabor operators always have some dc response (i.e., they will respond to a constant, zero contrast input), whereas sine phase Gabor operators do not. Even so, there are examples of quadrature pair operators that look very much like sine- and cosine-phase Gabor operators.

A mechanism that sums the squared outputs of a quadrature pair is called an energy mechanism (Adelson & Bergen, 1985). Its response depends only on the Fourier energy (squared magnitude of the Fourier transform) of the stimulus, not on the stimulus phase. The amplitude response of an energy mechanism is equal to the squared amplitude response of the component linear operators.

Spatiotemporal Mechanisms

The responses of striate cells depend not only on the spatial distribution of light in a stimulus, but also on its

temporal presentation. Striate cells are best thought of as spatiotemporal mechanisms.

A 3D (space-time) Gabor operator is a three-dimensional sine grating multiplied by a three-dimensional Gaussian window. The spatiotemporal frequency tuning of the operator is specified by the spatial and temporal frequencies of the underlying sine grating, and its spatial and temporal bandwidths are specified by the spread of the operator's Gaussian window. The operator looks something like a stack of plates with small plates at the top and bottom of the stack and larger plates in the middle of the stack. The plates correspond to excitatory subregions of the receptive field, and the spaces between the plates correspond to inhibitory subregions. The stack can be tilted in any orientation in space-time, each different orientation corresponding to an operator with a different spatiotemporal frequency tuning.

A spatiotemporal linear operator that is tilted along an oblique axis in space-time is direction selective. It is now well recognized that stimulus motion is like orientation in space-time, and that spatiotemporally oriented filters can be used to detect and measure it. A number of authors have proposed spatiotemporal linear operators and spatiotemporal energy mechanisms as models of the early stages of motion perception (for examples, see Adelson & Bergen, 1985; Heeger, 1987; van Santen & Sperling, 1985; Watson & Ahumada, 1983, 1985; see also chapter 16).

Mathematically, the response, $R(t)$, of a spatiotemporal linear operator is the inner product in space and the reverse correlation in time of a stimulus, $I(x, y, t)$, with the receptive field of the operator, $f(x, y, t)$,

$$R(t) = \int\int\int_{-\infty}^{\infty} f(x, y, \tau) I(x, y, \tau - t)\, dx\, dy\, d\tau. \qquad (1)$$

The response can also be expressed using convolution with the impulse response:

$$R(t) = h(x, y, t) * I(x, y, t)|_{(x,y)=(x_0,y_0)}$$

$$= \int\int\int_{-\infty}^{\infty} h(\xi, \eta, \tau) I(x - \xi, y - \eta, t - \tau)$$

$$\times\, d\xi\, d\eta\, d\tau \Big|_{(x,y)=(x_0,y_0)}$$

where $h(x, y, t) = f(-x, -y, -t)$ is the impulse response, and where (x_0, y_0) is the center of the receptive field. The response waveform, $R(t)$, is the model equivalent of a post-stimulus time histogram (PSTH). The PSTH is a measure of a cell's average response per unit time.

The receptive field of a spatiotemporal linear cell is measured as the time-varying response to impulses flashed at each point in the visual field. The transfer function of a spatiotemporal linear cell is measured using drifting sine gratings. For each stimulus spatial and temporal frequency, the response $R(t)$ is sinusoidal with frequency equal to the temporal frequency of the stimulus. The amplitudes (peak height) of the output sinusoids give the amplitude response, and the phases (relative peak latency) of the output sinusoids give the phase response.

Halfwave-Rectification and Half-Squaring

Cell firing rates are always positive, whereas linear operators can have positive or negative outputs. For a linear cell with a high maintained firing rate the positive and negative values correspond to responses above and below the maintained response. Cortical cells have very little maintained discharge so they can not truly act as linear operators.

The positive and negative outputs of a linear operator can be encoded by two halfwave-rectified linear operators. One mechanism encodes the positive outputs of the underlying linear operator, and the other one encodes the negative outputs. The two mechanisms are complements of one another, that is, the excitatory subregions of one receptive field are replaced by inhibitory subregions in the other. In other words, the two mechanisms are shifted 180° in phase relative to one another. Due to the rectification, only one of the two has a non-zero response at any given time.

For a halfwave-rectified linear cell, the receptive field of the underlying linear operator is measured using impulses of opposite polarity. Positive impulses (brighter than the mean luminance) are used to map the excitatory subregions of the receptive field, and negative impulses (darker than the mean luminance) are used to map the inhibitory subregions. The responses to dark impulses are interpreted with negative sign. The transfer function of the underlying linear operator is measured using sine grating stimuli. The response of the halfwave-rectified operator is a truncated sinusoid. The amplitude (peak height) and phase (relative peak latency) of the response are unaffected by halfwave-rectification.

A popular model of simple cells is that they are half-wave-rectified linear operators. It is suggested in this chapter that half-squaring (halfwave-rectification followed by squaring) is a more accurate model of the output nonlinearity. The output of a half-squared linear operator is given by:

$$A(t) = \left\lfloor \int\int\int_{-\infty}^{\infty} f(x, y, \tau) I(x, y, \tau - t)\, dx\, dy\, d\tau \right\rfloor^2, \qquad (2)$$

where $\lfloor x \rfloor = \max(x, 0)$ is taken to mean halfwave-rectification, $I(x, y)$ is the stimulus, $f(x, y, t)$ is the receptive field of the linear operator, and where the integral expression is copied from equation 1.

Care must be taken when interpreting measurements of the receptive field and transfer function of a half-squared linear cell. Responses to impulses and gratings do not give the receptive field and transfer function of the underlying linear operator. The amplitude (peak height) of the response waveform for a drifting grating stimulus is the square of the amplitude of the underlying linear operator. The phase (relative peak latency) of the response waveform is unaffected by the squaring. The response to positive impulses minus the response to negative impulses gives the signed-square of the receptive field of the underlying linear operator. That is,

$$R'(t) = \begin{cases} R^2(t) & \text{if } R(t) > 0 \\ -R^2(t) & \text{if } R(t) < 0, \end{cases}$$

where $R'(t)$ is the signed-square of $R(t)$, and $R(t)$ is the response of the underlying linear operator.

Although the term "amplitude response" should be reserved for linear operators, I use it in this chapter when writing about half-squared operators and energy mechanisms as well. In both cases, the "amplitude response" is measured using sine grating stimuli. For half-squared linear operators we measure the Fourier amplitude of the fundamental component of the response waveform (this is proportional to the peak height of the response waveform). Energy mechanisms give an unmodulated response to sine gratings so we measure the dc response.

An energy mechanism can be constructed as the average of the outputs of four half-squared linear operators, all four with the same "amplitude response," but with phases in steps of 90°. The energy output, $E(t)$, is expressed as:

$$E(t) = (1/4)[A^0(t) + A^{90}(t) + A^{180}(t) + A^{270}(t)], \qquad (3)$$

where $A^\phi(t)$ is the response of a half-squared linear operator, and where the superscript, ϕ, specifies the operator's phase in degrees. The "amplitude response" of the energy mechanism is the same as the "amplitude response" of each of the input half-squared linear operators.

Contrast Normalization

The responses of linear and energy mechanisms increase with stimulus contrast over the entire range of contrasts. Information about a visual stimulus, other than its contrast, is represented as the relative response of a collection of mechanisms. For example, the orientation of a grating is represented as the relative response of a collection of energy mechanisms, each with a different orientation tuning. The ratio of the responses of two of the mechanisms is fixed, independent of stimulus contrast. Likewise, consider the response of one mechanism when presented with two differently oriented gratings. If the contrast of both gratings is changed by the same factor then the ratio of the responses to the orientations remains unchanged.

It is possible that the visual system also represents visual information as the relative response of collections of cells. For this to be the case, it is crucial that the ratio of a cell's responses to two stimuli be independent of contrast. But cortical cells, unlike linear or energy mechanisms, have a limited dynamic range, so their responses saturate for high contrasts. How is it possible for response ratios to be independent of stimulus contrast, in the face of response saturation? I propose in this chapter that cell responses are normalized for stimulus contrast.

The contrast-normalization mechanism discussed here is analogous to models of retinal light adaptation and gain-control (see Sperling & Sondhi, 1968 for an example, and Shapley & Enroth-Cugell, 1984 for a review), the purpose of which is to keep the retinal response approximately the same when the level of illumination changes. That way, the brain can proceed to process visual information without having to attend to the light level. The consequence of retinal light adaptation is that much of our perception is invariant with respect to intensity, over a wide range of light levels. For example, the perceived contrast of a grating stimulus is largely invariant with respect to intensity.

Likewise, contrast-normalization allows the brain to process visual information without having to attend further to the contrast. For example, the perceived orienta-

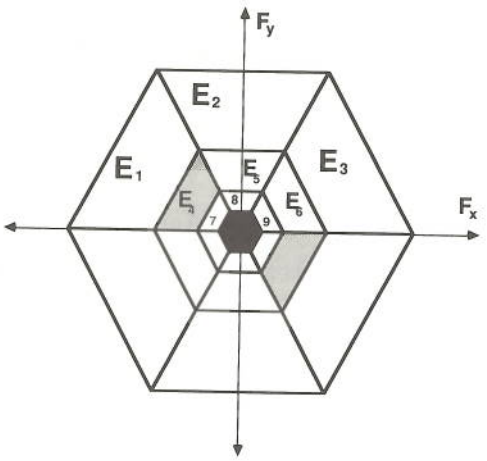

Fig. 9.1
Diagram of the frequency domain, partitioned by nine oriented energy mechanisms. Each subregion, labeled E_1 through E_9, represents the "amplitude response" of one energy mechanism. For example, the E_2 mechanism responds to high frequency horizontal gratings. The "amplitude responses" are drawn schematically here; for actual energy mechanisms the regions are more rounded and they overlap somewhat.

tion of a grating stimulus is largely invariant with respect to contrast.

Figure 9.1 illustrates the partition of the frequency domain by a collection of energy mechansims. Each region in the diagram corresponds to the "amplitude response" of one energy mechanism. Contrast normalization of each mechanism is realized by dividing its output by the total energy at all orientations and nearby spatial frequencies. The normalized energy, $\overline{E}_i(t)$, is given by:

$$\overline{E}_i(t) = \frac{E_i(t)}{\sigma^2 + \sum_{j=1}^{9} E_j(t)}, \tag{4}$$

where σ is a constant in the denominator, known as the semisaturation constant. As long as σ is nonzero, the normalized output will always be finite, even for a zero contrast stimulus. In fact, the normalized output will always have a value between 0 and 1, saturating for high contrasts.

The underlying linear operators can be chosen so that the sum of their squared amplitude responses is the unit constant function (everywhere equal to one). Then the summation in the denominator gives the total Fourier energy in an annulus of spatial frequencies.

Operators tuned to different orientations (e.g., corresponding to each of the regions labeled E_4, E_5, and E_6 in figure 9.1) are all normalized together. That is, each is divided by the sum of all nine energy outputs at all orientations and nearby spatial frequencies, plus σ^2. Within each spatial frequency band, the ratio of the outputs of two normalized operators is invariant with respect to stimulus contrast. The ratio is maintained even though the normalized outputs saturate at high contrasts. E_1, E_2, and E_3 are normalized by a different set including higher frequencies and excluding E_7, E_8, and E_9. Likewise for the lower frequency operators.

Feedback Normalization

Contrast normalization given by equations 2, 3, and 4 is expressed in feedforward manner. First, the A_i^{ϕ}'s are computed, then they are combined to give the E_i's, and then the E_i's are combined to give the $\overline{E}_i$'s. However, the unnormalized A_i^{ϕ}'s and E_i's cannot be represented by mechanisms (e.g., neurons or 8-bit computers) with limited dynamic range.

The solution is to use a feedback network to do the normalization. Then, the A_i^{ϕ}'s and E_i's need not be explicitly represented as cell output firing rates. The details of the feedback normalization network are beyond the scope of this chapter, but will be discussed in a forthcoming paper.

One consequence of using a feedback network to achieve the normalization is that the feedback signal must be averaged over space and/or time to avoid unstable oscillations in the output. For spatially extended periodic stimuli, the feedback reaches a steady state after a brief period of time. The spatial and temporal pooling of the contrast normalization signal is left unspecified in this chapter, since it only deals with steady state responses.

Model Simple and Complex Cells

The model in its entirety is depicted in figure 9.2. Simple cells are modeled as contrast normalized and half-squared linear operators. Model complex cells are built on outputs of model simple cells. The contrast normalization feedback signal is combined from all orientations and nearby spatial frequencies.

The various stages of the model are as follows. Linear operators of four different phases are applied to the stimulus. The outputs of these operators are then half-squared and normalized to give the model simple cell responses:

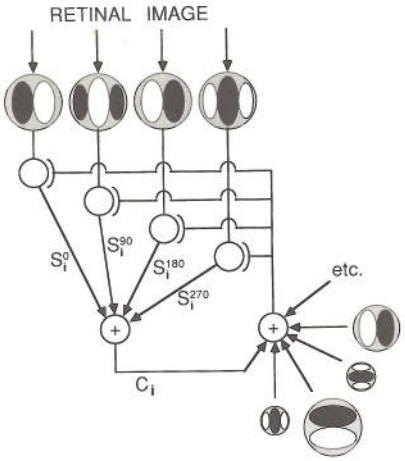

Fig. 9.2
Illustration of the various stages of the model. Linear receptive fields are depicted as circles, subdivided into bright and dark subregions. The S_i^ϕ labels represent simple cell outputs, and the C_i label represents a complex cell output. The feedback signal is the combined energy at all orientations and nearby spatial frequencies, averaged over space and time. The feedback signal suppresses the simple cell responses by way of divisive inhibition.

$$S_i^\phi(t) = k\frac{A_i^\phi(t)}{\sigma^2 + \frac{1}{4}\sum_j \sum_\phi A_j^\phi(t)}$$

$$= k\frac{A_i^\phi(t)}{\sigma^2 + \sum_j E_j(t)}, \qquad (5)$$

where S_i^ϕ is the response of a model simple cell with phase, ϕ, σ is the semisaturation constant, k is a constant scale factor that determines the maximum attainable firing rate, and $A_i^\phi(t)$ is defined in equation 2. The subscript i is an index used to specify the "amplitude response" of each operator. The model complex cell responses are computed by averaging the simple cell responses:

$$C_i(t) = (1/4)\sum_\phi S_i^\phi(t)$$

$$= (k/4)\sum_\phi \frac{A_i^\phi(t)}{\sigma^2 + \frac{1}{4}\sum_j \sum_\phi A_j^\phi(t)}$$

$$= k\frac{E_i(t)}{\sigma^2 + \sum_j E_j(t)}. \qquad (6)$$

Note again that if the underlying linear operators are chosen correctly then the denominator of equations 5 and 6 gives the Fourier energy of the stimulus within an annulus of spatial frequencies. As mentioned above, space-time averaging of the contrast-normalization is not included in the above equations because this chapter deals only with steady state responses.

Striate Cell Responses

The previous section describes a nonlinear model of striate cell responses. This section reviews some of the electrophysiological data on the responses of simple and complex cells in cat striate cortex and compares model cell responses, given by equations 5 and 6, with electrophysiological data.

The results in this section demonstrate that simple cells behave more like half-squared linear operators than like halfwave-rectified linear operators. The results in this section also demonstrate that response saturation can be attributed to contrast normalization.

Simple Cells

Receptive Field

Simple cells have clearly defined excitatory and inhibitory subregions. Bright (brighter than the mean luminance) light in an excitatory region or dim light in an inhibitory region enhances a simple cell's response, whereas bright light in an inhibitory region or dim light in excitatory region inhibits its response. The excitatory and inhibitory subregions are also called "on" and "off" regions; the cell responds to light increment (the onset of a bright stimulus) in an "on" region, and to light decrement (the offset of a bright stimulus) in an "off" region. Superimposed on each "on" subregion is "off" inhibition and vice versa.

Although many physiologists have used oriented bars, edges, and gratings to generate one-dimensional maps of receptive fields, relatively few (e.g., Jones & Palmer, 1987) have measured the two-dimensional spatial structure of receptive fields. Some researchers have mapped the three-dimensional spatiotemporal structure of simple cell receptive fields (Hamilton, Albrecht & Geisler, 1989; McLean & Palmer, 1989; see chapter 8).

Frequency Domain vs. Space Domain

Many experimenters (Andrews & Pollen, 1979; Kulikowski & Bishop, 1981; Maffei, Morrone, Pirchio & Sandini, 1979; Movshon et al., 1978a; Tadmor & Tolhurst, 1989) have tested for linearity of simple cells by comparing cells' transfer functions with their impulse responses. The

logic of these experiments is straightforward. The response of a linear cell to the sum of two stimuli is equal to the sum of the responses to each of the component stimuli. Since a grating is composed of the sum of a number of impulses, the response of a linear cell to a grating is predictable from its response to impulses. Likewise, since an impulse can be thought of as the sum of a number of gratings, the response to an impulse is predictable from the response to gratings.

The results of these experiments show that simple cell response to gratings and their response to impulses look very nearly like Fourier transforms of one another, up to an arbitrary scale factor. These results have been taken as evidence for linearity.

Most of these researchers found, however, that the response to gratings and the response to impulses are not precise transforms of one another. In many cases, the inverse transform of the response to gratings gives a receptive field profile with additional side bands beyond those measured directly. In addition, the measured response to gratings is often more narrowly tuned than predicted from the Fourier transform of the response to impulses.

Several experimenters (e.g., Tadmor & Tolhurst, 1989) suggest that the discrepancy between the frequency and space domain measurements can be explained by over-rectification. If the neuron has to reach a certain level of excitation before any activity is seen, there will be a disproportionate decrease in small responses.

The results are also predicted by half-squaring. With half-squaring instead of halfwave-rectification, the response to impulses and the response to gratings are not transforms of one another. But in spite of the nonlinearity, the inverse transform of the response to gratings still looks very similar to the response to impulses, when they are rescaled relative to one another by an appropriate scale factor. As shown in figure 9.3A, the inverse transform of the response to gratings has some extra (low amplitude) side bands, in agreement with much of the physiological data. As shown in figure 9.3B, the response to gratings is more narrowly tuned than predicted from the response to impulses, also in agreement with physiological data.

The point of this demonstration is that the experiments do not distinguish well between halfwave-rectification, over-rectification, and half-squaring. The experimental re-

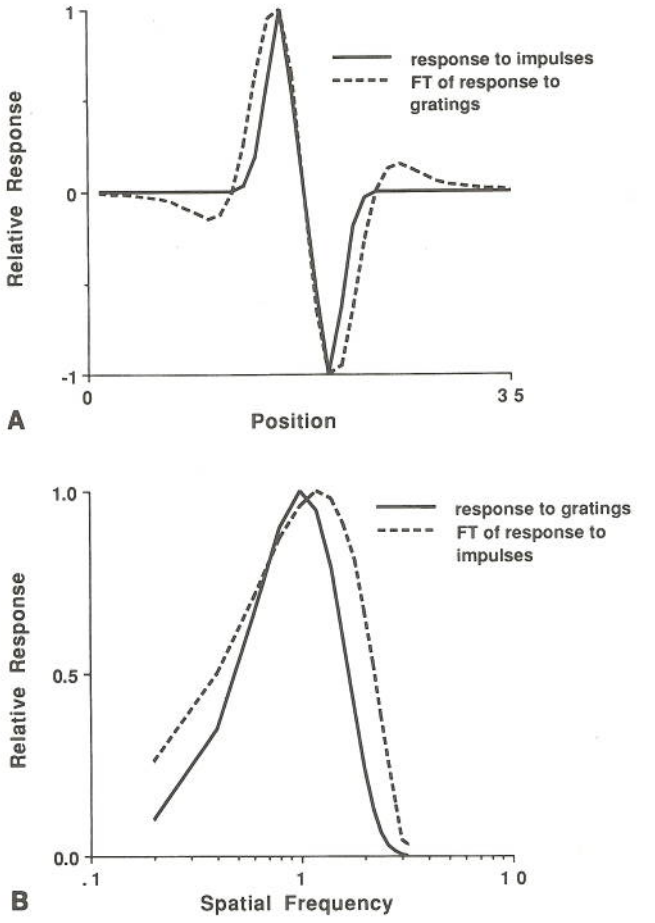

Fig. 9.3
A comparison between the response to impulses and the response to gratings of a half-squared Gabor operator. Response to gratings measured as the Fourier amplitude of the response, for each stimulus spatial frequency. Response to impulses measured as the signed-square of the underlying linear operator. Since the operator is nonlinear, the response to impulses and response to gratings are not transforms of one another. (A) Response to impulses superimposed with the inverse Fourier transform of response to gratings. (B) Response to gratings superimposed with the magnitude of the Fourier transform of response to impulses. The response to gratings is more narrowly tuned than predicted from the response to impulses, and the response to impulses is more narrowly tuned than predicted from the response to gratings. This is in agreement with much of the physiological data.

sults can not be taken as evidence to favor one model over the other.

Responses to Counterphase Gratings

Simple cells exhibit characteristic responses to temporally modulated (e.g., counterphase) sine gratings. The response varies over time with the temporal modulation of the stimulus, and the amplitude and phase of modulation both depend on the spatial phase of the grating (Kulikowski & Bishop, 1981; Maffei & Fiorentini, 1973; Movshon et al., 1978a; Reid et al, 1987; see chapter 8). Reid et al. (1987) and Movshon et al. (1978a) have measured response amplitude and response phase of simple cells while varying the spatial phase of counterphase gratings. They have both shown (mathematically) that for a halfwave-rectified spatiotemporal linear operator, a polar plot of the response amplitude as a function of the response phase is elliptical in shape. Their experimental results, however, are typically not quite elliptical. Rather the results have been described as "wasp-waisted ellipses" since the amplitudes near the minor axes are smaller than they should be to fit an ellipse. Movshon et al. proposed that this deviation could be explained by over-rectification.

The wasp-waisted elliptical shape is also predicted by half-squaring. In figure 9.4, the output of a spatiotemporal Gabor operator was computed for counterphase gratings of various spatial and temporal frequencies, and various spatial phases. The response waveforms were then half-squared, and the amplitude and phase of the fundamental Fourier component (equal to the temporal frequency of the stimulus) were measured. Figure 9.4 plots relative response amplitude as a function of response phase for two different spatial and temporal frequencies. The plots are similar in shape to physiological data. Like the results discussed above, this experiment does not distinguish well between over-rectification and half-squaring.

Counterphase vs. Drifting Gratings

By comparing the responses to counterphase stimuli with responses to drifting grating stimuli, Reid et al. (1987) have demonstrated that there is a nonlinear contribution to the direction selectivity of simple cells. They computed a directional index given by $(R_p - R_a)/(R_p + R_a)$, where R_p and R_a are, respectively, the responses for gratings drifting in the preferred and antipreferred directions. Reid et al. showed that for a halfwave-rectified spatiotemporal

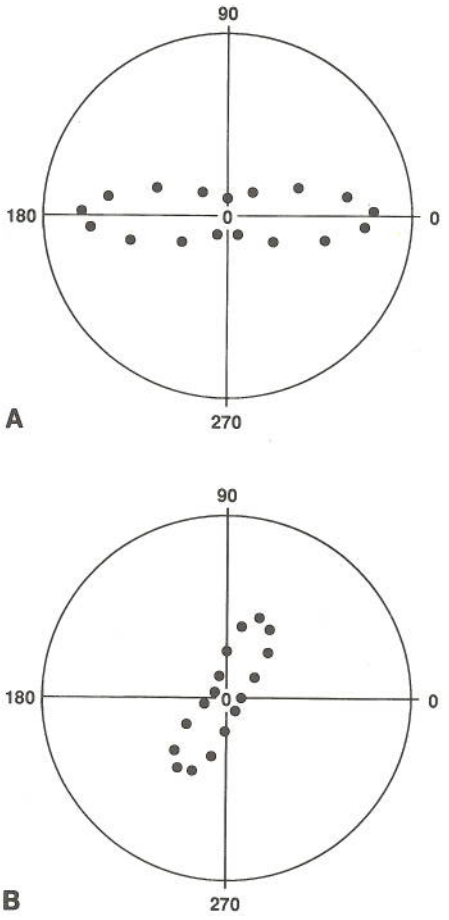

Fig. 9.4
Responses of a half-squared linear operator for counterphase grating stimuli of varying spatial phase. The amplitude of the fundamental component of the response is represented radially, while the angular coordinate indicates the temporal phase of the response. The stimulus spatial and temporal frequencies are different for the two plots. These polar plots of response amplitude versus response phase are shaped like wasp-waisted ellipses, in agreement with physiological data.

linear operator this directional index is predictable from the responses to counterphase gratings. Specifically, the directional index is equal to the ratio of the axes of the ellipse obtained, as described above, from counterphase grating stimuli. However, Reid et al. found that these two measures of the directional index do not agree; the prediction from counterphase stimuli underestimates the index by about half.

Half-squaring explains this result reasonably well. Figure 9.5 plots the ratio of the axes of the best fitting ellipse derived from counterphase stimuli against the directional

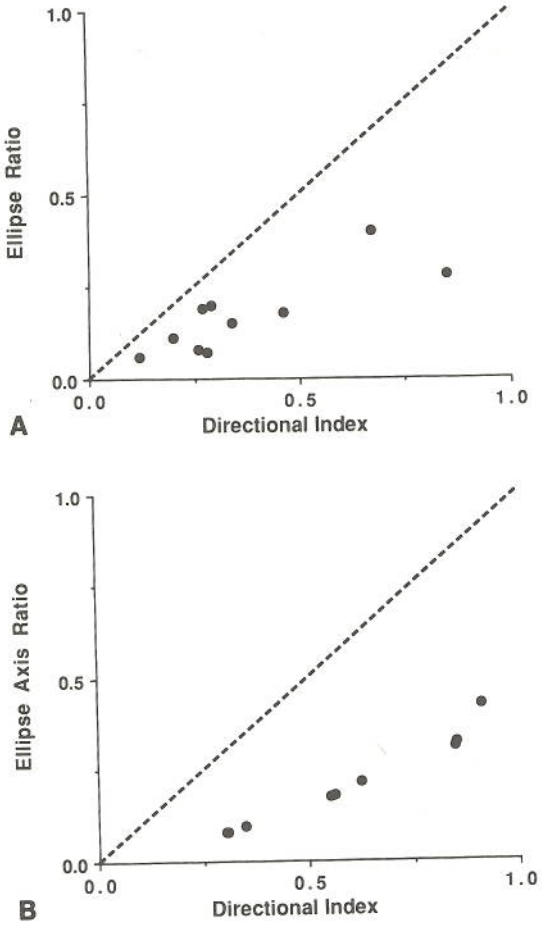

Fig. 9.5
Directional index predicted from counterphase grating stimuli versus that measured directly from drifting gratings for: (A) a direction selective simple cell (replotted from Reid et al., 1987), and (B) a half-squared spatiotemporal Gabor operator. Each point is for a different stimulus spatial and temporal frequency. The dotted line is the prediction of a halfwave-rectified linear operator. For both model cells and real cells, the directional index predicted from counterphase stimuli underestimates that measured directly with drifting gratings.

index measured with drifting grating stimuli. The figure shows that for a half-squared Gabor operator, the directional index predicted from counterphase stimuli underestimates that measured directly with drifting gratings, in a manner very similar to the experimental results.

Complex Cells

Complex cells are clearly nonlinear as they respond to either a bright or a dim stimulus placed anywhere within their receptive fields. In experiments performed by Movshon et al. (1978b) a bar fixed in one position was flashed simultaneously with a second bar, of the same or opposite polarity, that could appear in one of several positions around the location of the fixed bar. By measuring the influence of the second bar upon the response to the first, Movshon et al. were able to demonstrate that complex cell receptive fields are composed of subunits. The subunits have clearly defined spatial profiles with excitatory and inhibitory subregions. The subunit outputs are rectified before being combined into the complex cell response. In addition, Movshon et al. measured complex cell responses to grating stimuli. They determined that the inverse Fourier transform of the response to gratings matches the spatial profile of the underlying subunits, up to an arbitrary scale factor. The present model is in general agreement with these results. The subunits of model complex cells are model simple cells with identical "amplitude response."

Emerson and coworkers (1987, 1991) analyzed responses of complex cells to stimuli made up of pairs of bars flashed in sequence. The response to a pair of bars is different from the sum of the responses to each individual bar. For some spatial and temporal separations between the bars the cell response is greater than the linear prediction, and for other separations it is less than the linear prediction. Emerson et al. (1991) have shown that the nonlinear interaction between pairs of bars is consistent with that predicted by energy mechanisms.

A variety of other experiments also indicate that energy mechanisms are reasonable models of complex cells. An energy mechanism has an unmodulated response to drifting sine grating stimuli, as do the majority of complex cells (Maffei & Fiorentini, 1973; Movshon 1978b).

For counterphase grating stimuli, both complex cells and energy mechanisms have responses that vary over time at twice the temporal frequency of the stimulus (Movshon et al., 1978b). In addition, responses to counterphase gratings do not depend on the spatial phase of the stimulus (Maffei & Fiorentini, 1973; Movshon et al., 1978b).

Contrast-Response

The inclusion of contrast normalization in the model results in response saturation. The contrast-response function (that is, response as a function of log contrast for sine grating stimuli of optimal spatial frequency and orientation) for model cells is qualitatively similar to typical

experimentally measured contrast-response relationships in both cat and primate (Albrecht & Hamilton, 1982; Dean, 1981; Maffei & Fiorentini, 1973; Ohzawa, Sclar & Freeman, 1985; Sclar, Maunsell & Lennie, 1990).

The constrast-response functions for striate cells in both cat and primate have been fitted by the hyperbolic ratio function (Albrecht & Hamilton, 1982; Li & Creutzfeldt, 1984; Sclar et al., 1990):

$$R = R_{\max}\frac{c^n}{\sigma^n + c^n} + M, \qquad (7)$$

where R is the evoked response, c is the contrast of the test grating, M is maintained discharge, n is a constant, σ is the semisaturation constant, and $R_{\max}$ is the maximum attainable response.

With parameters $n = 2$ and $M = 0$, the contrast-response function given by equation 7 is equivalent to that of model cells given by equations 5 and 6. The equivalence is easily demonstrated by recalling that the summation in the denominator of equations 5 and 6 is proportional to c^2.

Physiological data from both cat and primate show that the exponent in the contrast-response function does not differ significantly between populations of simple and complex cells (Albrecht & Hamilton, 1982; Dean, 1981). The exponent is 2 on average, but there is variability from cell to cell (Albrecht & Hamilton, 1982; Sclar et al., 1990).

In the present model, the contrast-response functions of both simple and complex cells have exponents of 2, because of half-squaring. If the model simple cells were halfwave-rectified rather than half-squared, then their exponent would instead be 1.

Contrast Dependence of Tuning

The contrast-response curve of a model cells shifts downward (on log-log axes) if the orientation of the test grating is non-optimal. This property of model cells is easily explained by equations 5 and 6. In each of these equations the value of the numerator depends on stimulus orientation because the underlying linear operator is orientation tuned. The value of the denominator does not depend on stimulus orientation because the suppression is pooled equally over all orientations. If the suppression was broadly tuned for orientation then the contrast-response curves would shift downward and rightward. The relative amount of rightward shift depends on the breadth of the tuning.

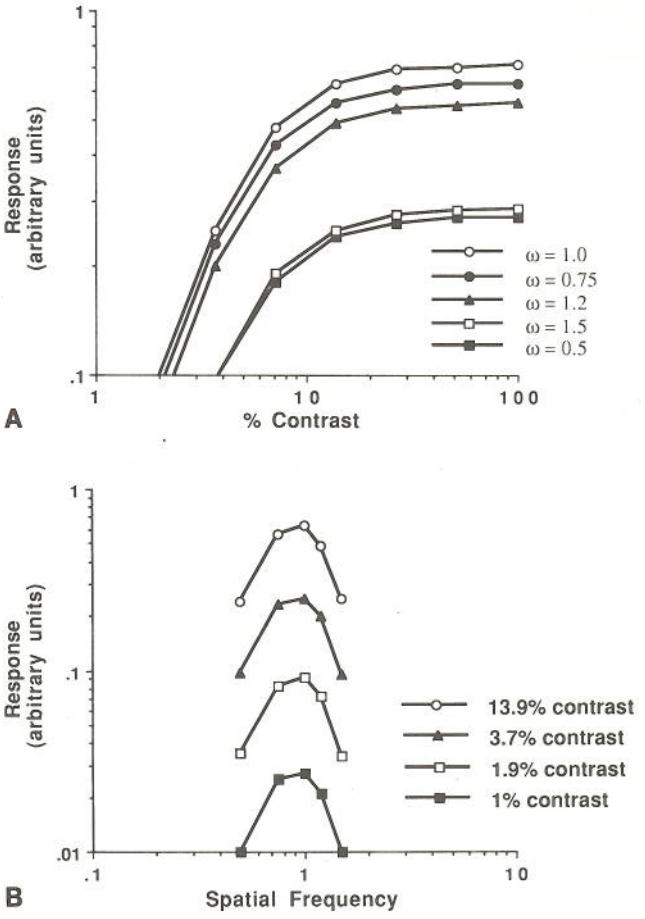

Fig. 9.6
Model complex cell response for grating stimuli of various contrasts and spatial frequencies. (A) and (B) show the same data plotted in different ways. (A) Response versus contrast as the spatial frequency, ω, of the stimulus is varied. The contrast-response curve shifts downward and very slightly rightward (not visible in the graph) in the log-log plot if the spatial frequency of the test grating is nonoptimal. (B) Spatial frequency tuning curves as the contrast of the stimulus is varied. Tuning width is largely invariant with respect to contrast. Width broadens very slightly (not visible in the graph) with increased contrast. Both of these results are in agreement with physiological data.

Figure 9.6A shows the contrast-response curves of a model cell for various stimulus spatial frequencies. The curves shift downward and slightly rightward for nonoptimal spatial frequencies. The small rightward shift occurs because the suppression from contrast normalization is broadly tuned for spatial frequency. If the suppression

were equal for all spatial frequencies then there would be no rightward shift.

Downward shifts of contrast-response have been measured physiologically by Li and Creutzfeldt (1984) for stimuli of nonoptimal orientation, for stimuli of nonpreferred direction of motion, and for stimuli in the nondominant eye. Li and Creutzfeldt, and other authors, interpreted these results as demonstrating that saturation of the contrast-response curve is already present at the precortical level and therefore not due to intracortical mechanisms. On the contrary, the model predicts the downward shift precisely because of the mutual suppression between cortical cells. Albrecht and Hamilton (1982) recorded similar downward shifts of contrast-response curves for stimuli of nonoptimal spatial frequency, but they also found a slight rightward shift of the curves, in agreement with the present model.

The orientation tuning width of model cells is invariant with respect to contrast. Changing the stimulus contrast scales the response by a constant factor over all orientations. In other words, the ratio of the responses produced by different orientations remains fixed, independent of contrast. Contrast has no impact on tuning width because the suppression from contrast normalization is equal for all stimulus orientations. If the suppression was broadly tuned for orientation then the model cell's tuning width would depend on contrast.

Figure 9.6B shows the spatial frequency tuning curves of a model cell for various contrasts. The spatial frequency bandwidth broadens very slightly with increased contrast, because the suppression from contrast normalization is broadly tuned for spatial frequency. Experimental results demonstrate that real cells behave similarly. In spite of response saturation, orientation and spatial frequency tuning widths vary little with contrast (Albrecht & Hamiltom, 1982; Li & Creutzfeldt, 1984; Sclar & Freeman, 1982).

Discussion

For some years, simple cells have been modeled as halfwave-rectified linear operators. It has also become popular to model complex cells as energy mechanisms. A variety of experimental results provide evidence in support of the linear/energy model, but a variety of other experimental results can not be explained by the linear/energy model.

In this chapter I have suggested two modifications to the linear/energy model in order to explain a larger body of physiological data. One modification is the use of half-squaring instead of halfwave-rectification (or over-rectification) at the output of the model simple cells. The other modification is to include a contrast normalization nonlinearity.

With contrast normalization and squaring, the contrast-response curves of model cells are very similar to physiological measurements. Contrast-response curves of model cells shift mostly downward for nonoptimal stimuli, and the tuning widths of model cells are largely independent of contrast (figure 9.6). In other words, the ratio of the responses produced by two different stimuli is largely invariant with respect to stimulus contrast. In this way, information about a visual stimulus, other than its contrast, can be represented as the relative responses of a collection of cells.

Contrast normalization explains some additional properties typical of striate cells including nonspecific suppression (e.g., Bonds. 1989) mentioned in the introduction to this chapter, and contrast adaptation (e.g., Ohzawa et al., 1985). These issues will be addressed in a forthcoming paper.

In addition, the results of a variety of psychophysical experiments have been modeled using contrast normalization. For example, both Bergen and Landy (see chapter 17), and Graham (see chapter 18) include contrast normalization steps in their models of texture discrimination.

In this chapter, I discuss three models of simple cell rectification: halfwave-rectification, over-rectification, and half-squaring. While there is ample evidence to reject halfwave-rectification, the experiments to date are generally consistent with both over-rectification and half-squaring. This is not surprising since these two nonlinearities are approximately the same over a restricted operating range. Mathematically, the two nonlinearities are expressed as:

$$N_1(x) = a_1 \lfloor x \rfloor^2$$

$$N_2(x) = a_2 \lfloor x - T \rfloor,$$

where $N_1(x)$ and $N_2(x)$ are half-squaring and over-rectification, respectively, a_1 and a_2 are scale factors, T is a threshold, and $\lfloor x \rfloor = \max(x, 0)$ is taken to mean half-

wave rectification. By choosing a_2 and T appropriately, $N_2(x)$ can be made to approximate $N_1(x)$ for a certain range of x values.

These two functions approximate one another only for a certain range of values, so it may be possible to discriminate between them experimentally. For example, the results of the Reid et al. experiment (figure 9.5) can be reanalyzed to test the different hypotheses. One can compensate for the effect of half-squaring by taking the square root of the responses. After taking square roots, the directional indices measured from drifting gratings should be equal to the ellipse ratios measured from counterphase gratings. Likewise, one can compensate for the effect of over-rectification by adding a fixed constant value to the responses. More generally, one could try fitting the experimental data using the following nonlinearity:

$$N(x) = a\lfloor x - T \rfloor^n. \tag{8}$$

The choice of exponent, n, and threshold, T, that yield the best fit would be measured from the data. Half-squaring corresponds to the case in which $n = 2$ and $T = 0$. Over-rectification corresponds to the case in which $n = 1$ and $T \neq 0$.

Although the present model explains a large body of physiological data, there is variability in the behavior from one cell to the next. Some, but not all, of the variability between cells can be accounted for by the present model's parameters. For example, there are some complex cells that, unlike energy mechanisms, have modulated responses to gratings of certain spatial frequencies.

As another example, half-squaring predicts that polar plots of response to counterphase gratings are shaped like wasp-waisted ellipses (figure 9.4). Half-squaring also predicts that the ellipse axis ratio underestimate the directional index (figure 9.5). For some cells, however, the ellipse axis ratio underestimates the directional index even though polar plots of response to counterphase gratings are well fit by ellipses (not wasp-waisted).

Future research will aim to enhance the model so that it might account for some of the variability between cells. For example, it is possible to account for some of the variability in contrast-response measurements by adding a threshold parameter to the model. Contrast-response of model cells would then be given by:

$$R = R_{\max} \frac{\lfloor c - T \rfloor^n}{\sigma^n + c^n} + M. \tag{9}$$

If T is negative and M is positive then the cell would have a nonzero maintained discharge. If T is positive and M is negative then the cell would be over-rectified.

Varying T in equation 9 not only changes the level of maintained discharge, but also changes the shape of the contrast-response curve. Varying T and σ simultaneously can look very much like a change in the exponent, n. This suggests the possibililty of fitting contrast-response data with a fixed exponent of 2, accounting for the variability from cell to cell by changing the value of T.

For a particular cell, the parameters of the model can be measured and compared using different experiments. For example, the exponent, n, and threshold, T, can be measured by fitting contrast-response data to equation 9. These parameters can also be measured, as discussed above, by using equation 8 to fit the data of Reid et al.

As another example, recall that the contrast-response curve of model cells shifts downward and slightly rightward if the spatial-frequency of the test grating is non-optimal. The amount of rightward shift depends on the breadth of tuning of the contrast-normalization suppression. For a particular cell, one could measure both the breadth of tuning and the shift in contrast-response, and then compare the two.

The ultimate goal of modeling visual neurons is to be able to predict the response of a cell to any stimulus, based on a limited number of measurements. The success of this endeavor hinges on pinpointing the regularities in the behavior of visual neurons. A good model will have a small number of easily measured parameters and will be able to predict the responses of a large number of cells.

Acknowledgments

This chapter benefitted greatly from discussions with Ted Adelson, Beau Watson, Jim Bergen, John Robson, A. B. Bonds, and Tony Movshon.

References

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

Albrecht, D. G. & Hamilton, D. B. (1982). Striate cortex of monkey and cat: Contrast response function. *Journal of Neurophysiology, 48,* 217–237.

Andrews, B. W. & Pollen, D. A. (1979). Relationship between spatial frequency selectivity and receptive field profile of simple cells. *Journal of Physiology (London), 287,* 163–176.

Bergen, J. R. & Adelson, E. H. (1988). Early vision and texture perception. *Nature, 333,* 363–367.

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, F. W., Cooper, G. F. & Enroth-Cugell, C. (1968). The angular selectivity of visual cortical cells to moving gratings. *Journal of Physiology (London), 198,* 237–250.

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

Campbell, F. W. & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology (London), 197,* 551–556.

Daugman, J. G. (1980). Two-dimensional spectral analysis of cortical receptive field profiles. *Vision Research, 20,* 846–856.

Dean, A. F. (1981). The relationship between response amplitude and contrast for cat striate cortical neurons. *Journal of Physiology (London), 318,* 413–427.

Emerson, R. C., Bergen, J. R. & Adelson, E. H. (1991). Directionally selective complex cells and the computation of motion energy in cat visual cortex. *Vision Research,* in press.

Emerson, R. C., Citron, M. C., Vaughn, W. J. & Klein, S. A. (1987). Nonlinear directionally selective subunits in complex cells of cat striate cortex. *Journal of Neurophysiology, 58,* 33–65.

Gabor, D. (1946). Theory of communication. *JIEE (London), 93,* 429–457.

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

Heeger, D. J. (1987). Model for the extraction of image flow. *Journal of the Optical Society of America A, 4,* 1455–1471.

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

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

Kulikowski, J. J. & Bishop, P. O. (1981). Linear analysis of the response of simple cells in the cat visual cortex. *Experimental Brain Research, 44,* 386–400.

Li, C. & Creutzfeldt, O. (1984). The representation of contrast and other stimulus parameters by single neurons in area 17 of the cat. *Pflugers Archives, 401,* 304–314.

Maffei, L. & Fiorentini, A. (1973). The visual cortex as a spatial frequency analyzer. *Vision Research, 13,* 1255–1267.

Maffei, L., Morrone, C., Pirchio, M. & Sandini, G. (1979). Responses of visual cortical cells to periodic and nonperiodic stimuli. *Journal of Physiology (London), 296,* 27–47.

Malik, J. & Perona, P. (1990). Preattentive texture discrimination with early vision mechanisms. *Journal of the Optical Society of America A, 7,* 923–931.

McLean, J. & Palmer, L. A. (1989). Contribution of linear spatiotemporal receptive field structure to velocity selectivity of simple cells in area 17 of cat. *Vision Research, 29,* 675–679.

Movshon, J. A., Thompson, I. D. & 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.

Movshon, J. A., Thompson, I. D. & Tolhurst, D. J. (1978b). Receptive field organization of complex cells in the cat's striate cortex. *Journal of Physiology (London), 283,* 79–99.

Ohzawa, I., Sclar, G. & Freeman, R. D. (1985). Contrast gain control in the cat's visual system. *Journal of Neurophysiology, 54,* 651–667.

Pentland, A. P. (1989). A possible neural mechanism for computing shape from shading. *Neural Computation, 1,* 208–217.

Reid, R. C., Soodak, R. E. & Shapley, R. M. (1987). Linear mechanisms of directional selectivity in simple cells of cat striate cortex. *Proceedings of the National Academy of Science, 84,* 8740–8744.

Robson, J. G. (1988). Linear and nonlinear operations in the visual system. *Investigative Ophthalmology and Visual Science Supplement, 29,* 117.

Ross, J., Morrone, M. C. & Burr, D. C. (1989). The conditions under which Mach bands are visible. *Vision Research, 29,* 699–715.

Sanger, T. (1988). Stereo disparity computation using Gabor filters. *Biological Cybernetics, 59,* 405–418.

Sclar, G. & Freeman, R. D. (1982). Orientation selectivity of the cat's striate cortex is invariant with stimulus contrast. *Experimental Brain Research, 46,* 457–461.

Sclar, G., Maunsell, J. H. R. & Lennie, P. (1990). Coding of image contrast in central visual pathways of the macaque monkey. *Vision Research, 30,* 1–10.

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

Sperling, G. & Sondhi, M. M. (1968). Model for visual luminance discrimination and flicker detection. *Journal of the Optical Society of America, 58,* 1133–1145.

Tadmor, Y. & Tolhurst, D. J. (1989). The effect of threshold on the relationship between the receptive-field profile and the spatial-frequency tuning curve in simple cells of the cat's striate cortex. *Visual Neuroscience, 3,* 445–454.

Turner, M. R. (1986). Texture discrimination by Gabor functions. *Biological Cybernetics, 55,* 71–82.

van Santen, J. P. H. & Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America A, 2,* 300–321.

Watson, A. B. & Ahumada, A. J. (1983). *A look at motion in the frequency domain.* Technical Report 84352, NASA-Ames Research Center.

Watson, A. B. & Ahumada, A. J. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America A, 2,* 322–342.