Motion selectivity and the contrast-response function of simple cells in the visual cortex

DUANE G. ALBRECHT AND WILSON S. GEISLER
Department of Psychology, University of Texas, Austin
(RECEIVED November 5, 1990; ACCEPTED April 8, 1991)

Abstract
The responses of simple cells were recorded from the visual cortex of cats, as a function of the position and contrast of counterphase and drifting grating patterns, to assess whether direction selectivity can be accounted for on the basis of linear summation. The expected responses to a counterphase grating, given a strictly linear model, would be the sum of the responses to the two drifting components. The measured responses were not consistent with the linear prediction. For example, nearly all cells showed two positions where the responses approached zero (i.e. two “null phase positions’’); this was true, even for the most direction selective cells. However, the measured responses were consistent with the hypothesis that direction selectivity is a consequence of the linear spatiotemporal receptive-field structure, coupled with the nonlinearities revealed by the contrast-response function: contrast gain control, halfwave rectification, and expansive exponent. When arranged in a particular sequence, each of these linear and nonlinear mechanisms performs a useful function in a general model of simple cells. The linear spatiotemporal receptive field initiates stimulus selectivity (for direction, orientation, spatial frequency, etc.). The expansive response exponent enhances selectivity. The contrast-set gain control maintains selectivity (over a wide range of contrasts, in spite of the limited dynamic response range and steep slope of the contrast-response function). Rectification conserves metabolic energy.

Keywords: Visual cortex, Receptive fields, Direction selectivity, Contrast gain control, Contrast response function, Motion

Introduction
The ability to sense motion is crucial for vision and visually guided behavior. Neurons in the visual cortex of monkeys and cats play a fundamental role in motion sensitivity and most are, to some extent, selective for the direction of motion (Hubel & Wiesel, 1962; 1968). The investigations of Barlow and Levick (1965) in the rabbit retina led them to propose that the basic mechanism of direction selectivity could be the result of “summation’’ over “simple excitatory and inhibitory connections” (pp. 498-500). However, other studies have approached the problem under the general assumption that direction selectivity is inherently nonlinear: for example, that it might involve a multiplicative or divisive interaction between inputs (for general reviews of this topic, see Nakayama, 1985; Hildreth & Koch, 1987).

Direction selectivity can be produced through strictly linear addition and subtraction of inputs (see for example the initial linear stages of the quadrature models developed by Watson & Ahumada, 1985; Adelson & Bergen, 1985). We have tested the usefulness and validity of a linear model (the linear quadrature model) for describing the properties of motion-sensitive neurons recorded from the visual cortex of monkeys and cats (Hamilton et al., 1989). The spatiotemporal transfer function of simple cells was measured using sine-wave grating patterns of variable spatial and temporal frequency, drifting first in one direction of motion and then in the opposite direction. The results of these experiments, using drifting gratings, showed that both the amplitude and the phase satisfied several strong constraints implied by the linear quadrature model. Quantitative measurements of the spatiotemporal receptive fields of simple cells provide further experimental support for the linear mechanism (McLean & Palmer, 1989).

On the other hand, the responses of direction-selective simple cells to stationary flickering grating patterns do not appear to be entirely consistent with a linear mechanism. Reid et al. (1987) developed a model of motion sensitivity based upon linear summation of lateral geniculate nucleus (LGN) inputs. The model was used to predict direction selectivity from the responses to stationary flickering gratings. For the 19 cells tested, they concluded that “about half of the direction selectivity is due to mechanisms that sum in a linear fashion (p. 8742).” Hamilton (1987) developed and tested similar predictions based upon a linear model of direction selectivity. From a sample of 27 cells, he showed examples of some cells which did agree with
the predictions and some which did not and then concluded that both linear and nonlinear mechanisms contribute to direction selectivity.

In many respects, simple cells are quite linear (for recent reviews, see Shapley & Lennie, 1985; De Valois & De Valois, 1988; Skottun et al., 1991). However, the contrast-response functions of simple cells reveal nonlinear behavior. Specifically, as the contrast increases, the response generally increases rapidly with a power-function exponent greater than 1.0 (the average for cat simple cells being approximately 2.5), and then saturates; furthermore, when measured at different spatial frequencies, the contrast-response functions shift vertically (on log-log coordinates) indicating a “contrast-set gain mechanism” (Albrecht & Hamilton, 1982). Although these nonlinear behaviors may not be inherent to the mechanism responsible for establishing direction selectivity, they may well influence the final output of a direction-selective mechanism, even if the direction-selective mechanism is based solely upon the simple principle of linear summation. Thus, it would seem reasonable to incorporate these contrast-response nonlinearities into any attempt to understand the responses of simple cells.

In the present study, we measured the responses of simple cells to stationary gratings and drifting gratings to quantitatively evaluate the linear and nonlinear components of direction selectivity. We also performed independent measurements of each cell’s contrast-response function. We developed and tested the predictions of the linear model and found, in agreement with Reid et al. (1987) and Hamilton (1987), that linear summation alone cannot account for the responses to stationary flickering gratings. However, if the nonlinearities revealed in the contrast-response function are combined with a linear spatiotemporal receptive field, then the predicted responses are consistent with the measured responses, for both drifting and counterphase stimuli.

Methods

The procedures for electrophysiological recording, stimulus display, and measurement of neural responses using linear systems analysis have been described in detail elsewhere (Albrecht & Hamilton, 1982; Albrecht et al., 1984; Hamilton et al., 1989).

Briefly, several days prior to an actual experiment, under deep barbiturate anesthesia, a preformed rigid plastic pedestal containing a recording chamber was attached to the animal’s skull. On the day of the recording, the animal was initially anesthetized with 20 mg/kg of the short-acting barbiturate thiamylal sodium and then maintained throughout the experiment on 75% NO₂/25% O₂ along with 1 mg/kg/h of thiamylal sodium. The head was held rigid in stereotaxic coordinates (without ear and eye bars) using the plastic pedestal. The eyes were immobilized by continuous infusion of gallamine triethiodide (10 mg/kg/h) and the animal was artificially respirated through an endotracheal throat tube. Single neurons were recorded using glass-coated tungsten or platinum–iridium microelectrodes.

The stimuli in the present set of experiments were spatial sine-wave grating patterns (presented on a Conrac studio monitor at a frame rate of 100 Hz), either drifting at a fixed temporal frequency or flickering in a stationary position at a fixed temporal frequency. The different stimulus conditions were randomly interleaved. One presentation consisted of a block of ten contiguous temporal cycles. Each block was separated by a period of time equal to the block length; during these separations, the animal viewed mean luminance with no contrast. A minimum of four blocks were obtained, which resulted in 40 repeated presentations of each stimulus condition. Action potentials were collected into 1-ms time bins and the resulting average peristimulus–time histograms were Fourier analyzed. The response measure was the amplitude and phase of the first harmonic component.

Once a single neuron was isolated and classified as a simple cell, its optimal orientation was determined, and held constant throughout the experiment. The spatial-frequency tuning, temporal-frequency tuning, and contrast-response function were quantitatively measured. Direction selectivity was determined by measuring the responses to the optimal grating drifting in the preferred and nonpreferred directions. The direction-selectivity index was the ratio of the measured responses subtracted from one (1 − nonpreferred/preferred). Thus, for a cell which only responded to movement in one direction, the index would be 1.0; for a cell which responded equivalently to movement in both directions, the index would be 0.0. The responses as a function of contrast, \( R(C) \), were fitted using least-squares criteria to the following function: \( R(C) = R_{\text{max}}C^n/(C^n + C_{50}) \), where \( R_{\text{max}} \) is the maximum response rate, \( C_{50} \) the semi-saturation contrast (the contrast required to produce 50% of the cell’s maximum response), and \( n \) the power-function exponent. This saturating power function provides a good description of the contrast-response function of neurons in the visual cortex.

Following these preliminary experiments, the responses to stationary counterphase flickering gratings were then measured at 12 different spatial/phase positions in 30-deg steps. The contrast, \( (I_{\text{max}} - I_{\text{min}})/(I_{\text{max}} + I_{\text{min}}) \), of the counterphase pattern was modulated sinusoidally. The spatial and temporal frequencies of the gratings were held constant at the optimal values. The full set of measurements were performed at a contrast level near the midpoint of the dynamic response range for the specific cell and then, as time allowed, at other contrasts around the midpoint. For ease of viewing, the graphs (of the responses as a function of the position of the counterphase flickering grating) were normalized in the following three ways: (1) the amplitude and phase responses were shifted horizontally so that the amplitude functions peaked at 90 and 270 deg, (2) the phase functions were shifted vertically so that they passed through the origin, and (3) the amplitudes were expressed as a percentage of the maximum response.

Results

The initial goal of this study was to assess whether the motion-sensitive properties of simple cells could be accounted for on the basis of a linear mechanism for direction selectivity. The basic strategy was to measure the direction selectivity of each cell using drifting gratings and then assess the degree to which the linear model could predict the responses to stationary counterphase flickering gratings.

Counterphase gratings are composed of two gratings of equal contrast moving in opposite directions. Given a linear mechanism for direction selectivity, and the principle of additivity in linear systems, the response to a counterphase grating would be a weighted sum of the responses to the individual components present in the stimulus. Although the weight of each component in the stimulus is equal, the weight of each
component in the response will depend upon the degree of selectivity. For example, given a linear cell that is completely direction selective (i.e., no response in the nonpreferred direction), the weight given to the component drifting in the nonpreferred direction would be zero, and the weight given to the component drifting in the preferred direction would be one. As the direction selectivity changes from cell to cell, the weight given to each component in the counterphase stimulus would vary accordingly.

Linear summation

This section describes the expected responses to counterphase gratings presented in different spatial positions, given a strictly linear model, where the direction selectivity is simply due to linear summation of inputs (see eqns. (A1–A5) in the Appendix). These expectations are then compared to the measured responses of 41 simple cells. (Note that adding halfwave rectification does not affect the predictions described in this section.)

The solid lines in Fig. 1A plot the predicted amplitude and phase of response for a linear nondirection selective cell, as a function of the position of a counterphase flickering grating. From the work of Enroth-Cugell and Robson (1966; see also, Hochstein & Shapley, 1976; Movshon et al., 1978; De Valois et al., 1982), we know that the amplitude of the response should be a sinusoidal function of spatial position with two “null phase positions,” where the response is zero. (We have adopted the convention of plotting all amplitudes as positive, with the phases ranging from 0–360 deg.) The temporal phase of the response should consist of two values separated by 180 deg: 90 over half of the range of spatial position (0–180 deg) and 270 over the other half (180–360 deg). The filled circles in Fig. 1A plot the measured amplitude and phase responses for a representative nondirection-selective simple cell with a direction selectivity index of 0.08. As can be seen, the expectations from a linear model provide an adequate fit for this cell. The measured amplitude varies as a function of position in a fashion which is similar to the absolute value of a sine function, with two positions where the response is near zero; the phase responses cluster around the two expected values, 90 and 270.

The solid lines in Fig. 1B plot the predicted amplitude and phase for a totally direction-selective cell. Given strict linearity,
the amplitude would not change as a function of position, and the phase would change continuously, with a slope of one. As noted above, these linear predictions are a simple consequence of the fact that while the stimulus is actually composed of two gratings drifting in opposite directions, a linear direction-selective cell will only be affected by the component moving in the cell's preferred direction. If the cell is only responding to one of the drifting components, changing the position of the counterphase grating would be equivalent to changing the starting position of that component. Hence, changing position would not affect the response amplitude (only the response phase).

The filled circles in Fig. 1B plot the measured amplitude and phase responses for a representative direction-selective simple cell with a direction index of 0.95. As can be seen, the expectations from a strictly linear model do not provide an adequate fit for this cell. Like the nondirection-selective cell, the measured amplitude varies as a function of position in a fashion which approximates the absolute value of a sine function, and there are two locations where the response is near zero. The phase responses do change continuously, as expected from the linear model; however, they cluster into two distinct groups separated by 180 deg, and the slope within each group is closer to 0.5 than 1.0.

In order to quantify these response patterns for the total population of cells, we indexed the size of the change in amplitude as a function of position by taking the maximum and minimum values of the response and computing the following ratio: \((\text{max} - \text{min})/(\text{max} + \text{min})\). The value of this amplitude index was then compared to the direction-selectivity index. In general, given a linear model, the amplitude index should be equal to one minus the direction-selectivity index.

Consider the expected changes in the amplitude index for cells with different degrees of direction selectivity. Given a completely nondirection-selective linear cell, there will be two positions of the grating where the inputs sum to a maximum value and two positions where the inputs sum to a minimum value (the "null phase positions"). Thus, for a cell with a direction index near 0.0, the amplitude changes would be large: the amplitude index would approach 1.0. Given a totally direction-selective linear cell, the response amplitude will be constant independent of position. Thus, for a cell with a direction index near 1.0, the amplitude changes would be small: the amplitude index would approach 0.0. Given a direction-selective linear cell between these two extremes, the amplitude will change with position; however, the differences between the maximum and minimum values will decrease as the direction selectivity increases. (Note that, strictly speaking, the term "null phase position" would be inappropriate for a linear cell with any degree of direction selectivity because the inputs do not perfectly cancel at any position of the counterphase grating; "minimum phase position" would be perhaps a more accurate term.)

Figure 2A shows the distribution of the amplitude index for the total sample of cells. The amplitude for most cells showed sizable changes as a function of the position of the counterphase grating (only four cells had an index less than 0.5). The mean of the distribution was 0.79. Figure 2B shows the scatter plot for the amplitude index vs. the direction index for all cells. The solid line, with a slope of -1.0, is the prediction of the linear model. As can be seen, nearly all of the cells in this sample showed larger changes in amplitude than was expected from a linear filter. The slope of the regression line for the data points was -0.21; the coefficient of correlation was -0.32. Even the most direction-selective cells showed sizable changes in amplitude as a function of position. For example, the direction-selective cell illustrated in Fig. 1B had an amplitude index of 0.76, yet its direction-selectivity index was 0.95.

The slope of the phase function was measured for the total population of cells and then compared to the direction-selectivity index. (The slope was determined by fitting separate straight lines to the data points which fell within the x axis ranges of 20–160 deg and 200–340 deg.) As noted above, for a linear nondirection-selective cell, the phases should cluster into two distinct groups separated by 180 deg and the slope within each group should be
0.0. For a totally direction-selective cell, the phases should change continuously from 0-360 with a slope of 1.0. In general, given a strictly linear filter, the slope of the phase function would be equal to the direction-selectivity index.

Figure 3A shows the distribution of the slope for the total sample of cells. For 70% of the population, the slope was less than 0.5; only two cells had a slope greater than 0.7. The average slope for the distribution was 0.35. Figure 3B shows the scatter plot for the slope vs. the direction selectivity. As can be seen, the majority of cells fall below the value expected from a strictly linear mechanism (indicated by the solid line with a slope of 1.0). The slope of the regression line for the data points was 0.52; the coefficient of correlation was 0.70. Even the most direction-selective cells did not have a slope of 1.0; for example, the slope of the direction-selective cell shown in Fig. 2B was 0.61.

**Linear summation, rectification, and response exponent**

The results presented above clearly indicate that simple cells cannot be described as strictly linear filters. One possible explanation is that the direction-selective mechanism itself is nonlinear. Another explanation, one that we consider in the Discussion, is a phase-specific direction-selective inhibition: a nonlinear silent inhibition from the nonpreferred component. However, these types of explanations require postulating relatively complicated and undocumented (although see Dean et al., 1980) nonlinear processes. A more parsimonious approach is to first consider other well-known nonlinearities.

One obvious nonlinearity of simple cells is seen in the responses to sine-wave stimuli: because simple cells have little or no maintained discharge, and they cannot produce negative responses, the response waveform is “halfwave rectified.” Any realistic model of simple cells should include this nonlinearity and thus in the sections which follow, the models incorporate halfwave rectification. However, halfwave rectification following linear summation of inputs cannot account for the results presented in the section above because the responses to both drifting and counterphase gratings are merely diminished by a fixed amount (the amplitude of the fundamental harmonic is diminished by 50%).

Another prominent nonlinear property of simple cells is the expansive power-function growth in response to contrast, before the onset of compression and saturation (Albrecht & Hamilton, 1982; Sclar et al., 1990). The power-function exponent varies from cell to cell; it is generally greater than 1.0; the average value is 2.5. If the exponent reflects an expansive nonlinearity occurring after linear summation, it could potentially explain the discrepancy between the responses to drifting vs. counterphase gratings. In this section, we develop and test the hypothesis that simple cell direction selectivity is based upon linear summation of inputs and that this linear summation is then followed by halfwave rectification and an expansive nonlinearity—the exponent model (cf. Heeger, in press, for a similar proposal). The exponent model is presented formally in the Appendix (see eqns. (A6–A11)).

Consider the potential effects of an expansive nonlinearity on the responses to counterphase flickering gratings presented in different positions. Recall that given a strictly linear filter, the magnitude of variation in the counterphase amplitudes (as a function of position) is inversely related to the degree of direction selectivity (the amplitude index is 1.0 minus the direction index). If the output of a linear direction-selective mechanism passes through an expansive nonlinearity, the variation in the counterphase amplitudes would be enhanced—the stronger responses at the peaks (90 and 270 deg) would be disproportionately increased relative to the weaker responses at the troughs (0 and 180 deg). Similarly, the direction selectivity of the cell would be enhanced; the stronger responses in the preferred direction would be disproportionately increased relative to the weaker responses in the nonpreferred direction. The phases would be unaffected by the expansive nonlinearity and thus the slope of the phase functions would directly reflect the linear direction-selective mechanism.

To illustrate the basic effect of an expansive nonlinearity on
Fig. 4. The effect of applying an expansive response exponent after linear summation is illustrated for a highly direction-selective cell (direction index 0.95). The solid line in (A) shows the amplitudes expected from a strictly linear direction-selective cell with a direction index of 0.72. A cell with this degree of direction selectivity shows some variation in amplitude (amplitude index 0.28) and the phases cluster into two groups with a slope equal to the direction index. The dashed lines in (A) show the effect of applying a power-function exponent of 4.2 (taken from the contrast-response function for this cell) to the expected responses. As can be seen, the exponent considerably enhances the variation in the amplitude and increases the peakedness. Furthermore, the exponent enhances the expected direction selectivity from 0.72-0.99. The solid line in (B) shows the expected phases before and after applying the exponent: the exponent has no effect on the phases. (Stimulating contrast/peak response = 0.1/65.9.)

the responses to counterphase flickering stimuli, Fig. 4 plots the measured amplitudes and phases of a direction-selective cell (direction index 0.95, amplitude index 0.87, and slope index of 0.65). The solid curve in Fig. 4A shows the predicted amplitudes of a strictly linear model with a direction index of 0.72; the dashed curve shows what happens to the predicted amplitudes if they are passed through an exponentiation stage (the exponent for this particular cell, taken from its contrast-response function, was 4.18). Note that the expansive nonlinearity results in a nonsinusoidal relationship between response amplitude and position (an exponentiated sinusoidal relationship); specifically, the exponent increases the difference between the maximum and minimum values, narrows the width of each half-cycle, and increases the peakedness. For this cell, the exponent enhanced the predicted variation in the amplitude from 0.28-0.83, which more closely matches the measured variation of 0.87. The exponent also enhanced the predicted direction selectivity from 0.72-0.99, which more closely matches the measured direction selectivity of 0.95. Because the exponent has no effect on the phases, the solid line in Fig. 4B shows the predicted phases with and without the exponent; the predicted slope, 0.72, is similar to the measured slope, 0.65.

In the exponent model, the predictions for normalized counterphase responses are completely determined by two parameters: (1) the direction selectivity of the linear mechanism (which determines the weight given to each of the drifting components) and (2) the nonlinear exponent (which determines the degree of response expansion). Both of these parameters can be estimated from the responses to drifting gratings. Specifically, by measuring the exponent $n$ of the contrast-response function and by measuring the ratio of the responses in the two directions, it is possible to make a parameter-free prediction of the normalized counterphase responses as a function of position for both amplitude and phase. (The direction selectivity of the linear mechanism is given by the $n$th root of the ratio of the responses in the preferred and nonpreferred directions; see eqns. (A9) and (A10) in the Appendix.)

Figure 5 replots the measured responses to counterphase gratings shown in Fig. 1; the nondirection-selective cell is shown in Fig. 5A and the direction-selective cell is shown in Fig. 5B. The expansive exponent (taken from the contrast response function) was 1.86 for the nondirection-selective cell and 3.22 for the direction-selective cell; the measured direction-selectivity indices were 0.08 and 0.95. The smooth curves through the data points represent the parameter-free predictions expected from these measured exponents and direction selectivities. As can be seen, the fit to both the amplitude and the phase data is quite good.

Figure 6 shows the results of this analysis for nine typical cells. For each cell, the exponent and the direction selectivity of the linear mechanism were predetermined as described above and thus there were no free parameters. The measured direction selectivity $d$, along with the exponent $n$ and the direction selectivity of the linear mechanism $D$, are shown in the lower right corner of each plot. These nine cells were chosen to illustrate a range of direction selectivities and a range of contrast-response exponents. From left to right, the measured direction selectivity decreases; this is reflected as a reduction in the slope of the phase functions. From top to bottom the exponent decreases; this is reflected as a reduction in the amplitude variation and, as expected, the effect of the exponent is more evident for the cells with a greater degree of direction selectivity. The fit to all 41 cells was quite good: for over half of the cells the predictions accounted for more than 90% of the variance (the average value was 83%).

An alternate approach to evaluating this model (similar to the approach taken by Reid et al., 1987) is to compare the direction selectivity measured from the responses to drifting gratings with the direction selectivity expected from the responses to stationary gratings. This analysis was also performed on the entire sample of cells. Specifically, for each cell, the amplitude and phase responses (as a function of the position of the counterphase flickering grating) were fitted using the linear-summation model, followed by rectification and the exponent of the contrast-response function. The exponent for each cell was fixed (from the independent set of measurements of the contrast-response function); however, the direction selectivity of the linear mechanism was free to vary. The optimized value of the linear direction selectivity, along with the measured exponent of the contrast-response function, were then used to predict the
A. NONDIRECTION SELECTIVE

The scatter plot shown in Fig. 7 illustrates the results of this analysis for all 41 cells; the direction selectivity predicted from the responses to stationary flickering gratings is plotted as a function of the direction selectivity measured from the responses to drifting gratings. As can be seen, the predicted and measured values were, on average, quite similar. The straight line, with a slope of 1.0, illustrates perfect correspondence. The slope of the regression line for the data points was 0.86; the coefficient of correlation was 0.89.

Contrast gain, linear summation, rectification, and response exponent

The model described above (linear summation followed by an expansive exponent) provides a good description of the responses to counterphase gratings presented at a fixed contrast (within the dynamic range of a cell). However, it does not take into account response compression and saturation. A more accurate model will have to include saturating nonlinearities.

The simplest hypothesis is that the expansive behavior at lower contrasts and the compressive behavior at higher contrasts are due to the same response nonlinearity—the response nonlinearity model. However, this hypothesis seems unlikely, based upon the results of previous studies. Albrecht and Hamilton (1982) measured the contrast-response function of striate cells at different spatial frequencies. As they argue, if the saturation were due to a final response nonlinearity, then saturation would occur at higher contrasts for nonoptimal spatial frequencies, the magnitude of the saturated response would be the same for all frequencies, and the spatial frequency tuning would change with contrast. Instead, they found that saturation occurred at approximately the same contrast for all spatial frequencies, the magnitude of the saturated response was different for each frequency, and the spatial-frequency tuning remained relatively invariant. They point out that this behavior would be produced if a saturating "contrast-set gain" mechanism preceded response generation. This earlier work, in conjunction with the results presented above, lead us to propose that there are two separate nonlinearities, a contrast-gain control nonlinearity that is responsible for saturation and a final response-exponent nonlinearity, that is responsible for expansion—the contrast-gain/exponent model.

These two alternative hypotheses (the response nonlinearity model and the contrast-gain/exponent model) can be discriminated by measuring contrast-response functions for counterphase gratings presented at different spatial positions. On the one hand, if the saturation and expansion are part of the same final response nonlinearity, then the contrast-response functions...
Fig. 6. Measured responses of nine different cells as a function of the position of a counterphase grating pattern. The smooth curves through the data points are the predictions from the model which assumes a linear mechanism for direction selectivity followed by a nonlinear response exponent (the exponent model). The exponent for each cell was determined from an independent measurement of the contrast-response function. As can be seen, the linear-summation model for direction selectivity, followed by the contrast-response exponent, provides a good fit to the measured responses for these nine cells. The nine cells were organized in this figure such that from left to right direction selectivity decreases and from top to bottom the exponent decreases. As would be expected, the slopes of the phases tend to decrease from left to right, and the amplitude change tends to decrease from top to bottom. In the lower right corner of each graph, the following three values are listed: \( n \) = measured exponent of the contrast response function, \( d \) = measured direction selectivity, and \( D \) = the direction selectivity of the linear mechanism. The direction selectivity of the linear mechanism \((D)\) was determined by removing the effect of the exponent on the measured direction selectivity \((d)\). (Stimulating contrast/peak firing rate for each cell: from left to right, top row, 0.1/65.9, 0.15/77.5, 0.05/57.0, middle row, 0.15/49.5, 0.04/47.5, 0.13/74.8, and bottom row, 0.2/24.5, 0.12/30.4, 0.06/27.3.)
measured at different spatial positions would reach saturation at different contrasts and the final saturated response levels would be identical (for both optimal and nonoptimal positions). On the other hand, if the contrast-gain/exponent model is correct, then the contrast-response functions would reach saturation at the same contrasts and the final saturated response levels would vary with the spatial position of the counterphase grating.

Figure 8 plots the responses of a typical cell to counterphase gratings presented at different positions and contrasts. In Fig. 8A, response is plotted as a function of position at five different contrast levels (0.02, 0.04, 0.08, 0.16, and 0.32). In Fig. 8B, response is plotted as a function of contrast at four different phase positions (52, 82, 142, and 172). The smooth curves through the data points are what would be expected if both the saturation and the expansive exponent were due to a final response nonlinearity (in this case, a saturating power function). As can be seen, the measured responses do not follow the predicted trend; instead, the responses tend to saturate at the same contrast and the magnitudes of the saturated responses are different for each position (responses at nonoptimal positions remain nonoptimal, independent of contrast).

To quantitatively test the hypothesis that there are separate contrast-gain and expansive response nonlinearities, we evaluated a specific version of the contrast-gain/exponent model. This version consisted of a multiplicative contrast-gain mechanism, a linear-summation mechanism, halfwave rectification, and a final expansive exponent. The rectification and expansive exponent occur after linear summation and the contrast-gain control occurs prior to the expansive exponent (the model is formally defined in the Appendix, see eqns. (A12-A16)). Figures 8C and 8D plot the same data points shown in Figs. 8A and 8B; the smooth curves are the best fit of the contrast-gain/exponent model. As can be seen, the predictions are quite accurate. The model correctly predicts that saturation occurs at a fixed contrast, that the magnitude of the saturated response varies with position, that the shape of the curve which describes amplitude as a function of position remains invariant with contrast (i.e. the curves at different contrasts only differ by a scale factor), and that the responses at nonoptimal positions remain nonoptimal.

The predictions are determined by four parameters: the direction selectivity of the linear mechanism (D), the expansive exponent of the final nonlinearity n, and two associated with the contrast-gain mechanism—a semi-saturation constant ($C_{50}$) and an exponent (m). The optimized values for these parameters were $D = 0.58$, $n = 2.47$, $C_{50} = 0.025$, and $m = 1.63$. (When the contrast-response functions were fitted using a saturating power function, the exponent was found to be 2.19, slightly less than 2.47, because, in the contrast-gain/exponent model, the slope of the contrast-response function is determined by both the contrast-set gain nonlinearity and final response nonlinearity.) If contrast is held constant, the two parameters of the contrast-gain mechanism are fixed and the model becomes equivalent to the simpler model described in the section above (the exponent model). Thus, the predictions of the contrast-gain/exponent model for all of the data measured at a fixed contrast (e.g., Figs. 4-7) are identical to the predictions of a model consisting of linear summation followed by rectification and an expansive exponent.

The contrast-gain/exponent model is consistent with other known characteristics of simple cells. The model correctly predicts the contrast-response functions measured at different spatial frequencies and orientations. Furthermore, it predicts the fact that spatial-frequency selectivity, orientation selectivity, direction selectivity, and degree of ocular dominance are relatively invariant with contrast. It can also predict the phase and amplitude responses, measured as a function of spatial and temporal frequency.

Discussion

A linear mechanism for direction selectivity

The first goal of this investigation was to test the hypothesis that the direction selectivity of simple cells is based upon linear summation. To do this, we presented stationary flickering gratings at different positions and contrasts. The linear model predicts that the responses to a counterphase grating should be the sum of the responses to the two drifting components. For example, the responses of a highly direction-selective cell to a counterphase grating should be no different than the responses to a drifting grating moving in the preferred direction. The results clearly showed that a model of simple cells which is strictly linear (i.e. not only the direction-selective mechanism is linear, but all other aspects as well) does not produce accurate predictions. However, when the nonlinearities evident in the contrast-response function are taken into account (contrast-gain control and expansive response exponent), a linear direction-selective mechanism does produce accurate predictions.
Contrast-gain/exponent model

We are led to suggest that the responses of simple cells might be usefully modeled, as illustrated in Fig. 9, with four basic stages (cf. Heeger, in press, for a similar proposal). The fundamental stimulus selectivities (for direction, orientation, spatial frequency, etc.) are established through a linear spatiotemporal filter (i.e., linear summation over space and time). Halfwave rectification and an expansive exponent are applied after linear summation; contrast-gain control is applied before the expansive exponent. This model is an extension of the strictly linear model presented in Hamilton et al. (1989) and is consistent with the results presented there.

At this point in time, we cannot be certain about the specific physiological and anatomical mechanisms which constitute these four stages. The contrast-set gain control observed in cortical cells may be the result of mechanisms in the retina, the LGN, the cortex, or all three. Some of this contrast-gain control is likely to be occurring in the retina (for a review see, Shapley & Enroth-Cugell, 1984); some may be feedforward gain control from LGN to cortical cells; and some may be due to a cooperative network in the cortex (see Bonds, 1990, for a
The contrast-gain/exponent model for simple cells. The model consists of four components. The linear component (a linear spatiotemporal filter) performs summation over space and time; it establishes the fundamental stimulus selectivities (such as direction, orientation, spatial frequency). The contrast-gain control component, a saturating multiplicative mechanism, scales the responses as a function of contrast; it maintains the stimulus selectivities independent of contrast. The rectification component places a threshold at zero response; it conserves metabolic energy. The expansive response exponent enhances stimulus selectivity. The wide arrows indicate global inputs dependent upon responses pooled across multiple cells. The thin arrows indicate local inputs dependent upon the responses within a single cell (or sequence of cells). The contrast-gain control is drawn below the other stages with early mechanisms, each selective to the opposite directions of motion, are halfwave rectified, differentiated, and then halfwave rectified again. Because of the rectification, one of the direction-selective mechanisms is silent in the sense that it produces no overt response. This model predicts that the responses to drifting and counterphase gratings are halfwave-rectified sine waves which modulate at the frequency of the stimulus. In response to counterphase gratings presented in different positions, the model predicts responses somewhat similar to the responses illustrated in Fig. 6. However, the exponent model and the contrast-gain/exponent model fit the data more accurately and they are more parsimonious (in that they are based upon simple linear summation coupled with the well-known nonlinearities of the contrast-response function).

Contrast-set gain control maintains stimulus selectivities

Cortical neurons have a limited dynamic response range. As stimulus strength increases, response magnitude generally increases rapidly as a power function with an exponent greater than one and then saturates. While the expansive relationship between stimulus and response can enhance stimulus selectivity (see discussion below), the saturation introduces a problem which could potentially diminish stimulus selectivity.

If the response saturation found in simple cells were determined solely by the overall magnitude of the response, it would cause nonpreferred stimuli (presented at high contrasts) to produce responses as large as preferred stimuli (see Robson 1975, for a discussion of this issue). However, previous work has shown that for most simple cells, spatial-frequency tuning, orientation tuning, eye preference, and direction selectivity are to a large extent unaffected by increases in contrast (Albrecht, 1978, Figs. 17-19; Albrecht & Hamilton, 1982, Figs. 7-10; Sclar & Freeman, 1982; Li & Creutzfeldt, 1984). Taken together with the results of the present experiments, these facts argue for contrast-set gain control which attenuates responses dependent upon the overall level of contrast.

In order for a multiplicative gain control to produce response saturation as a function of steady-state contrast, the gain factor must decrease as contrast increases so as to exactly cancel the increased stimulation. Because this saturation is achieved within a few temporal cycles, the gain control would have to occur quite rapidly (within a few hundred milliseconds) and thus it should perhaps be distinguished from the rather slow type of cortical cell adaptation which occurs following prolonged exposure to high contrasts (see, for example, Albrecht et al., 1984; Ohzawa et al., 1985). This rapid contrast-gain control is perhaps more similar to the “global inhibition” documented by Bonds (1990), the “rapid suppression” documented by Nelson (1991), and the “contrast normalization” proposed by Heeger (in press).
The net result of the contrast-set gain control mechanism is that the fundamental stimulus selectivities of a particular cell are maintained independent of the overall level of contrast—in spite of the limited response range and the steep slopes of the contrast-response function.

Another consequence of contrast-set gain control is that, once above the saturation contrast (which is often quite low), the stimulus–response relationship is highly constrained: each stimulus can only evoke one particular response level and thus, in comparison to a linear system, the response rate is more uniquely tied to the stimulus. For example, the maximum response level signals a unique stimulus (the stimulus that is simultaneously optimized on all relevant dimensions). Other response levels signal particular sets of nonoptimal stimuli which are determined by the tuning function of the cell along the relevant dimensions. Consider a polar plot of the two dimensions of spatial frequency and orientation. Each response level defines a concentric ring obtained by slicing the tuning function at that response level; only those stimuli on this ring could have generated that particular response level. The size of the ring decreases as the response level increases, collapsing to a point at the maximum response, thus illustrating that the higher the response rate the more unique the relationship between stimulus and response.

Rectification conserves energy

Simple cells have little or no maintained discharge and, when stimulated with a sine wave, only half of the modulation appears in the response (see, for example, Movshon & Tolhurst, 1975; Albrecht & De Valois, 1981). Thus, the responses appear to be halfwave rectified. If simple cells did have maintained discharges, at half their maximum response rates, then both halves of the modulation could potentially be represented. Because they do not, two cells (i.e. “reversed-polarity pairs”) are required for transmission of the information—one for the top half of the signal, and one for the bottom half. Although doubling the number of neurons might seem inefficient, it greatly reduces the number of action potentials produced, the amount of synaptic transmitter substance released, and hence metabolic effort.

To compare the efficiency of these two alternatives (one cell on at half-max vs. a pair of rectifying cells), consider the metabolic effort associated with encoding the luminance variations of sine-wave stimuli. For a cell to maintain a continuous response at half the maximum rate, the metabolic activity would obviously be quite high. Furthermore, since the average response of the cell would not vary with stimulus contrast, the metabolic rate would remain high whether the cell was responding to an actual visual contrast or simply waiting for a potential visual contrast.

This would not be true for pairs of rectifying cells. When the contrast is zero, the average response rate will be zero—a clear savings. However, even if the contrast of the stimulus is high and the two rectifying cells are responding at their maximum levels, the total response generated by both cells would be approximately half that of the one cell. Furthermore, whenever the responses of the two cells are below saturation, then even smaller responses will be generated by the rectifying cells.

In general, the metabolic savings factor $s$ for sine-wave stimuli is inversely related to the fraction $f$ of the saturated response evoked by an optimal sine wave ($s = \pi/4f$). Given that simple cells are so selective (for position, orientation, spatial frequency, direction of motion, etc.), their preferred stimulus is not always present and thus most cells will be far below the maximum saturated response and the savings in metabolic effort will be enormous. Suppose that the visual cortex as a whole is responding on average at 1% of its maximum level (the actual number is undoubtedly much smaller), then using rectified pairs would reduce the metabolic activity by a factor of 78.

Exponent enhances stimulus selectivities

The responses of cortical cells in both cat and monkey increase rapidly as a function of the strength of stimulation—the response increases with a power-function exponent greater than one. The value of the exponent in this relationship varies from cell to cell; the mean for cat simple cells is 2.5 and thus the slopes are very steep (in log-log coordinates). The steep slopes of the contrast-response function could potentially serve to enhance any and all previously established stimulus selectivities. By itself, an exponentiation stage (with an exponent greater than one) could not produce spatial-frequency tuning or orientation tuning; nor could it produce direction selectivity or ocular dominance. Furthermore, it may or may not enhance discriminability, depending upon where noise is introduced (see Geisler et al., 1991, for a detailed analysis of discrimination in single neurons). However, any stimulus preferences inherent in the input to the exponentiation stage will be greatly magnified.

Consider what happens to the responses to two separate stimuli, a gratings drifting to the left and a gratings drifting to the right, as these responses pass through the exponentiation stage of a particular direction-selective cell with an exponent of, say, 2.5. For one of these stimuli, the grating drifting to the right, linear summation of inputs results in a net sum of, say, 2.0; for the other stimulus, a grating drifting to the left, linear summation of inputs results in a net sum of, say, 6.0. When these two linear sums are passed through the exponentiation stage their relative magnitude is dramatically altered: the grating drifting to the right produces a response of 5.6 and the grating drifting to the left produces a response of 88.2. Thus, before the exponentiation stage, the preferred stimulus produced a net linear sum which was only three times greater than the nonpreferred stimulus, whereas after the exponentiation stage, the preferred stimulus produced a response which was nearly 16 times greater than the nonpreferred stimulus.

The enhancement of stimulus selectivity introduced by an expansive exponent eases the structural requirements for establishing highly selective receptive fields. Given a strictly linear relationship between input strength and response magnitude, the structural requirements are more demanding. For example, consider what would be required to produce a cell with a direction selectivity index of 0.95—a very direction-selective cell. If the contrast–response relationship was strictly linear, then the spatiotemporal receptive field would have to be very oriented in space–time. Although not impossible, this linear receptive field is actually quite difficult to realize physically (see Watson & Ahumada, 1985; Adelson & Bergen, 1985). On the other hand, given an exponentiation stage, the demands on the spatiotemporal receptive field are considerably less, and much easier to realize. This point is clearly illustrated in Fig. 10 where direction selectivity after expansion is plotted as a function of direction selectivity before expansion, given different exponents.

Similar logic applies to other stimulus selectivities. In the
case of orientation tuning, a greater degree of selectivity could be attained with a shorter receptive field. In the case of spatial-frequency tuning, a greater degree of selectivity could be attained with fewer numbers of excitatory and inhibitory regions. If the rather broad spatial-frequency tuning of the center/surround inputs to the cortex are passed through a cascade of neural elements (each with exponents greater than one), the tuning at the end of the series will be very narrow.

Figure 11 illustrates the potential effect of an expansive exponent on spatial-frequency selectivity. The circles connected by solid lines plot the measured responses of a center/surround receptive field, recorded from the LGN of a cat, as a function of spatial frequency. Although this cell does show a certain degree of spatial-frequency tuning, it is not very selective: the bandwidth at half-height is approximately 2.5 octaves. The circles connected by dashed lines illustrate what would happen to the responses if they were passed through a cell with a contrast-response exponent of 2.5; the circles connected by dotted lines illustrate the consequences of cascading through a second cell with an exponent of 2.5 (or equivalently, instead of two cells in series, one cell with an exponent of 5.0). As can be seen, the mere addition of an exponentiation stage produces considerable narrowing of the spatial-frequency tuning. An exponent of 2.5 reduced the bandwidth to approximately 1.5 octaves, which is close to the average value for cat simple cells. A second exponent of 2.5 reduced the bandwidth still further to approximately 1.0 octaves.

The longstanding discrepancies between the degree of selectivity and the exact shape of the receptive field (for recent review of this issue, see Tadmor & Tolhurst, 1989) could perhaps be accounted for by the exponent of the contrast-response function. The basic discrepancy is that the number of parallel excitatory and inhibitory regions in the measured receptive field does not always correspond to the number expected from the measured spatial frequency response function. An expansive exponent could help explain this discrepancy in two ways. First, it would narrow the bandwidth of the measured spatial-frequency response function (as described above). Second, it would diminish the responses of the weaker peripheral regions of the receptive field and enhance the responses of the dominant central regions. Both of these effects would contribute to the discrepancy between measurements in the space domain and spatial-frequency domain.

Barlow (1972) and others (for a review, see Maunsell & Newsome, 1987) have suggested that as sensory information flows from the peripheral receptors through the thalamus and then on through the various regions of the cerebral cortex, the stimulus selectivity of the sensory neurons tends to increase: the receptors respond to a very broad range of stimuli, whereas the higher level neurons respond to a very narrow range of stimuli. As a consequence, the responses of the higher level neurons could potentially signal the presence of complex stimulus features. The exponent of the contrast-response function can help achieve this type of stimulus selectivity: if each stage of processing along a sensory pathway simply exponentiates the final output (using exponents greater than one), then the stimulus selectivity inherent in the summation of inputs would be enhanced. The recent report from Sclar et al. (1990) is consistent with this basic notion. They demonstrated that the average value of the exponent of the contrast-response function for neu-
rons in the visual pathway of monkeys increases from 1.6 in the parvocellular layer of the LGN to 2.4 in the striate cortex to 3.0 in the middle temporal visual area. Although one benefit of these steep contrast-response functions might be enhanced contrast sensitivity (as they point out), another benefit might be generalized enhancement of stimulus selectivity for higher level neurons.

Acknowledgments

We thank the University of Texas. We also thank Larry Stern for assisting with data analysis and Gail Geisler for drafting the figures. This research was supported in part by NIH Grant EY02688 to W.S. Geisler.

References


Appendix

This Appendix describes the quantitative models that were considered in trying to account for the responses of simple cells to drifting and counterphase gratings.

Linear model

To begin, let \( L(x,t) \) represent the spatiotemporal stimulus, where \( x \) is space and \( t \) is time. The linear mechanism integrates/filters the input stimulus by a spatiotemporal impulse-response function, \( h(x,t) \). Without loss of generality, we can assume that the receptive field is positioned at zero. Thus, the predicted response over time, \( R(t) \), is

\[
R(t) = L(x,t) * h(x,t) \big|_{x=0} \tag{A1}
\]

where ** represents the operation of two-dimensional convolution. The predicted responses to gratings drifting in opposite directions are

\[
R_1(t) = |H(u,-w)| C \cos(-2\pi wt + \phi(u,-w)) \tag{A2}
\]
where $C$ is the contrast of the gratings, $|H(u, -w)|$ and $|H(u, w)|$ are the values of the amplitude spectrum of $h(x, t)$ at the spatial ($u$) and temporal ($w$) frequencies of the drifting gratings, and $\phi(u, -w)$ and $\phi(u, w)$ are the values of the phase spectrum at these frequencies. Note that changing the sign of $w$ reverses direction of motion; without loss of generality, we can assume that $-w$ is associated with motion in the nonpreferred direction. The response to the mean luminance does not appear in the above equations because we are only considering the amplitude and phase of the fundamental.

Comparing eqns. (A2) and (A3) shows that the ratio of the response amplitudes in the two directions is $|H(u, -w)| / |H(u, w)|$. Thus, the ratio of the response amplitudes equals the ratio of the amplitude spectrum values of the linear summation mechanism. The direction-selectivity index used in the present paper is defined to be one minus the ratio of the response in the nonpreferred direction to that in the preferred direction. Therefore, the predicted direction-selectivity index $D$ is

$$D = 1 - |H(u, -w)| / |H(u, w)|.$$  (A4)

Because a counterphase grating can be written as the sum of two gratings drifting in opposite directions (scaled by 0.5 to preserve the mean luminance and contrast), the predicted response is given by

$$R_3(t) = 0.5 |H(u, -w)| C \cos(-2 \pi w t + \phi(u, -w)) + 0.5 |H(u, w)| C \cos(-2 \pi u x_0 - w t + \phi(u, w)),$$  (A5)

where $x_0$ is the position of the counterphase grating. Once normalized to the peak response and to a response phase of zero, the predicted responses (as a function of $x_0$) depend only on the direction-selectivity index $D$ of the linear mechanism and the exponent $n$ of the final response nonlinearity. These two parameters can be estimated from the responses to drifting gratings, making it possible to generate parameter-free predictions of the measured direction selectivity (see Figs. 5 and 6). Alternatively, $n$ can be estimated from the contrast-response function for drifting gratings and $D$ can be estimated from the responses to counterphase gratings in order to make parameter-free predictions of the measured direction selectivity (see Fig. 7).

Exponent model

This model is identical to the linear-summation model except that the output is rectified and then an exponent $n$ is applied. Thus, the predicted response is given by

$$R(t) = \max[L(x, t) \ast h(x, t), 0]^n.$$  (A6)

where taking the $\max$ with respect to 0 performs the halfwave rectification. The predicted responses to drifting gratings are

$$R_1(t) = \max[|H(u, -w)| C \cos(-2 \pi w t + \phi(u, -w)), 0]^n$$  (A7)

and

$$R_2(t) = \max[|H(u, w)| C \cos(2 \pi w t + \phi(u, w)), 0]^n.$$  (A8)

For this model, the ratio of the response amplitudes in the two directions is $|H(u, -w)|^n / |H(u, w)|^n$. Thus, the ratio of the response amplitudes equals the ratio of the amplitude spectrum values of the linear mechanism raised to the exponent of the response nonlinearity. Therefore, the direction selectivity index $d$ is

$$d = 1 - |H(u, -w)|^n / |H(u, w)|^n.$$  (A9)

Combining eqns. (A4) and (A9) shows that the relationship between the direction-selectivity index $D$ of the linear mechanism and the direction-selectivity index $d$ of the final response is given by

$$D = 1 - (1 - d)^{1/n}.$$  (A10)

This relationship was used to estimate $D$ from the measured responses to drifting gratings.

The predicted response to counterphase gratings is

$$R_3(t) = \max[0.5 |H(u, -w)| C \cos(-2 \pi u x_0 + w t) + \phi(u, -w), 0] + 0.5 |H(u, w)| C \cos(-2 \pi u x_0 - w t) + \phi(u, w), 0]^n.$$  (A11)

Once normalized to the peak response and to a response phase of zero, the predicted responses (at a fixed spatial and temporal frequency), as a function of $x_0$, depend only on the direction-selectivity index $D$ of the linear mechanism and the exponent $n$ of the final response nonlinearity. These two parameters can be estimated from the responses to drifting gratings, making it possible to generate parameter-free predictions for responses to counterphase gratings (see Figs. 5 and 6). Alternatively, $n$ can be estimated from the contrast-response function for drifting gratings and $D$ can be estimated from the responses to counterphase gratings in order to make parameter-free predictions of the measured direction selectivity (see Fig. 7).

Contrast-gain/exponent model

This model is identical to the above model except that contrast-gain control is applied before the response exponent. The contrast-gain control mechanism scales the input stimulus by a gain factor $g(\bar{C})$ that is a function of the average (rms) spatiotemporal contrast $\bar{C}$ over some region of space and time. Thus, the predicted response is given by

$$R(t) = \max[g(\bar{C}) L(x, t) \ast h(x, t), 0]^n.$$  (A12)

The contrast-gain factor is given by the following equation:

$$g(\bar{C}) = \bar{C}^{m-1}/(\bar{C}^m + \bar{C}_{50}^m)$$  (A13)

where $m$ is an exponent greater than zero, and $\bar{C}_{50}$ is the half-saturation constant.

There are several points to make about this model before considering the predictions for drifting and counterphase gratings. (1) Because the gain-control factor is determined by the average contrast, it acts like a constant in eqn. (A12) and can be factored out. Thus, it is apparent that the predictions of the model are the same whether the contrast-gain control is before the linear summation, after the linear summation, after the half-
wave rectification, or distributed across all three levels. (2) On
the other hand, the response exponent is dependent on the re-
response magnitude and hence occurs last. (3) There are several
ways that the average contrast might have been computed. For
simplicity, we used the rms calculation over one temporal and
spatial cycle, but the predictions for the present experiments are
not strongly dependent on the type of averaging.

The predicted responses to drifting gratings are

\[
R_1(t) = \max[|H(u,-w)|g(\tilde{C})C \cos(-2\pi wt + \phi(u,-w))],0]^n
\]

(A14)

and

\[
R_2(t) = \max[|H(u,w)|g(\tilde{C})C \cos(2\pi wt + \phi(u,w))],0]^n.
\]

(A15)

Combining eqns. (A14-A15) shows that once again the ratio of
the response amplitudes in the two directions is \( |H(u,-w)|^n / |H(u,w)|^n \), and thus eqn. (A10) still holds.

The predicted response to counterphase gratings is

\[
R_3(t) = \max[0.5|H(u,-w)|g(\tilde{C})C \cos(-2\pi (u\omega_0 + wt)
+ \phi(u,-w))
+ 0.5|H(u,w)|g(\tilde{C})C \cos(-2\pi (u\omega_0 - wt)
+ \phi(u,w))],0]^n.
\]

(A16)

Once normalized to the peak response and to a response phase
of zero, the predicted response to counterphase gratings as a
function of \( \omega_0 \), depend only on the direction-selectivity index \( D \)
of the linear mechanism and the final response exponent \( n \). The
normalized predictions are identical to those of the exponent
model at a fixed contrast, but differ if contrast is varied.