Neural Mechanisms for Processing Binocular Information
I. Simple Cells

AKIYUKI ANZAI, IZUMI OHZAWA, AND RALPH D. FREEMAN

Group in Vision Science, School of Optometry, University of California, Berkeley, California 94720-2020

Anzai, Akiyuki, Izumi Ohzawa, and Ralph D. Freeman. Neural mechanisms for processing binocular information. I. Simple cells. J. Neurophysiol. 82: 891–908, 1999. The visual system integrates information from the left and right eyes and constructs a visual world that is perceived as single and three dimensional. To understand neural mechanisms underlying this process, it is important to learn about how signals from the two eyes interact at the level of single neurons. Using a sophisticated receptive field (RF) mapping technique that employs binary m-sequences, we have determined the rules of binocular interactions exhibited by simple cells in the cat’s striate cortex in relation to the structure of their monocular RFs. We find that binocular interaction RFs of most simple cells are well described as the product of left and right eye RFs. Therefore the binocular interactions depend not only on binocular disparity but also on monocular stimulus position or phase. The binocular interaction RF is consistent with that predicted by a model of a linear binocular filter followed by a static nonlinearity. The static nonlinearity is shown to have a shape of a half-power function with an average exponent of ~2. Although the initial binocular convergence of signals is linear, the static nonlinearity makes binocular interaction multiplicative at the output of simple cells. This multiplicative binocular interaction is a key ingredient for the computation of interocular cross-correlation, an algorithm for solving the stereo correspondence problem. Therefore simple cells may perform initial computations necessary to solve this problem.

INTRODUCTION

Neural signals from the left and right eyes are segregated until they reach the striate cortex and converge onto single cells to form binocular neurons. Therefore it is believed that binocular neurons in the striate cortex perform initial computations for mediating binocular fusion and stereoscopic depth perception (e.g., Barlow et al. 1967; Pettigrew 1965; Pettigrew et al. 1968). To identify the neural computations carried out by the binocular neurons, it is essential to obtain rules of how signals from the two eyes are combined at the level of single neurons, i.e., the binocular interaction of signals.

Hubel and Wiesel (1959) were the first to describe binocular interactions exhibited by simple cells in the cat’s striate cortex. They observed that stimulating ON (or OFF) subregions of the left and right eye receptive fields (RFs) simultaneously results in response summation, whereas stimulating an ON subregion in one eye and an OFF subregion in the other eye cancels the response. This suggests that the binocular interaction of signals may be linear. They also reported that some cells respond only when stimulated binocularly (Hubel and Wiesel 1962), which is indicative of a nonlinear binocular interaction. However, this still could be attributed to a subthreshold summation that is linear (Ohzawa and Freeman 1986). Because they found that left and right eye RFs occupy corresponding positions on the two retinas and are strikingly similar in their organization, they thought that retinal images of objects either in front of or behind the point of visual fixation would not be effective for evoking responses from the cells (Hubel and Wiesel 1959, 1962). Therefore they concluded that binocular cells in the striate cortex are probably not involved in stereoscopic depth discrimination (Hubel and Wiesel 1959, 1962, 1970, 1973). Instead, it was thought that such cells may be related to mechanisms of binocular fixation (Hubel and Wiesel 1959).

Other studies also found that the binocular interaction of signals results in response facilitation, summation, or occlusion, but contrary to Hubel and Wiesel’s claim, these studies reported that a substantial number of cells are selective to binocular disparity (Barlow et al. 1967; Bishop et al. 1971; Blakemore 1969; Ferster 1981; Fischer and Kruger 1979; Kato et al. 1981; LeVay and Voigt 1988; Maske et al. 1986a,b; Pettigrew 1965; Pettigrew et al. 1968; von der Heydt 1978). For instance, Pettigrew et al. (1968) measured the tuning for binocular disparity of cells in the cat’s striate cortex using moving bright bars of various binocular disparities. They found that some cells are narrowly tuned to binocular disparity and that the optimal disparity and the width of the tuning vary from cell to cell. Others found similar results (Barlow et al. 1967; Blakemore 1969; Bishop et al. 1971; Ferster 1981; Fischer and Kruger 1979; Kato et al. 1981; LeVay and Voigt 1988; Maske et al. 1986a,b; von der Heydt 1978).

Cells selective to binocular disparity also are found in monkey striate cortex (Cumming and Parker 1997; Gonzalez et al. 1993; Poggio 1990; Poggio and Fischer 1977; Poggio and Talbot 1981; Poggio et al. 1985, 1988). A proportion of these cells are shown to respond to dynamic random-dot stereograms (Cumming and Parker 1997; Gonzalez et al. 1993; Poggio 1990; Poggio et al. 1985, 1988), which suggests that the stereo correspondence problem may be solved, at least partially, at the striate cortex (Gonzalez et al. 1993; Poggio et al. 1985; but see Cumming and Parker 1997). Indeed, some of the cells are sensitive to binocular image correlation (Gonzalez et al. 1993; Poggio et al. 1985, 1988).

Although these studies have established that responses of binocular cells are modulated depending on the binocular disparity of a stimulus, there are some problems that make interpretation of the results difficult. First of all, the use of moving bars confounds spatial and temporal factors. When the binocular disparity of the stimulus is changed, the timing at which left and right eye bars reach the corresponding positions of the...
retinae also is changed. In other words, a binocular disparity introduces an interocular temporal offset as well as a spatial offset. Therefore it is not clear whether binocular disparity tuning results from differential responses to binocular disparity, the temporal sequence of bar stimulation, or both.

Second, there are many pairs of monocular bar positions that yield the same binocular disparity. Therefore it is possible that cells respond differently to the same binocular disparity depending on the monocular positions of the bars. The previous studies ignored this possibility either by averaging responses over space using moving bars or by the use of extended stimuli such as dynamic random-dot stereograms (but see Ohzawa et al. 1990).

Another problem is that there is some evidence that suggests that binocular disparity tuning is stimulus dependent. Maske et al. (1986a) measured the tuning for binocular disparity of cells in the cat’s striate cortex using bright and dark bars. They found that tuning curves obtained with these stimuli are different for some cells. Ohzawa et al. (1990) measured binocular interaction profiles of cells in the cat’s striate cortex using not only bright and dark bars but also a combination of the two, i.e., a bright bar in one eye and a dark bar in the other eye. They found that the profiles depend on the stimulus (see also Cumming and Parker 1997). Therefore binocular disparity tuning measured with only bright or dark bars/dots, as in most of the previous studies, is incomplete.

There is also an important issue that most of the previous studies could not address (but see Ferster 1981): what are the neural mechanisms underlying binocular interactions that make these cells selective to binocular disparity? Ohzawa and Freeman (1986) measured the tuning for interocular phase disparity of simple cells in the cat’s striate cortex using drifting sinusoidal gratings. They found that most cells show a phase-specific binocular interaction that is consistent with the predictions of linear binocular summation. Therefore they concluded that the binocular interaction exhibited by simple cells is linear. This suggests that a simple linear mechanism is responsible for a cell’s selectivity to binocular disparity.

On the other hand, there is also evidence for nonlinear binocular interactions. Ferster (1981) measured the tuning of cells in areas 17 and 18 of cats for binocular disparity using moving bright bars. He compared the binocular disparity tuning with the profiles of left and right eye RFs and found that the binocular disparity tuning can be predicted by taking a cross-correlation between the left and right eye RF profiles. This indicates that the binocular interaction is multiplicative and suggests that the mechanism underlying binocular disparity selectivity is nonlinear. This result appears to be at odds with Ohzawa and Freeman’s result that binocular interaction is linear. A resolution of this apparent contradiction requires a more detailed analysis of binocular interaction and monocular RFs.

To avoid the problems of the previous studies and address the issue of neural mechanisms underlying binocular interaction, white noise analysis (e.g., Marmarelis and Marmarelis 1978) is conducted in this study. Spatiotemporal white noise generated according to binary m-sequences (Sutter 1987, 1992) is used to measure binocular interaction RFs and monocular RFs of simple cells in the cat’s striate cortex. The binocular interaction RF represents how signals from the left and right eyes are combined at each pair of monocular positions. It describes how a cell responds to stimuli of various binocular disparities and how that depends on monocular stimulus position. Therefore the question of whether binocular disparity tuning depends on the monocular position of a stimulus can be addressed. The noise stimulus covers the entire left and right eye RFs and is updated rapidly so that binocular disparity exists everywhere in the RFs all the time. This ensures that spatial and temporal parameters of the stimulus are not confounded. Moreover, the stimulus contains all binocular combinations of bright and dark bars (bright-bright, dark-dark, bright-dark, and dark-bright), so that the measurement is complete.

The use of white noise also allows one to examine the system structure of cells and estimate parameters for the components of the system (see Anzai 1997 for a review on this topic). It has been proposed that simple cells can be modeled as a system that has a structure of a linear filter followed by a static nonlinearity (e.g., Albrecht and Geisler 1991; Andrews and Pollen 1979; DeAngelis et al. 1993; Hamilton et al. 1989; Heeger 1992b; Jagadeesh et al. 1993, 1997; Mancini et al. 1990; Movshon et al. 1978; Ohzawa and Freeman 1986; Pollen et al. 1988; Tadmor and Tolhurst 1989; Tolhurst and Dean 1987, 1990), and that the static nonlinearity is a half-squaring function (e.g., Emerson et al. 1989; Heeger 1992b; Mancini et al. 1990). However, most of the studies that examined the linearity of simple cells conducted rather relaxed tests of linearity (see Anzai 1997 and Heeger 1992b for a review), and the linearity was not tested for each point of the RF in space and time. In addition, any deviation from a linear prediction often was attributed to a static nonlinearity without an appropriate analysis of the nonlinearity. White noise analysis offers an alternative method of identifying the system structure of cells (e.g., Billings and Fakhouri 1978; Chen 1995; Chen et al. 1986; Hunter and Korenberg 1986; Korenberg and Hunter 1986; Marmarelis and Marmarelis 1978). For example, if a cell has a system structure of a linear binocular filter followed by a static nonlinearity, its binocular interaction RF and monocular RFs are expected to show a certain relationship. Therefore by examining the relationship among the RFs, one can determine if the system structure of binocular simple cells is consistent with the model. A similar analysis has been applied to temporal interaction (Emerson et al. 1989; Mancini et al. 1990) and spatiotemporal interaction (Jacobson et al. 1993; Emerson 1997) for monocular responses of simple cells. Once the system structure is identified, one can estimate parameters for the system components (Emerson et al. 1989; Mancini et al. 1990). In particular, parameters for nonlinear components of the system (e.g., the shape of the static nonlinearity) are important because they represent the underlying nonlinear computations performed by the cell.

Here, by determining the system structure for binocular simple cells and describing the nature of nonlinearities in the system, neural mechanisms underlying binocular interactions are identified. Thus the issue of whether binocular interaction is linear or nonlinear is resolved. This analysis also provides important clues as to what kind of neural computations are performed by binocular simple cells. Possible roles of binocular simple cells in binocular fusion and stereopsis are considered.
METHODS

Surgical and histological procedures, apparatus, and recording procedures are identical to those described in the preceding paper (Anzai et al. 1999a). Binocular interaction RFs of simple cells are obtained, along with their monocular RFs, using dichoptic one-dimensional (1D) binary m-sequence noise (for details of the stimulus configuration, see Anzai et al. 1999a). The relationship between the binocular interaction RF and monocular RFs is analyzed for each cell to determine whether binocular simple cells behave in a way that is consistent with a model of a linear binocular filter followed by a static nonlinearity. Then for those cells that are consistent with the model, the shape of the static nonlinearity is estimated.

Construction of RF maps and their interpretation

Each spike train recorded as a response to binary m-sequence noise is cross-correlated with the stimulus sequence to obtain RF maps. The cross-correlation between the stimulus sequence in the left eye and a spike train yields a left eye RF (L). Substituting the stimulus sequence for the right eye into the cross-correlation yields a right eye RF (R). The cross-correlation among the stimulus sequences in the left and right eyes and the spike train yields a binocular interaction RF (B). The cross-correlations are computed by means of the fast m-transform (Sutter 1991), which is a very efficient algorithm for the computations. Operationally, these computations can be described as follows.

To obtain a monocular RF, first a spike train is cross-correlated with a binary m-sequence at each position of the stimulus elements. This yields a cross-correlogram that represents a temporal response profile (in steps of 5 ms) of the RF for each position (Fig. 1). Then a spatial response profile of the RF is constructed by taking a value from each correlogram at a correlation delay (\(\tau\)). The monocular RF represents the responses to bright bars minus the responses to dark bars and provides the best linear approximation, in a mean-squared error sense, to the stimulus-response relationship of the cell.

A binocular interaction RF is constructed as illustrated in Fig. 2. There is a region in space that is covered by both left and right eye stimuli, which is labeled as the binocular view field in the figure. Any point in the binocular view field can be specified by two stimulus bar locations—one in each eye. If this region is filled with bright dots when the corresponding left and right eye bars have the same polarity and with dark dots when the polarities are different, a two-dimensional (2D) noise pattern like that shown in the figure is obtained. This pattern changes every 40 ms according to the same m-sequence used to generate the dichoptic 1D noise stimulus. The sequence has a different time shift for each point in the binocular view field so that the synthesized pattern is uncorrelated in space and time for the purpose of RF mapping. Then one can compute a cross-correlation profile of the RF at \(\tau\).
between a spike train and the sequence of the synthesized pattern and obtain a 2D activity map in the same way that the monocular RF is obtained. The map is called a binocular interaction RF and represents the responses to stimuli of matched polarity in the two eyes minus the responses to stimuli of mismatched polarity in the two eyes. This map reflects only responses due to nonlinear binocular interaction; i.e., if the left and right eye signals are summed linearly without any further nonlinear processing, the map is uniformly zero. As illustrated in Fig. 2, right, the binocular interaction RF has axes of left eye bar position, $X_L$, and right eye bar position, $X_R$. The vertical axis, $D$, represents binocular disparity, and the frontoparallel axis, $X_F$, runs in the horizontal direction.

Identification of the system structure for binocular simple cells

White noise analysis allows one to determine the system structure of cells (see Anzai 1997 for a review). For a binocular simple cell that has a structure of linear binocular filter followed by a static nonlinearity, as depicted in Fig. 3, a relationship exists between the binocular interaction RF and monocular RFs. That is, such a cell satisfies the following condition (see Appendix for derivation)

$$B_{ij}(\tau) = \alpha \cdot L_i(\tau) \cdot R_j(\tau)$$

(1)

where $L$, $R$, and $B$ are the left eye, right eye, and binocular interaction RFs, respectively. The subscripts, $i$ and $j$, denote stimulus bar positions in the left and right eyes, respectively. The variable, $\tau$, indicates a cross-correlation delay, and $\alpha$ represents a constant. This equation states that if a binocular simple cell has the system structure illustrated in Fig. 3, then its binocular interaction RF is approximately proportional to the product of the left and right eye RFs (note that for a Gaussian input stimulus this relationship is exact, not an approximation). In other words, if one plots the binocular interaction RF, $B_{ij}(\tau)$, against the product of the left and right eye RFs, $L_i(\tau) \cdot R_j(\tau)$, then the data points should fall on a straight line through the origin with a slope $\alpha$. A linear regression analysis is conducted on such a plot for each cell to examine if the condition stated in Eq. 1 is met.

In this analysis, data points outside the cell’s monocular RFs are excluded; the values of $i$, $j$, and $\tau$ in Eq. 1 are restricted to be within the extent of the cell’s RFs. The extent of each monocular RF is defined by the smallest region in space and time outside of which the squared value of each data point is $<5\%$ of the squared value of the peak data point.

Estimating the shape of the static nonlinearity

For cells that are well described by a linear binocular filter followed by a static nonlinearity (i.e., those that satisfy Eq. 1), the shape of the static nonlinearity is estimated from its input-output relationship. The input to the static nonlinearity, i.e., the output of a linear binocular filter [denoted by $W(t)$ in Fig. 3], is estimated by convolving the monocular RFs ($L$ and $R$) with the noise stimuli ($S_L$ and $S_R$) used to obtain the RFs.\footnote{For a cell with a structure illustrated in Fig. 3, measured RFs, $L$ and $R$, are actually equivalent to impulse response functions of the left and right eye linear filters only up to some unknown scaling factor (see Eqs. A9 and A11 in the Appendix). Therefore the input to the static nonlinearity $W(t)$, when estimated using $L$ and $R$, is also scaled by the same factor. For this reason, $W$ will be presented as a normalized quantity. However, the shape of the static nonlinearity depends neither on the scaling factor nor on the normalization.}

The output of the static nonlinearity $Y(t)$ in Fig. 3 is the spike train recorded as a response to the noise stimuli. Both $W$ and $Y$ represent a time series. A value of $W$ indicates an input to the static nonlinearity summed over a period of 40 ms, which is the stimulus update period. Likewise, a value of $Y$ represents a total spike count for a 40-ms period. The input-output relationship of the static nonlinearity then is obtained by plotting $Y$ values against $W$ values for the entire record of the spike train ($\sim 20$ min long). Because spike generation is a stochastic process, the same input value of $W$ does not necessarily yield the same spike count $Y$. Therefore the axis for the

![Diagram of binocular simple cells](image.png)

**FIG. 3.** Model of binocular simple cells. Model consists of 2 major parts: a linear binocular filter (the dashed box) and a static nonlinearity (N). Binocular filter is subdivided into a left eye linear filter (L) and a right eye linear filter (R). These monocular filters receive inputs $[S_L(t)$ and $S_R(t)$] from multiple spatial locations in the left and right eyes. Outputs from the monocular filters are combined linearly to produce the output of the binocular filter, $W(t)$. Finally, the static nonlinearity converts $W(t)$ to $Y(t)$, the output response of the cell to the stimuli $S_L(t)$ and $S_R(t)$. The variables $S_L$, $S_R$, $W$, and $Y$ are all functions of time $t$.\footnote{A. ANZAI, I. OHZAWA, AND R. D. FREEMAN}
input $W$ is divided into bins (see the legend of Fig. 9 for details of binning) and a mean $Y$ value and a mean $W$ value of the data points are computed for each bin. A curve connecting the mean values for all bins defines the shape of the static nonlinearity. See Fig. 9 for an example.

As shown in RESULTS, the static nonlinearity turns out to be an expansive function. Such a function can be well described by a half-power function of the form

$$Y = \max \{ \beta(W + \theta), 0 \}^n$$

This equation is a combination of a threshold and an expansive (for $n > 1$) power function. The parameter $\theta$ represents a threshold value, and $\beta$ scales the input $W$. Max[] denotes the largest of the values inside the brackets and acts as a rectifier. The exponent $n$ determines the steepness of the expansive function. When $n$ is 1 and $\theta$ is 0, Eq. 2 reduces to a half-rectification. If $n$ equals 2, then it is a half-squaring function (Heeger 1992b). Taking the logarithm of Eq. 2 (for $\beta(W+\theta) > 0$) yields

$$\log Y = n \log \beta + n \log (W + \theta)$$

This equation indicates that the relationship between $\log Y$ and $\log W$ can be described by a straight line with a slope $n$ for $W \gg \theta$. It also predicts that $\log Y$ values at $W = \theta$ would deviate from the straight line (see Fig. 9B for an example). To estimate the exponent of the static nonlinearity, we plot the data on log-log coordinates and fit a straight line to three consecutive data points to find the maximum slope, which is taken as an estimate of the exponent.

RESULTS

Monocular RFs and binocular interaction RFs have been obtained for 85 binocular simple cells in 16 adult cats. Of these, 49 cells exhibited substantial binocular interactions and are analyzed here. The remaining 36 cells showed only weak binocular interactions. Ten of these cells were responsive to stimulation of either eye, but their signal-to-noise ratios for the binocular interaction RF are low due to relatively low spike counts. The rest of the cells were either ocularly very unbalanced or not very responsive to stimulation of either eye. Therefore these 36 cells have been excluded from the analysis.

Examples of monocular RFs and binocular interaction RFs

Figure 4 shows examples of monocular RFs ($L$ and $R$) and binocular interaction RFs ($B$) for six simple cells. For each cell,

![Diagram of RFs](image-url)

**FIG. 4.** Examples of monocular RFs and binocular interaction RFs. One-dimensional profiles of left (L) and right (R) eye RFs and contour plots of binocular interaction RFs (B) are shown for 6 simple cells (A–F). The symbols, $X_L$ and $X_R$, denote positions in the left and right eyes, respectively. The axes, $X_F$ and $D$, represent position along the frontoparallel axis and binocular disparity, respectively. For each cell, monocular RFs and the binocular interaction RF are plotted on the same scale. Contour lines are drawn such that they divide the response amplitude between zero and either a positive or negative peak, whichever is greater, into 8 equally spaced levels. Solid and dashed contours indicate positive and negative values, respectively. Checkered pattern of the binocular interaction RF is a characteristic of simple cells. Cross-correlation delays: 44 (A), 49 (C and F), and 54 ms (B, D, and E).
the RFs are constructed at a common correlation delay, which is chosen from optimal correlation delays of the RFs. The optimal correlation delay of an RF is defined as the delay at which the sum of squared values of all data points in the RF is maximum. For a given cell, two monocular RFs and a binocular interaction RF generally had the same optimal correlation delay. When they had different optimal delays (differences never exceeded 20 ms), the one that maximizes signal-to-noise ratios of the RFs was chosen to be the common correlation delay.

Monocular RFs of the cells shown in Fig. 4, A and D, have similar profiles in the two eyes, indicating relatively small RF phase disparities. On the other hand, the rest of the cells have clearly different RF profiles in the two eyes, i.e., some degree of RF phase disparity. By definition, these RFs represent spatial structures that characterize the best linear transformation between stimulus and response (in a mean-squared error sense). In other words, if the left and right eye signals were summed linearly without any further nonlinear processing, these RFs would be sufficient to characterize the cell’s responses to binocular stimulation. However, binocular simple cells also exhibit nonlinear response properties, as evidenced by the binocular interaction RFs shown in the figure. This indicates that a cell’s response to binocular stimulation is determined not just by the structure of the monocular RFs but also by the structure of the binocular interaction RF.

The common feature of the binocular interaction RFs is their checkered patterns. The checker elements indicated by the solid and dashed contours represent combinations of positions in the left and right eyes at which the cell responds preferentially to the interocular polarity matched and mismatched stimuli, respectively. The polarity of the checker elements changes along the axes of stimulus position in the left eye ($X_L$) and in the right eye ($X_R$). This indicates that the binocular interaction of simple cells depends on the monocular stimulus position or phase. Because of this, the strength of binocular interaction depends not only on the stimulus binocular disparity ($D$), but also on the stimulus position or phase along the frontoparallel axis ($X_F$) where binocular disparity is constant. However, because the polarity of the checker elements does not change along the frontoparallel axis, integrating the binocular interaction RF along the constant disparity axis would yield a binocular disparity tuning function. Therefore binocular disparity tuning exhibited by simple cells is a consequence of their tuning for monocular phase in each eye and not for binocular disparity per se.

The checkered pattern also suggests that the binocular interaction RF is separable into left and right eye functions, i.e., the RF is described as the product of two functions—one for each eye. In fact, locations of checker elements seem to be aligned with locations of peaks and troughs of monocular RFs, implying that the left and right eye RFs may be the two functions. As described in methods, if a binocular simple cell has a system structure of a linear binocular filter followed by a static nonlinearity, the binocular interaction RF should be proportional to the product of the left and right eye RFs. This prediction is examined next.

**Structure analysis of binocular simple cells**

To determine if binocular interaction RFs are proportional to the product of left and right eye RFs, first qualitative comparisons are made between the predictions and raw data in Fig. 5. In the figure, binocular interaction RFs of three cells from Fig. 4 are shown on the left (Raw data). The product of the left and right eye RFs is computed for each cell and is shown on the right side of the figure (Prediction), along with 1D profiles of the left and right eye RFs. Contour plots for the predictions are quite similar qualitatively to those for the raw data, suggesting that they are proportional to each other. This is consistent with the results of Ferster (1981), who showed that the binocular disparity tuning of simple cells can be predicted by taking a cross-correlation between the left and right eye RF profiles (dot products of the left and right eye RF profiles at various interocular RF shifts).

This finding is further confirmed by the following quantitative comparisons. In Fig. 6, the value of each data point in the binocular interaction RF is plotted against that of the corresponding point in the predicted interaction RF, i.e., the product of left and right eye RFs, for each of the cells shown in Fig. 4. The solid lines indicate linear regression lines fitted to the data. Clearly, a straight line provides a good fit. Pearson’s correlation coefficient $r$ is indicated at the top right of each plot. The coefficients are very high ($>0.9$) for all cells shown here, suggesting that binocular interaction RFs are proportional to the product of left and right eye RFs.

Figure 7 shows a histogram of correlation coefficients for a population of cells examined. The distribution is strongly biaised toward high values, and ~80% of the cells have an $r$ value either equal to or $>0.75$. Therefore most binocular simple cells behave in a manner that is consistent with the model of a linear binocular filter followed by a static nonlinearity, as depicted in Fig. 3. Similar results have been obtained for temporal interaction data of simple cells (Emerson et al. 1989; Mancini et al. 1990).

Figure 7 also indicates that 10 cells (20% of sample) have correlation coefficients of $<0.75$; their binocular interaction RFs are correlated only moderately with the products of left and right eye RFs. One example of such cells is shown in Fig. 8. The binocular interaction RF (Raw data in Fig. 8A) of this cell is somewhat elongated along the frontoparallel axis and therefore cannot be described by the product of the left and right eye RFs (Prediction in Fig. 8B). When data points of the binocular interaction RF are plotted against those of the predicted RF (Fig. 8C), they scatter vertically around a linear regression line (—), resulting in only a moderate correlation ($r = 0.7$). Of 10 cells with correlation coefficients $<0.75$, 8 exhibit a left-right inseparable binocular interaction RF at one or more cross-correlation delays (see also Emerson 1997; Jacobson et al. 1993). Therefore these cells have a system structure that is different from that depicted in Fig. 3. However, it is not clear if they are real variations of simple cells or simple cell-like complex cells because binocular interaction RFs of complex cells are inseparable (Anzai et al. 1999b). In the following section, only those cells with a correlation coefficient of $\geq 0.75$ ($n = 39$) are considered to have the system structure illustrated in Fig. 3 and are subjected to further analysis.

**Shape of the static nonlinearity**

Having identified the system structure for binocular simple cells, one can proceed to estimating parameters for the com-
ponents of the system. There are three components in the system: a left eye linear filter, a right eye linear filter, and a static nonlinearity (Fig. 3). Because monocular RFs are already in hand, the shapes of the left and right eye filters are known. Only the shape of the static nonlinearity needs to be determined.

The shape of the static nonlinearity is obtained from its input-output function (see METHODS for details). As shown in Fig. 3, the input to the static nonlinearity $W(t)$ is the output of the linear binocular filter and can be estimated by convolving monocular RFs with the stimulus used to obtain the RFs. The output of the static nonlinearity $Y(t)$ is the spike train obtained as the cell’s response to the stimulus. Figure 9A shows an example of the input-output function plotted on-linear coordinates. Each dot represents a data point for a pair with input value $W$ and output value $Y$. Because $Y$ is a spike count, it is always positive and takes discrete values, whereas $W$ is continuous and can be negative. The horizontal axis is divided into bins (see the legend of Fig. 9A for details of binning), and mean $W$ and $Y$ values of the data points are computed for each bin. The mean data are indicated by open circles and open triangles in the figure. Note that the mean $Y$ values do not necessarily fall on the middle of the data ranges along the vertical axis. This is because the distribution of the data points along the axis is generally heavily biased toward zero. Solid lines connecting the open symbols represent the shape of the static nonlinearity. As seen in this example, the static nonlinearity has the shape of an expansive function.

Because the data points in Fig. 9A are scattered widely, one might wonder if the static nonlinearity is actually a half-rectification but noise in the system makes it look like an expansive function on average. It is also possible that the threshold for spiking changes from time to time. Then the shape of the static nonlinearity would be smeared when averaged over time and a half-rectification could look like an expansive function. Although we cannot rule out these possibilities, it is nonetheless important to characterize the shape of the static nonlinearity as a functional description of the cell.
The result shown in Fig. 9A indicates that, regardless of its true shape, the static nonlinearity acts like an expansive function on average.

As described in METHODS, an expansive function like seen in Fig. 9A can be well described by a half-power function (Eq. 2). The degree of expansion is represented by an exponent (n in Eq. 2) of the power function, which can be estimated quite easily by plotting the input-output function on log-log coordinates as shown in Fig. 9B. On log-log coordinates, the exponent corresponds to the slope of a straight line (see Eq. 3). We fit a straight line to three consecutive data points to find the maximum slope, which is taken as an estimate of the exponent for the static nonlinearity (see METHODS for details). For the example shown in Fig. 9B, the three data points indicated by filled circles yield the maximum slope of 2.08 for a straight line fit (—). The static nonlinearity of this cell is, therefore, approximately a half-squaring function. Note that the deviation from the straight line of the data points at \( W < 0.2 \) is predicted by the effect of a threshold (\( \theta \) in Eq. 3). It is also interesting that this cell does not show clear response saturation, despite the fact that instantaneous spike rates could exceed 400 spikes/s (Fig. 9A). This is in marked contrast to the response saturation seen in contrast response functions (e.g., Albrecht and Hamilton 1982; Anzai et al. 1995; Dean 1981; Maffei and Fiorentini 1973; Movshon and Tolhurst 1975; Sclar et al. 1990; Tolhurst et al. 1981). It is possible that response saturation is a consequence of adaptation to a prolonged exposure of the cell to a band-limited stimulus (such as a sinusoidal grating) of high contrast, and that, without such adaptation, cells can produce a much higher spike rate instantaneously.

Figure 10 shows more examples of the input-output function on log-log coordinates. The effect of a threshold is apparent at...
low $W$ values in any of the cells shown, but only a slight hint of response saturation can be seen in some cells (D–F). The maximum slope ($n$) of a straight line fit (—) varies from cell to cell, indicating that each cell has a different exponent. In Fig. 11, a histogram of exponents is shown for the population of simple cells examined. The exponent ranges from 1.32 to 3.11. The distribution has a mean of $2.17 \pm 0.53$ SD. Therefore the exponent of the static nonlinearity for binocular simple cells is, on average, $\sim 2$. Emerson and his collaborators (Emerson et al. 1989; Mancini et al. 1990) conducted a similar analysis on temporal interaction data of simple cells and found that a second-degree polynomial captures the main characteristic of the shape of the static nonlinearity. They concluded that the static nonlinearity is basically a half-squaring function, which is concordant with the results presented here.

**DISCUSSION**

In this study, white noise analysis has been applied to measurements of binocular interaction RFs and monocular RFs for simple cells in the cat’s striate cortex. Binocular interaction RFs of most simple cells are found to be proportional to the product of the left and right eye RFs. This indicates that the binocular interaction depends not only on stimulus binocular disparity but also on stimulus position or phase in the left and right eyes. The binocular interaction RF is consistent with that of a linear binocular filter followed by a static nonlinearity. The static nonlinearity is well characterized by a half-power function with an average exponent of $\sim 2$, i.e., a half-squaring function. This squaring nonlinearity is an implementation of a multiplicative operation and may play a fundamental role in computations performed by simple cells. In the context of binocular information processing, the squaring nonlinearity makes the initial linear convergence of signals from the left and right eyes multiplicative at the output of simple cells. This multiplicative binocular interaction is a key ingredient for the computation of interocular cross-correlation, an algorithm for solving the stereo correspondence problem. Therefore the process of solving the stereo correspondence problem may begin with these binocular simple cells.

**Binocular interaction RF and binocular disparity tuning**

The binocular interaction RF is a response map of nonlinear binocular interaction. It describes how a cell responds to stimuli at various positions in the left and right eyes, compared with the prediction from a linear sum of responses to stimulation of either eye. Therefore it represents the tuning of a cell for binocular disparity and how it depends on stimulus position in each eye.

In most previous studies, disparity tuning was measured with moving bars or extended stimuli such as random-dot stereograms. As a result, the dependency of the tuning on monocular stimulus position could not be examined. Binocular interaction RFs reported in this study reveal that binocular interactions exhibited by simple cells do depend on monocular stimulus positions and in a predictable manner. The binocular interaction RF is proportional to the product of the left and right eye RFs. Therefore the binocular disparity tuning of a simple cell, which may be obtained by integrating its binocular...
interaction RF along the frontoparallel axis, is predictable from its monocular RFs.

Ferster (1981) described how the binocular disparity tuning of simple cells can be predicted from monocular RFs. He computed dot products of left and right eye RFs for various interocular RF shifts (i.e., binocular disparities) and obtained the predicted tuning for binocular disparity. He found that the predicted and measured tuning matched very well. This computation corresponds to a cross-correlation between left and right eye RFs and is operationally equivalent to deriving a binocular interaction RF as the product of left and right eye RFs (as shown in Fig. 5, right) and integrating the binocular interaction RF along the frontoparallel axis (see Fig. 12). Because the measured binocular interaction RF is proportional to the product of left and right eye RFs, this is, in fact, an appropriate way of predicting binocular disparity tuning from monocular RFs.

This predictability of binocular disparity tuning from monocular RFs implies that there will be some relationships between the parameters of binocular disparity tuning and monocular RFs. First of all, the cell’s optimal disparity should correspond to the distance between the peaks of left and right eye RFs, i.e., the RF phase disparity. When the RF phase disparity is small, i.e., left and right eye RFs are similar in shape, the disparity tuning function is expected to be symmetric around the peak of the tuning function because the binocular interaction RF is symmetric around the axis of a constant binocular disparity that goes through the peak (Fig. 12A).

Therefore the profiles of the monocular RFs are important in determining the shape of the binocular disparity tuning for simple cells.

Because binocular disparity tuning depends on monocular RFs, simple cells are not truly tuned to binocular disparity per se. They are simply tuned to the spatial phases of left and right eye stimuli. However, because the polarity of the checkered pattern in the binocular interaction RF does not change along the frontoparallel axis (see Fig. 4), a group of simple cells that have the same RF phase disparity but different monocular RF phases can represent a binocular disparity independent of monocular stimulus phase. As shown in the following paper (Anzai et al. 1999b), binocular interaction RFs of complex cells are consistent with this scheme.

2 For RFs that are modeled as a Gabor function, the distance between peaks in the left and right eye RFs is actually slightly smaller than the RF phase disparity; as the RF phase disparity increases, the interocular peak distance increases slightly less. However, the difference is generally insignificant.
Do simple cells respond to random-dot stereograms?

The behavioral demonstration through the use of random-dot stereograms that binocular disparity alone is sufficient to mediate the perception of depth (Julesz 1960) illuminated a fundamental aspect of stereoscopic depth perception. It revealed that recognition of object form is not necessary to solve the stereo correspondence problem, which in turn implies that the correspondence problem may be solved at very early stages of binocular information processing.

Physiological evidence supporting this implication was provided by Poggio and his collaborators, who found that some cells in V1 and V2 of macaque monkeys respond to cyclopean stimuli embedded in random-dot stereograms (Poggio 1990; Poggio et al. 1985, 1988). Interestingly, these cells were predominantly complex cells, which suggests that mostly complex cells are responsible for solving the correspondence problem. It also suggests that the hierarchical notion of Hubel and Wiesel (1962) that simple cells feed into complex cells may not be correct because stimuli that complex cells respond to must, by that model, also be effective for simple cells. Are simple cells really not responsive to random-dot stereograms?

In fact, random-dot stereograms are not the only stimuli that are reportedly ineffective for simple cells. Hammond and MacKay (1975, 1977) claimed that complex cells but not simple cells respond to monocularly presented moving random-dot patterns (see also Morrone et al. 1982). If simple cells do not respond to monocular random-dot patterns, then it is not surprising that they do not respond to random-dot stereograms, either. However, Hammond and MacKay’s results later were challenged by studies that demonstrated that simple cells do...
respond to random-dot patterns (Skottun et al. 1988; see also Casanova et al. 1995; Gulyas et al. 1987). There is also a theoretical framework that predicts that simple cells should respond to such patterns (Grzywacz and Yulli 1990, 1991), and simple cell behavior matches with that predicted by the theory (Skottun et al. 1994). Furthermore the fact that RFs of simple cells can be mapped with the 2D white noise used in the previous paper (Anzai et al. 1999a; see also Jacobson et al. 1993; Reid et al. 1997) is compelling evidence that they do in fact respond to random-dot patterns.

Likewise, the fact that simple cells respond to dichoptic white noise and exhibit binocular interactions as shown in Fig. 4 strongly suggests that they should respond to random-dot stereograms. Random-dot stereograms can be considered a special case of white noise; the left and right eye patterns are both monocularly white but are interocularly correlated. Non-linear interactions exhibited by simple cells are, in general, of low orders (perhaps the first few), and the strength of interactions declines progressively as the order of interaction increases (Mancini et al. 1990). Therefore responses of simple cells to cyclopean stimuli in random-dot stereograms are likely due to low-order interactions, mostly of the second order. Then the binocular interaction RF, which represents second-order binocular interactions, should indicate how a cell responds to random-dot stereograms. In other words, the disparity tuning obtained by integrating the binocular interaction RF along the frontoparallel axis (as illustrated in Fig. 12) should be very similar, if not identical, to the tuning obtained with random-dot stereograms.

Then why did Poggio’s group find that the overwhelming majority (90%) of cells that respond to random-dot stereograms are complex cells (Poggio 1990; see also Cumming and...
Parker 1997; Gonzalez et al. 1993)? The binocular interactions exhibited by simple cells depend on monocular positions, as shown in Fig. 4. This indicates that simple cells will not respond well to stereograms with monocular phases that are not optimal for them. Because the monocular phase of dynamic random-dot stereograms changes constantly, it is to be expected that simple cells would respond in a sporadic rather than a sustained fashion to the stereograms. Therefore it may be difficult to associate their responses to the binocular disparity of the cyclopean stimulus. However, if monocular spatial phases of the stimulus are distributed evenly over the stimulus presentation period and measurements are repeated many times, responses of simple cells to random-dot stereograms should become apparent, and the binocular disparity tuning would emerge as was the case for a minority of simple cells reported in the previous studies.

System structure for binocular simple cells

In this study, the system structure for most binocular simple cells has been identified as a linear binocular filter followed by a static nonlinearity. This result is concordant with the results of previous studies that conducted similar analyses on simple cells. Mancini (Mancini 1983; Mancini et al. 1990) measured responses of simple cells in the cat’s striate cortex to temporal white noise generated according to binary m-sequences at various positions over the RF. He obtained a temporal profile of the RF as well as a profile for second-order temporal interaction and found that responses of simple cells can be well described by a model of a linear temporal filter followed by a static nonlinearity.

Emerson et al. (1989) successfully applied a more general structure comprising a cascade of a linear filter, a static nonlinearity, and another linear filter to describe responses of simple cells. Although one would not expect a linear filter to follow the output nonlinearity of simple cells, certain aspects of spike generation (e.g., a slow inactivation of sodium channels) (French and Korenberg 1989) and temporal binning of spikes in the analysis could introduce additional temporal filtering. Therefore some deviations from the proportionality condition of Eq. 1 seen in our data could be accounted for by the linear filter after the static nonlinearity. Nonetheless the fact that most simple cells satisfy Eq. 1 indicates that the second linear filter is a minor component, if necessary, to model simple cells. Indeed, the model’s performance does not change very much with or without it (Jacobson et al. 1993).

In contrast to these findings, Jacobson et al. (1993) found that the structure of a linear filter followed by a static nonlinearity can explain, on average, only ~60% of the responses of simple cells in the striate cortex of macaque monkeys. They measured responses of simple cells to white noise and obtained monocular RFs as well as monocular spatiotemporal (second-order) interaction RFs. They show in their paper some examples of interaction RFs that are elongated (i.e., inseparable) and therefore cannot be described by the product of monocular RFs. A minority of the simple cells examined in our current study also exhibit inseparable binocular interaction RFs at one or more cross-correlation delays. These results suggest that some simple cells are not consistent with a model of a linear filter followed by a static nonlinearity; their structure may consist of parallel streams of a linear filter followed by a static nonlinearity (Jacobson et al. 1993). However, it is not clear if these cells are real variations of simple cells or simple cell-like complex cells since complex cells exhibit inseparable binocular interaction RFs (Anzai et al. 1999b).

It should be pointed out that the system structure estimated in this study is by no means complete. It has been known that simple cells exhibit various other nonlinear properties. For example, responses of cells are normalized according to stimulus contrast, which is known as contrast gain control or contrast normalization (e.g., Albrecht and Geisler 1991; Bonds 1991; Geisler and Albrecht 1992; Heeger 1992a; Ohzawa et al. 1982, 1985). The gain control signal presumably is provided by a group of other cortical cells as a feedback signal. Because the noise stimuli used in this study have an average contrast that is relatively constant over time, the response gain of the cell is also expected to be relatively steady. Therefore the effect of the feedback signal can be considered constant, and the feedback circuitry can be separated effectively from the feedforward circuitry. In other words, the structure studied here only applies to the feedforward circuitry. There are also known inhibitory influences originating outside of the classical RF such as end and side inhibition (e.g., DeAngelis et al. 1994; DeValois et al. 1985; Hubel and Wiesel 1968; Kato et al. 1978; Maffei and Fiorentini 1976). In this study, the stimuli used were only slightly larger than the classical RF. Therefore inhibitory surrounds were not stimulated to any great extent. These nonlinear mechanisms need to be examined separately to build a more complete model of simple cells.

Static nonlinearity of simple cells

The results of this study show that the static nonlinearity of simple cells is a half-power function with an exponent of ~2. This suggests that the static nonlinearity of simple cells performs a nonlinear computation that is more than just thresholding. If the static nonlinearity were to serve as only a threshold, a half-rectification (an exponent of 1) would be sufficient. In that case, the output would be proportional to the input that exceeds a threshold, and therefore the underlying computation represented by the static nonlinearity would be essentially linear above the threshold. The fact that the exponent of a half-power function ranges approximately from 1.32 to 3.11 (much larger than 1) suggests that the expansive nonlinearity may be fundamental to the computations performed by simple cells.

It is interesting that the range of exponents is rather small. Any exponent other than 1 signifies some sort of nonlinear computation, but is there any reason why the exponent needs to be in this range? Obviously, the exponent should be significantly higher than 1 for simple cells to perform nonlinear computations without restricting the response dynamic range (exponent values <1 also represent nonlinearities, but they are of a compressive type). However, if exponents are too high, then a half-power function becomes similar to an over-rectification, i.e., a rectification (the exponent is 1) with a high-threshold and high gain (slope). Therefore it may be approximated as linear for inputs above the threshold. Although the sensitivity to small change in input would increase, a high gain also has an undesirable effect of reducing the input range that cells can encode because the output reaches the maximum quickly as input increases. Taken together, the range of expo-
ments seen among simple cells may reflect a range suitable for nonlinear computations that can be implemented within the limitation imposed by the maximum firing rate.

Given that the exponent is somewhere \( \sim 2 \), what kind of computations can be achieved by the static nonlinearity? The exponent of 2, a squaring, is an attractive operation from a computational point of view. First of all, because simple cells are selective to spatial frequency and phase, their output, if squared, corresponds to something analogous to a phase specific component of Fourier energy in a local region of the stimulus. This may be an ideal way of preserving local amplitude and phase information (Pollen and Ronner 1982). Second, the squared output of a linear filter is a building block for an energy model (Adelson and Bergen 1985; Ohzawa et al. 1990; Watson and Ahumada 1985). Third, the squaring enhances stimulus selectivity (Albrecht and Geisler 1991; J. L. Gardner, A. Anzai, R. D. Freeman, and I. Ohzawa, unpublished data); the tuning of cells for stimulus parameters such as orientation and spatial frequency becomes narrower, and the tuning band edges steeper, than would be without squaring. Finally, the squaring makes second-order interactions multiplicative. This is an important consequence of having an exponent near 2 because multiplication is a fundamental nonlinear operation. The implication of this multiplicative nonlinearity for functional roles of simple cells in binocular information processing will be discussed later. It should be noted that the above arguments should not depend critically on the exponent being exactly 2.

The neural bases and/or biophysical mechanisms responsible for the expansive nonlinearity are not known. One possibility is that spike generation at the soma is a function of the square of the average membrane potential over time. Another possibility is that the expansive nonlinearity can be a form of a dynamic nonlinearity, such as contrast normalization (Heeger 1992a,b). In this scheme, the static part of the nonlinearity is considered a half-rectification, i.e., the exponent is 1. However, the response gain (the slope of the half-rectification) and threshold (the position of the rectification) change dynamically according to stimulus contrast (assuming that the contrast is relatively low to avoid response saturation) such that the time average of the dynamic nonlinearity mimics a static nonlinearity with an exponent near 2 (see Suarez and Koch 1989 for a similar model). Because the gain normalization signal is thought to come from a group of other cortical neurons (Heeger 1992a), a feedback circuitry is involved in mediating the multiplicative nonlinearity in this scheme. There is also a suggestion that recurrent cortical excitation could amplify input signals (e.g., Douglas et al. 1995; Somers et al. 1995).

Multiplicative operations can be performed at dendritic trees as well. Mel (1992, 1993) showed that a model pyramidal cell driven by strong \( N \)-methyl-d-aspartate synaptic currents and/or containing dendritic Ca\(^{2+}\) or Na\(^+\) channels, responds more strongly to synaptic inputs that are spatially clustered than to those distributed diffusely. Therefore such a neuron could perform multiplications among the neighboring synaptic inputs and sum the results along the dendritic trees. This type of neuron is equivalent to what is known as a Sigma-pi neuron (Rumelhart et al. 1986), and its potential importance in nonlinear computations has been suggested (e.g., Durbin and Rumelhart 1989; Koch and Poggio 1992; Rumelhart et al. 1986). Whether or not real neurons, including simple cells in the striate cortex, are Sigma-pi neurons remains to be seen.

**Is the binocular interaction exhibited by simple cells linear or multiplicative?**

Ohzawa and Freeman (1986) measured the tuning for interocular phase disparity using drifting sinusoidal gratings to study the binocular interactions exhibited by simple cells in the cat’s striate cortex. They found that most cells show tuning that is consistent with the predictions of linear binocular summation. On the other hand, Ferster (1981) measured the binocular disparity tuning of simple cells in the cat’s striate cortex using moving bright bars and found that the disparity tuning can be predicted by taking a cross-correlation between left and right eye RF profiles. This result suggests that binocular interaction is multiplicative.

The results obtained in our current study offer a resolution to this apparent contradiction regarding the binocular interaction exhibited by simple cells. The system structure for binocular simple cells has been identified as a linear binocular filter followed by a half-power function with an exponent near 2. This can be formulated as

\[
(W_L + W_R)^2 = W_L^2 + W_R^2 + 2W_LW_R \quad \text{if } (W_L + W_R) > 0 \text{ and }
\]

\[
= 0 \quad \text{otherwise}
\]

where \( W_L \) and \( W_R \) are the output of the left and right eye linear filters, respectively. This equation indicates that the initial convergence of binocular signals is linear (the term \( W_L + W_R \) on the left hand side of the equation), but a squaring nonlinearity makes the binocular interaction multiplicative (the term \( 2W_LW_R \) on the right hand side of the equation). Therefore the answer to the question of whether binocular interaction is linear or nonlinear depends on whether one looks at the input stage (before the static nonlinearity) or the output stage (after the static nonlinearity). Although linear filtering is an important function of simple cells, one should incorporate the static nonlinearity into the description of their functional roles because what is important is the spike activity (after the static nonlinearity) that is sent to the next processing stages. Thus it seems appropriate to consider the binocular interaction as nonlinear (multiplicative). This interpretation offers an important functional role for binocular simple cells, as discussed next.

**Functional roles of binocular simple cells**

The fact that the binocular interactions exhibited by simple cells is multiplicative has an important implication as to their functional role in processing binocular information. It has been suggested that the stereo correspondence problem can be solved by taking the interocular cross-correlation of stereo images (Jenkin and Jepson 1988; Sanger 1988). For cortical cells to compute an interocular cross-correlation, they must be able to perform multiplication between left and right eye signals. Because simple cells exhibit multiplicative binocular interactions, they potentially could compute something analogous to an interocular cross-correlation to solve the stereo correspondence problem.

As formulated in Eq. 4, simple cells sum the outputs of left and right eye linear filters. The results then are rectified and
squared. Because the output of a linear filter is a weighted sum of the stimulus over space, the dot-product of the stimulus and the RF, the first line of Eq. 4 (for the positive output of the linear binocular filter) can be rewritten as

\[
\left\{ \int L(X_i)S_i(X_i)dx_i + \int R(X_i)S_a(X_a)dx_a \right\}^2
\]

\[
= \left\{ \int L(X_i)S_i(X_i)dx_i \right\}^2 + \left\{ \int R(X_i)S_a(X_a)dx_a \right\}^2 + 2 \int L(X_i)S_i(X_i)dx_i \int R(X_i)S_a(X_a)dx_a
\]  

(5)

where \(L\) and \(R\) denote left and right eye RFs, respectively. The variables \(S\) and \(X\) represent stimulus and stimulus position, respectively, and their subscripts \(L\) (left) and \(R\) (right) denote the eye of origin. The first two terms of the right hand side of the equation represent responses due to monocular stimulation. Therefore these terms do not depend on binocular disparity. The third term depends on binocular disparity and is presumably relevant to the computations leading to binocular fusion and stereopsis. Using the relationship that the binocular interaction RF is proportional to the product of left and right eye RFs, the third term can be rearranged as

\[
2 \int L(X_i)R(X_i)S_i(X_i)S_a(X_a)dx_i dx_a
\]

\[= \alpha \int B(X_i, \delta)S_i(X_i)S_a(X_i + \delta)dx_i dx_a\]  

(6)

where \(B\) denotes a binocular interaction RF and \(\alpha\) indicates a constant. By changing variables \(X_i L\) and \(X_i R\) to \(X\) and \(X + \delta\), respectively, Eq. 6 becomes

\[
\alpha \int B(X, X + \delta)S_i(X)S_a(X + \delta)dx\delta
\]

(7)

where \(X\) indicates a monocular stimulus position and \(\delta\) denotes a binocular disparity. Comparing Eq. 7 with an interocular cross-correlation, \(\Phi_{iR}(\delta)\), of the stimulus given by

\[
\Phi_{iR}(\delta) = \int S_i(X)S_a(X + \delta)dx
\]

(8)

it is clear that Eq. 7 is an interocular cross-correlation that is weighted according to the binocular interaction RF, which then is integrated over all binocular disparities. Therefore the computation performed by binocular simple cells includes interocular cross-correlation due to the static nonlinearity.

This interpretation requires the following comments. First, the output of simple cells contains monocular terms as indicated in Eq. 5 while the interocular cross-correlation defined in Eq. 8 does not. Therefore strictly speaking, simple cells do not compute interocular cross-correlation. However, because the monocular terms in Eq. 5 are independent of binocular disparity, responses to cyclopean stimuli entirely depend on the binocular term in Eq. 5 (i.e., Eq. 7). In this sense, simple cells can be considered to be computing something analogous to the cross-correlation of left and right eye images that are band-pass filtered.

It is also important to realize that the interocular cross-correlation performed by simple cells is local, i.e., the computation is restricted within its RF. Therefore they do not provide the complete solution to the stereo correspondence problem (Cumming and Parker 1997). In fact, they signal false matches as well. However, false matches can be rejected easily by combining the local solutions at various spatial locations, scales/frequencies, and orientations (Fleet et al. 1996).

Finally, the computation of interocular cross-correlation depends on the monocular spatial phase of a stimulus because simple cells are sensitive to monocular spatial phase. However, complex cells, which are not phase-sensitive, can provide phase independent interocular cross-correlation, as demonstrated in the following paper (Anzai et al. 1999b).

APPENDIX

Derivation of Eq. 1

Suppose that a binocular simple cell has the system structure of a linear binocular filter followed by a static nonlinearity, as depicted in Fig. 3. The output of the linear filter \(W(t)\) in Fig. 3) is described by a sum of convolution integrals at various positions in space

\[W(t) = \sum_{i} L_i(t)S_i(t - \tau)dt \]

(Al)

where \(L\) and \(R\) are impulse response functions for the left and right eye linear filters, respectively, and \(S\) denotes the input stimulus generated according to a binary m-sequence. The subscripts, \(i\) and \(j\), indicate stimulus positions in the left and right eyes, respectively. The static nonlinearity \(N\) in Fig. 3 can be represented, with any desired accuracy in a mean-squared error sense, by a polynomial of an arbitrary degree, provided that it is continuous within a given interval (Billings and Fakhouri 1978). Therefore the output of the neuron \(Y(t)\) in Fig. 3) can be described as

\[Y(t) = \sum_{j} \gamma_d |W(t)|^d\]

(A2)

where \(\gamma_d (d = 0,1,2,\cdots)\) is a constant. Substituting Eq. Al into Eq. A2 yields

\[Y(t) = Y_0 + \gamma_1 \left( \sum_{i} L_i(t)S_i(t - \tau)dt + \sum_{j} R_j(t)S_{j,1}(t - \tau)dt \right) + \gamma_2 \left( \sum_{i} L_i(t)S_i(t - \tau)dt + \sum_{j} R_j(t)S_{j,1}(t - \tau)dt \right)^2 + \gamma_3 \left( \sum_{i} L_i(t)S_i(t - \tau)dt + \sum_{j} R_j(t)S_{j,1}(t - \tau)dt \right)^3 + \cdots \]

(A3)

The left eye RF of the neuron is obtained by cross-correlating the output of the neuron and the stimulus for the left eye

\[\hat{L}_i(t) = E[Y(t)S_i(t - \tau)]\]

(A4)

where \(E[\ ]\) indicates the expected value of the quantity inside the brackets, and \(\hat{L}_i\) is the measured left eye RF

A binary m-sequence stimulus with a power density \(P\) and a
stimulus update period $\Delta$ has the following rth order correlation property

$$E\left[ \prod_{i} \{S_{i}(t - \tau_{i})\}^{n_{i}} \right] = \left( \frac{P}{\Delta} \right)^{n_{1} \cdots n_{r}}$$

where $r = \sum_{k} m_{k}$ (A5)

if $m_{k}$ is even for all possible values of $k$, and

$$E\left[ \prod_{i} \{S_{i}(t - \tau_{i})\}^{n_{i}} \right] = 0$$

(A6)

if $m_{k}$ is odd for at least one value of $k$.

Substituting Eq. A3 into Eq. A4, and using the property described above, the left eye RF becomes

$$L_{i}(\tau) = \gamma_{L}(\frac{P}{\Delta}) L_{i}(\tau) + 3 \gamma_{L}(\frac{P}{\Delta}) \sum_{j} L_{j}(\tau_{j}) \delta_{\tau_{j}}$$

$$+ \sum_{k} \int R_{k}^{L}(\tau_{k}) \delta_{\tau_{k}} + 2 \gamma_{L}(\frac{P}{\Delta}) L_{i}(\tau) + \cdots$$

(A7)

Assuming that no single $L_{i}(\tau)$ dominates, so that

$$L_{i}(\tau)\left\{ \sum_{j} \int L_{j}(\tau_{j}) \delta_{\tau_{j}} + \sum_{k} \int R_{k}^{L}(\tau_{k}) \delta_{\tau_{k}} \right\} \gg L_{i}(\tau)$$

(A8)

for any $i$ and $r$ and similar relationships hold for higher-order ($>3$) polynomial terms in Eq. A7, Eq. 7 becomes

$$L_{i}(\tau) = \beta_{n} L_{i}(\tau)$$

(A9)

where $\beta_{n}$ is a constant defined by

$$\beta_{n} = \gamma_{L}(\frac{P}{\Delta}) + 3 \gamma_{L}(\frac{P}{\Delta}) \left\{ \sum_{j} L_{j}(\tau_{j}) \delta_{\tau_{j}} \right\}$$

$$+ \sum_{k} \int R_{k}^{L}(\tau_{k}) \delta_{\tau_{k}} + \cdots$$

(A10)

Likewise, the right eye RF $\tilde{R}_{i}$ is described as

$$\tilde{R}_{i}(\tau) = \beta_{n} R_{i}(\tau)$$

(Equations A9 and A11 state that the measured monocular RFs are approximately proportional to the impulse response functions of the monocular filters.

Similarly, a binocular interaction RF $\tilde{B}_{i,j}$ is obtained by taking a cross-correlation among the output of the neuron and the left and right eye stimuli

$$\tilde{B}_{i,j}(\tau) = E[Y(t)S_{i}(t - \tau)S_{j}(t - \tau)]$$

(A12)

Substituting Eq. A3 into Eq. A12 and using the property described in Eqs. A5 and A6 and assuming Eq. A8 and similar relationships for higher order ($>3$) polynomial terms, Eq. A12 becomes

$$\tilde{B}_{i,j}(\tau) = 2 \gamma_{L}(\frac{P}{\Delta}) L_{i}(\tau) R_{j}(\tau) + 4 \gamma_{L}(\frac{P}{\Delta}) L_{i}(\tau) R_{j}(\tau) \left\{ \sum_{k} \int L_{k}^{2}(\tau_{k}) \delta_{\tau_{k}} \right\}$$

$$+ \sum_{k} \int R_{k}^{2}(\tau_{k}) \delta_{\tau_{k}} - 2 \left[ L_{i}(\tau) + R_{j}(\tau) \right] + \cdots = \beta_{b} L_{i}(\tau) R_{j}(\tau)$$

(A13)

where $\beta_{b}$ is a constant defined by

$$\beta_{b} = 2 \gamma_{L}(\frac{P}{\Delta}) + 12 \gamma_{L}(\frac{P}{\Delta}) \left\{ \sum_{k} \int L_{k}^{2}(\tau_{k}) \delta_{\tau_{k}} \right\} + \cdots$$

(A14)

Equation A13 states that the binocular interaction RF is approximately proportional to the product of the left and right eye impulse response functions. From Eq. A9, A11, and A13,

$$\tilde{B}_{i,j}(\tau) = a L_{i}(\tau) R_{j}(\tau)$$

(A15)

where $a = \beta_{b} / \beta_{m}$. In other words, the binocular interaction RF is approximately proportional to the product of the left and right eye RFs. Note that for a Gaussian white noise stimulus Eqs. A9, A11, A13, and A15 are exact (e.g., Marmarelis and Marmarelis 1978), not an approximation.

We are grateful to Dr. Eric Sutter for advice on binary m-sequences and their applications to receptive field mapping and to Dr. Stanley Klein for advice on nonlinear systems analysis and help with Appendix. We also are indebted to Dr. E. J. Chichilnisky for kindness in sharing with us his unpublished results regarding the shape of the static nonlinearity in retinal ganglion cells. We also thank Drs. Russel DeValois and Edwin Lewis for discussions and helpful comments and suggestions.

This work was supported by research and CORE Grants EY-01175 and EY-03176 from the National Eye Institute.

Address for reprint requests: R. D. Freeman, 360 Minor Hall, School of Optometry, University of California, Berkeley, CA 94720-2020.

Received 2 June 1998; accepted in final form 2 April 1999.

REFERENCES


