

A Neural Model of the Cortical Representation of Egocentric Distance

Alexandre Pouget and Terrence J. Sejnowski

Howard Hughes Medical Institute, Computational Neurobiology Laboratory, The Salk Institute, La Jolla, California 92037, and Department of Biology, University of California, San Diego, La Jolla, California 92093

Neurons in the visual cortex of monkeys respond selectively to the disparity between the images in the two eyes. Recent recordings have shown that some of the disparity-selective neurons in the primary visual cortex and the posterior parietal cortex are modulated by the distance of fixation. A population of such gain-modulated, disparity-selective neurons forms a set of basis functions of horizontal disparity and distance of fixation that can be used as an intermediate representation for computing egocentric distance. This distributed representation is consistent with psychophysical studies of human depth perception; in contrast, neurons explicitly tuned to distance are not consistent with how we perceive distance. In a population model that includes noise in the firing rates of neurons, the perceived distance is shown to be the estimate of geometrical distance that minimizes the variance of the estimation.

There are many visual cues to estimate the depth of an object, including stereopsis, motion parallax, shape from shading, and occlusion. These are relative depth cues and do not provide sufficient information to recover the position of an object with respect to the viewer (egocentric distance). For stereopsis, the depth of an object is measured relative to the fixation point of the two eyes, as illustrated in Figure 1. Because each eye sees an object from slightly different angles, the images are slightly displaced. The horizontal displacement, or horizontal disparity, is proportional to the distance between the object and the fixation point: zero disparity corresponds to an object at fixation, whereas positive and negative disparities correspond, respectively, to locations in front of or behind the fixation point (Fig. 1).

If the distance to the fixation point, also called the viewing distance, can be recovered, then the position of an object in egocentric coordinates can be estimated by combining this information with horizontal disparity. Psychophysical experiments indicate that at least two cues are used by the visual system to recover the viewing distance: the vergence angle and vertical disparities (von Hofsten, 1976, 1977; Ritter, 1977; Cumming et al., 1991; Rogers and Bradshaw, 1993). For simplicity, we consider only vergence in this report, but our approach can be easily extended to include vertical disparity. For an object directly in front of the viewer, the distance, \bar{D} , as a function of disparity, δ , vergence, ν , and interocular distance, I , is given by

$$D = \frac{I}{2 \tan\left(\frac{\nu - \delta}{2}\right)} \quad (1)$$

A three-dimensional plot of this function, with respect to ν and δ , is shown in Figure 8A. Although psychophysical experiments have demonstrated that humans can estimate absolute distance, this problem has only recently been studied with physiological methods (Gnadt and Mays, 1991; Trotter et al., 1991, 1992).

The pioneering work on the cat visual cortex (Nikara et al., 1968; Pettigrew et al., 1968) and later studies in a variety of species revealed that a large percentage of cells in the primary visual cortex and extrastriate areas are selective to horizontal disparity (Hubel and Wiesel, 1970; Poggio and Fischer, 1977; Ferster, 1981; Maunsell and van Essen, 1983; Poggio,

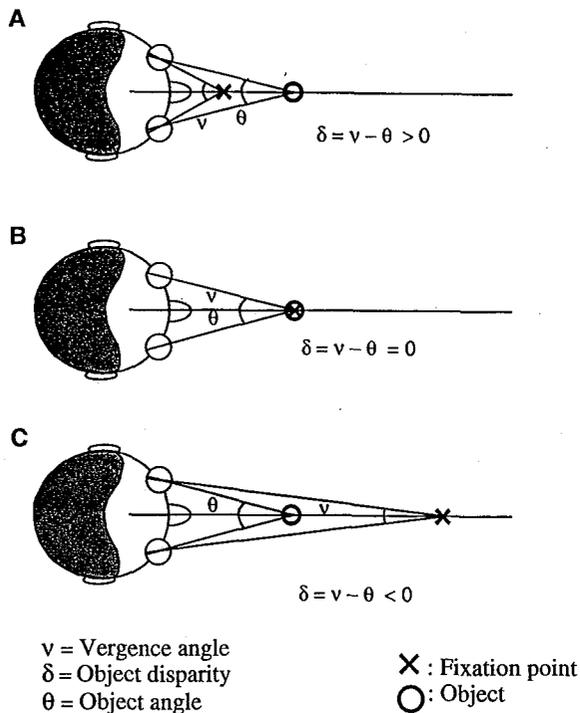


Figure 1. Viewing geometry for stereopsis. The horizontal disparity, δ , of a stimulus is equal to the difference between the vergence angle, v , and the object angle θ (sometimes called the absolute disparity of the object). The disparity δ depends on the position of the object with respect to the fixation point, which is shown in three positions: beyond fixation (A), at fixation (B), and in front of fixation (C). The egocentric distance of an object can be determined by combining a disparity measurement with vergence angle or any cue related to distance of fixation.

1984; Poggio et al., 1985, 1988; LeVay and Voigt, 1988). Poggio classified disparity-selective neurons into three groups, the near, tuned, and far cells. Tuned cells were characterized by narrow tuning with a peak close to zero disparity whereas near and far neurons showed broad tuning with peak responses at large values of disparity, positive for near cells and negative for far cells. LeVay and Voigt (1988) and Lehky and Sejnowski (1990) suggested these classes were only the extremes in a continuum of selectivities, and recent psychophysical studies provide confirming evidence (Stevenson et al., 1992; Cormack et al., 1993). Figure 2 shows the full spectrum of disparity-tuned responses that are typically found in the visual cortex.

In the above studies, the disparity selectivity of neurons was tested at a fixed distance, which confounded disparity and distance. The influence of these two variables can be distinguished by measuring the disparity selectivity of cells over a range of fixation distances. Figure 3 illustrates the two types of idealized responses that one might expect to find in such an experiment. One possibility is that the disparity tuning curve would be invariant under changes in distance of fixation: this would imply that the peak of the distance tuning curve should vary with the distance of fixation (Fig. 3A). The opposite extreme would be a cell that was tuned to distance, but with a disparity tuning that varied with distance to fixation (Fig. 3B).

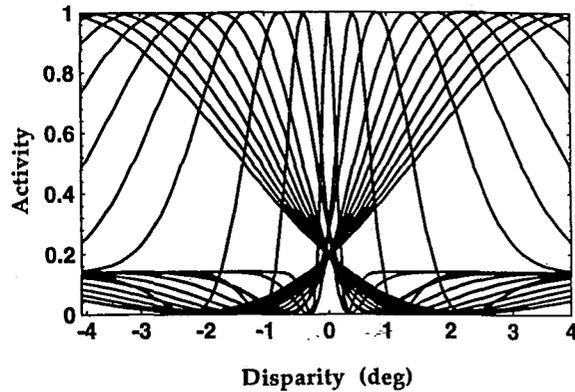
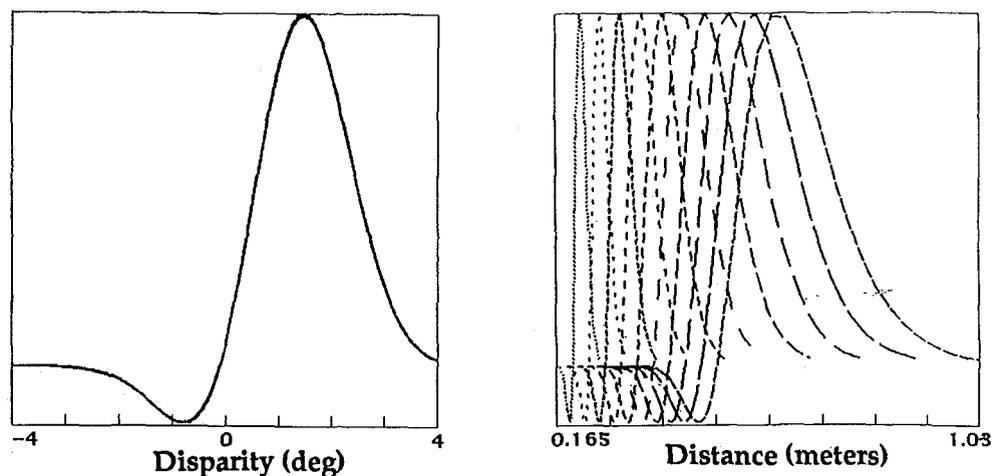


Figure 2. Idealized disparity tuning curves of cortical neurons. Each curve corresponds to the response of a single cell to disparity (see Eqs. 3–5). The curves illustrated here were sampled from a continuum of tuning curves. The widths of the tuning curves increase as the best disparities at the response peaks increase.

This experiment has been performed by Trotter et al. (1991, 1992) in VI, and Gnadt and Mays (1991) in the parietal cortex. They recorded the disparity tuning curves of cortical neurons for various viewing distances and found selectivities intermediate between these two extremes. The magnitudes of the responses to disparity were modulated by the distance to fixation but the shape and position of the peaks of the tuning curves were unchanged. Figure 4 shows four examples of these neurons (Trotter et al., 1992). Since the position of the peak of the disparity selectivity is invariant with changes in viewing distance, these neurons were not tuned to distance, but neither were they unambiguously tuned to disparity. This type of response was not unexpected since a previous neural network model trained to compute distance from pairs of vergence angles and disparity-tuned units exhibited such “gain” fields for distance (Lehky et al., 1990). However, this network model only provided limited insight into the nature of the representation used by the brain (see Discussion). The goal of this report is formally to characterize the representation found in the brain (Gnadt and Mays, 1991; Trotter et al., 1992) and in neural network models (Lehky et al., 1990) and to explore the computational advantages of this representation.

Our analysis relies on the theory of basis functions. This theory takes its root in 19th century mathematics, but it has only recently been applied to neural networks (Casdagli, 1989; Moody and Darken, 1989; Poggio, 1990). Specifically, we show that the gain-modulated neurons reported in the visual cortex form a set of basis functions for the representation of distance as well as other functions of disparity and vergence that are needed for guiding actions. The way humans perceive distance is consistent with this representation and corresponds to a low variance estimate of distance. Although gain-modulated neurons may indeed be involved in an intermediate representation of distance, our theoretical analysis shows that they may also be used to represent other functions of disparity and eye position.

A- Disparity Detector



B- Distance Detector

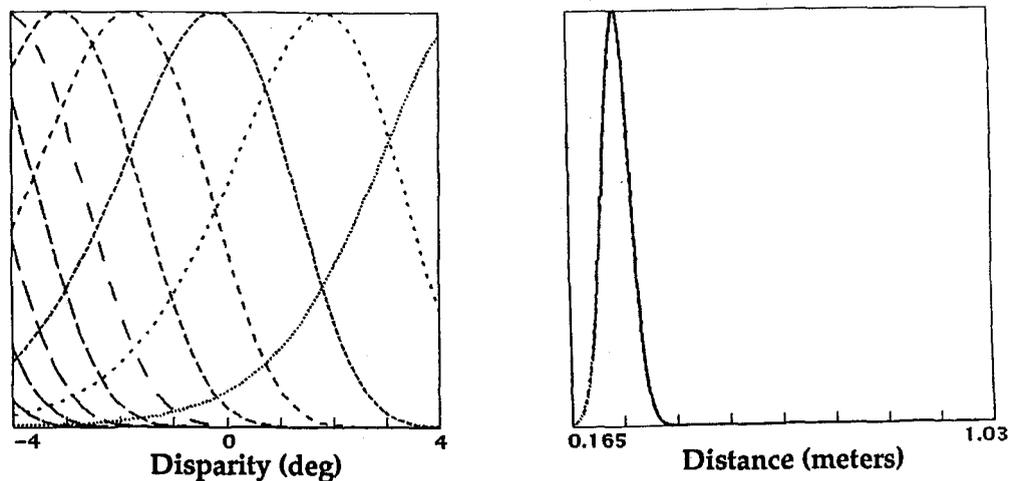


Figure 3. Responses of idealized disparity and distance tuning curves. *A*, Response of a pure disparity detector as a function of disparity (*left*) and as a function of distance (*right*), for 10 fixation distances (*dashed lines*). The response of a disparity detector to horizontal disparity would be unaffected by changes in fixation distance. *B*, Response of a pure distance detector is shown as a function of disparity (*left*) and as a function of distance (*right*), for 10 fixation distances (*dashed lines*). The response of a distance detector should only depend on egocentric distance.

Representing Egocentric Distance with Basis Functions

Basis Functions

A basic problem in the theory of function approximation is to approximate any function in a class of functions by a linear combination of fixed basis functions. For example, any well-behaved function can be approximated arbitrarily closely by a linear combination of sines and cosines summed over all possible frequencies and phases. The Fourier series is only one of many possible basis sets.

Sigmoid functions, commonly used in neural networks, and many radially symmetric functions can also be used as a basis set and have various advantages (Hornik et al., 1989; Moody and Darken, 1989; Poggio, 1990). The familiar gaussian function is radially sym-

metric and a family of gaussians with all possible centers and variances forms a basis set. Sigmoid functions with all possible discriminant surfaces in the input space also form a basis set, but one that breaks up the input space into extended regions rather than the compact regions of support formed by gaussian basis functions.

There is a close connection between approximation theory and neural networks (Lapedes and Farber, 1988; Poggio, 1990). In a three-layer feedforward neural network, the hidden layer of processing units represents input patterns by a population code that serves as an intermediate transformation between the inputs and outputs. There are many possible choices of output functions for the hidden units, depending on the nature of the transformation the network needs to perform and the number of hidden units available.

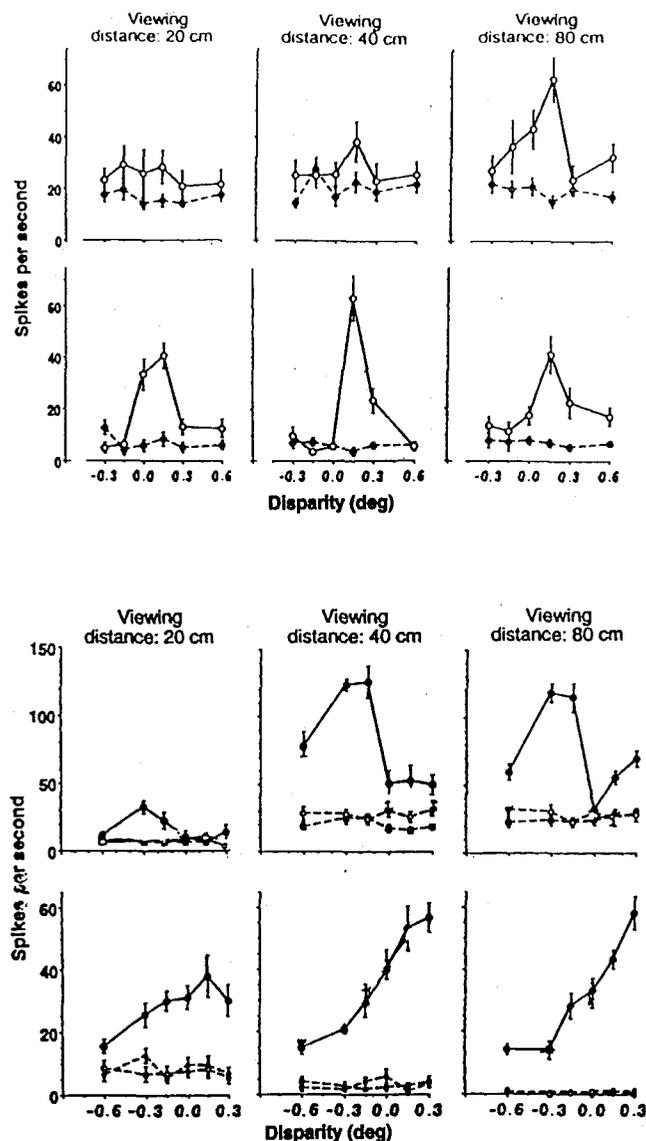


Figure 4. Experimental recordings showing that the disparity selectivity of V1 neurons is modulated by distance of fixation. Each row of panels is from one neuron and each panel gives the tuning curve for disparity at a different viewing distance. In all cases, only the amplitude of the response varies with fixation distance while the general shape of the disparity tuning curve and the position of response peak remain the same. These neurons are neither pure disparity nor pure distance detectors (see Fig. 3). Response to visual stimuli is shown by the solid lines and spontaneous activity by the broken lines. Adapted from Trotter et al. (1992).

One way to determine an appropriate set of hidden units is to use an optimization procedure such as backpropagation (Rumelhart et al., 1986). An alternative is to handcraft a good set of hidden units, such as a set of basis functions. Choosing the hidden units in advance greatly simplifies optimization since the input weights are fixed and only the weights from the hidden to the output units need to be determined.

Whether the brain makes use of basis functions is not yet clear from physiological data, but the problem of spatial localization may be one area where it can be tested (Poggio, 1990). We show in this section that the observed cortical representation of distance can be considered a new type of basis function that combines some of the best features of radial basis functions and sigmoids. Each basis function corresponds to the response of one gain-modulated neuron of the

type reported by Trotter et al. (1992) and Gnadt and Mays (1991). Figure 5 shows a neural network diagram of our model.

Methods

Our goal is to show that gain-modulated neurons can be used to approximate distance. We first specify a set of basis functions, $B_i(\delta, \nu)$, of disparity, δ , and vergence, ν , that are consistent with the observed responses of gain-modulated neurons in the visual cortex. The next step is to show that a linear combination of these basis functions can be used to estimate the distance, $D(\delta, \nu)$:

$$D(\delta, \nu) = \sum_{i=1}^n w_i B_i(\delta, \nu), \quad (2)$$

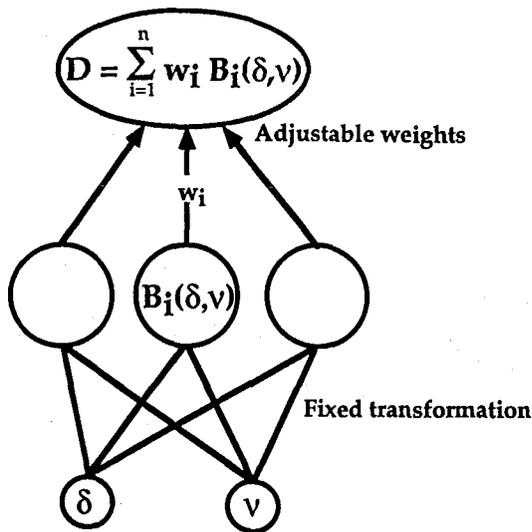


Figure 5. Neural network representation of a basis function model for computing egocentric distance from vergence and disparity. Each hidden unit receives as inputs the vergence angle, ν , and the disparity, δ , and computes a function of these two variables. The basis functions $B_i(\delta, \nu)$ on the hidden layer represent the gain-modulated neurons found in V1. In a backpropagation network, the hidden unit is usually a sigmoid function of a weighted sum of the inputs. The hidden units project to the output layer, which consists of a single unit that represents distance by its continuous output, D .

where the w_i are the weights of the network depicted in Figure 5.

Gain-modulated Units

We specified 41 tuning curves for disparity shown in Figure 2 (only every other curve is shown). These profiles were idealizations of real disparity tuning curves from monkeys and cats and were similar to the ones previously used by Lehky and Sejnowski (1990): *near units* (units with preferred disparity, δ_i , less than -1°):

$$a_i(\delta) = 1.12e^{-(\delta-\delta_i)^2/\sigma_i^2} - 0.35e^{-(\delta+(\sigma_i/2))^2/\sigma_i^2}, \quad (3)$$

far units (units with preferred disparity, δ_i , greater than 1°):

$$a_i(\delta) = 1.12e^{-(\delta-\delta_i)^2/\sigma_i^2} - 0.35e^{-(\delta-(\sigma_i/2))^2/\sigma_i^2}, \quad (4)$$

excitatory tuned units (units with preferred disparity, δ_i , within $[-1^\circ, 1^\circ]$):

$$a_i(\delta) = 1.5e^{-(\delta-\delta_i)^2/\sigma_i^2} - 0.5e^{-(\delta-\delta_i)^2/2\sigma_i^2}, \quad (5)$$

where δ_i is the peak of response for a given curve, and σ_i^2 is related to the width of the tuning curve. The response peaks were evenly spread over the disparity range $[-4^\circ, 4^\circ]$ and the σ_i^2 were chosen to equal the absolute value of the disparity corresponding to the peak response, except for the curves whose peak were within the disparity range $[-10', 10']$, for which the variance was set to 10 min of arc. We did not include tuned inhibitory cells since they could be modeled by tuned excitatory cells with negative output weights.

Disparity tuning curves were not available for neurons in humans, so these values were chosen to reflect data from psychophysical studies and recordings from other species. In cats, LeVay and Voigt recorded neurons that had peaks of disparity tuning that were dis-

tributed within this interval (LeVay and Voigt, 1988). Most recordings from monkeys have been from neurons in the foveal representation of primary visual cortex, $[-1^\circ, 1^\circ]$. Humans are above chance in perceiving disparity in the range $[-4^\circ, 4^\circ]$ (Westheimer and Tanzman, 1956; Richards, 1971), which suggests that the range of disparities over which cells respond in humans must extend well beyond $[-1^\circ, 1^\circ]$. In any case, the results presented in this report do not depend critically upon this range.

Vergence tuning curves were modeled as sigmoids and 10 different curves were chosen. The gain modulation of cortical neurons has not yet been fully characterized, but the sigmoid shape of the input-output function of most neurons is at least consistent with our choice. The expression used for the vergence selectivity, $z_i(\nu)$, was

$$z_i(\nu) = \frac{1}{1 + e^{-(\nu-\nu_i)/T_i}}, \quad (6)$$

where ν is the vergence angle, and ν_i and T_i are, respectively, the thresholds and the slopes, or "temperatures," of the sigmoids. The threshold controls the position of the sigmoid on the vergence axis, whereas the temperature controls the steepness of the sigmoid. We used 10 combinations of five thresholds ($7.78^\circ, 11.26^\circ, 14.76^\circ, 18.24^\circ, 21.73^\circ$) and two temperatures (3.45 and 1.15), which correspond to shallow modulation of the disparity responses with viewing distance.

Some of the gain modulation of neurons reported by Trotter et al. were nonmonotonic and would be better fit by a gaussian (see the second cell in Fig. 4). Using gaussian modulation in addition to sigmoid modulation would not affect the results presented in this report and for simplicity we only used sigmoids. We did not include monotonically decreasing sigmoids for they are equivalent to increasing ones multiplied by a negative weight.

Finally, by multiplying combinations of disparity and vergence selectivities, we obtained 10×41 gain-modulated responses, that is, a total of 410 basis functions:

$$B_i^g(\delta, \nu) = a_i(\delta)z_i(\nu) \quad (7)$$

Figure 6, A-C, shows three typical examples of these functions, plotting the responses as functions of vergence and disparity. These plots are idealizations of the neuronal responses shown in Figure 4. Each basis function models the response of a single cell to disparity and vergence.

Distance-tuned Units

For some of our simulations, we used a second set of functions tuned to distance rather than disparity in order to compare their properties with those of the above basis functions. The responses of these units were gaussian functions of distance similar to the ones shown in Figure 3B:

$$B_i^d(\delta, \nu) = e^{(D(\delta, \nu) - D_i)^2/2\sigma_i^2}, \quad (8)$$

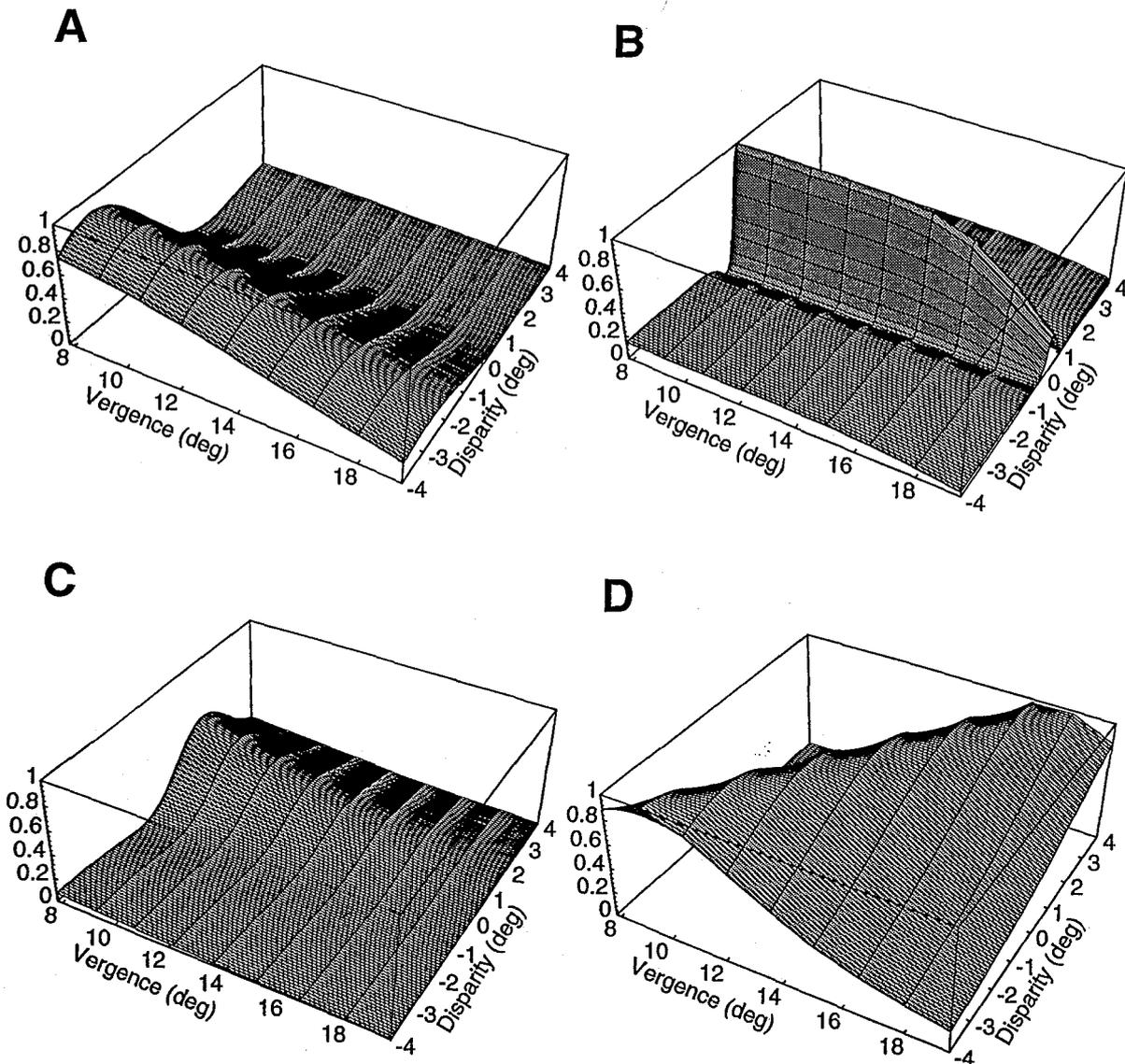


Figure 6. Typical tuning curves for the hidden units as functions of vergence and disparity. A-C, Three examples of idealized gain-modulated units. The tuning with disparity is a difference of gaussians and the vergence tuning is sigmoidal (see Eqs. 6, 7). D, A distance-tuned unit. Isodistance lines run along the diagonal in the vergence-disparity plane (see Eq. 8).

where $D(\delta, \nu)$ is the function given in Equation 1, and D_i is the preferred distance for basis function i . The σ were set to 0.07 in all $B_i^p(\delta, \nu)$.

A graph showing the response for one of these units is shown in Figure 6D. There were 410 distance-tuned units and the peaks of their responses were evenly distributed from 0.16 to 1.08 meters.

Output Representations: Geometrical versus Perceived Distance

Two different functions of distance were used for the output, one in which the output was proportional to geometric distance (Eq. 1), and a second function in which the output was proportional to perceived distance.

The need for this second function was motivated by experiments showing that the quantitative percept of depth is not a linear function of the stimulus disparity, as one might have expected (Richards, 1971;

Richards and Kaye, 1974). The linear relationship between perceived and geometrical disparity holds only in the range of $[-40', 40']$, outside of which the estimated disparity tends to decrease as the stimulus disparity increases (Fig. 7A). Consequently, humans cannot possibly perceive the actual distance of an object for all possible pairs of vergence and disparity. The perceived distance function can be obtained by replacing the disparity, δ , with the perceived disparity, δ_{perc} , in Equation 1. The perceived disparity was obtained by parameterizing the psychometric curves (Richards, 1971) (Fig. 7A,B):

$$\delta' = 1.8\delta e^{-11.5|\delta|^{0.9}}$$

$$\delta_{\text{perc}} = \begin{cases} \delta & \text{if } \delta' > \delta \\ \delta' & \text{otherwise} \end{cases} \quad (9)$$

Figures 8A and 9A show, respectively, geometrical distance and perceived distance as a function of disparity and vergence.

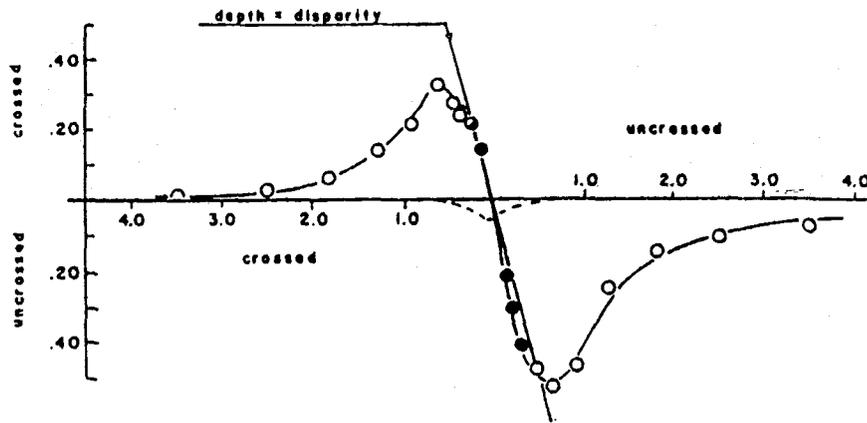
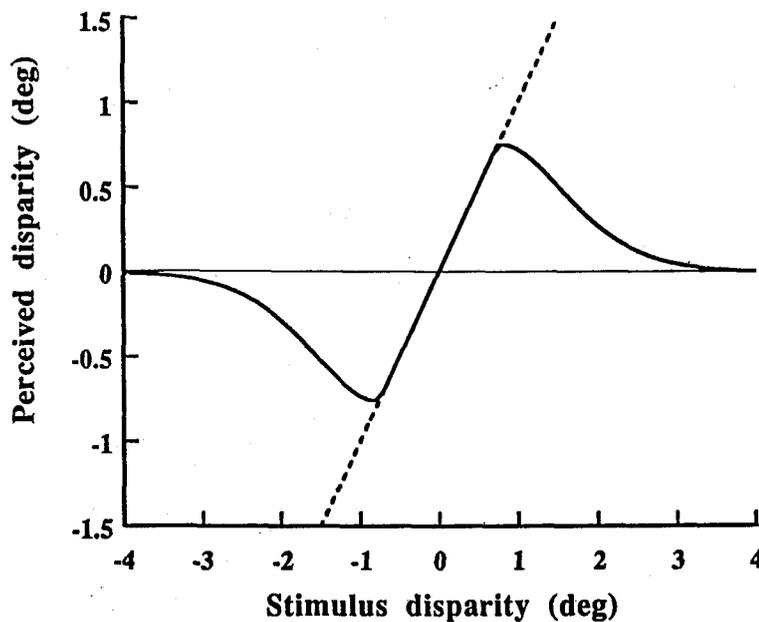
A**B**

Figure 7. Perceived disparity as a function of geometric disparity of the stimulus. *A*, Experimental data from Richards (1971). Perception is accurate (*straight line*) only within $[-40^\circ, 40^\circ]$. *B*, Parameterized curve in Equation 9. The *dashed line* shows the curve geometrical disparity. The curve we used is the mirror image of the one obtained by Richards because he used negative values for uncrossed disparities and positive values for crossed disparities.

The optimal set of weights were determined for all four combinations of the two input representations and the two output functions. The delta rule (Widrow and Hoff, 1960), an iterative optimization technique, was used to find the optimal set of weights for the network shown in Figure 5 (see Appendix). Because the second layer of the network is effectively a linear transformation, the delta rule is not subject to local minima and the weights are guaranteed to converge to the optimal network (Widrow and Stearns, 1985).

Results

Geometrical Distance

We first determined the set of weights that provided the best approximation to geometrical distance as a function of vergence and disparity using gain-modulated units as a basis set. In Figure 8, *A* and *B* show the actual function and the approximated function. The approximation had only 0.63% error, demonstrating that gain-modulated neurons form a set of basis

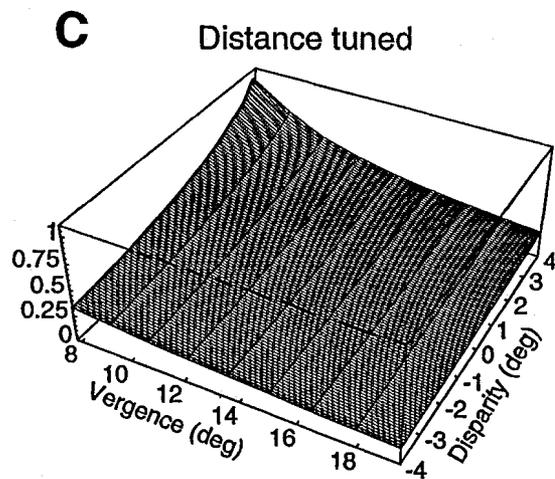
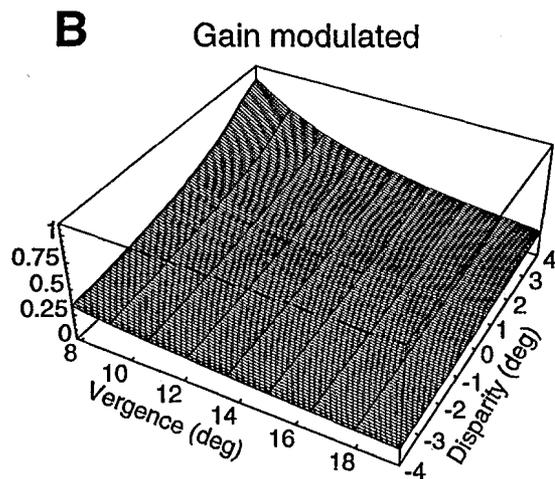
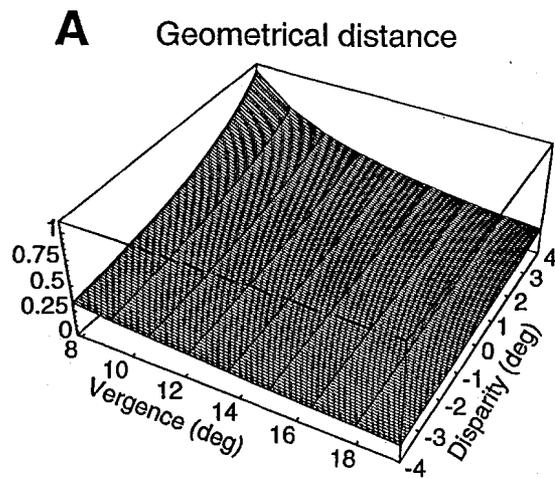


Figure 8. Network approximations of geometrical distance as a function of vergence and disparity. *A*, Geometrical distance. *B*, Approximation by a set of gain-modulated units. *C*, Approximation by a set of distance-tuned units.

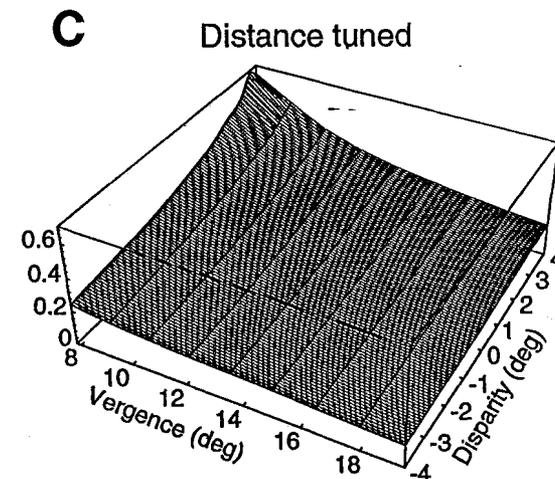
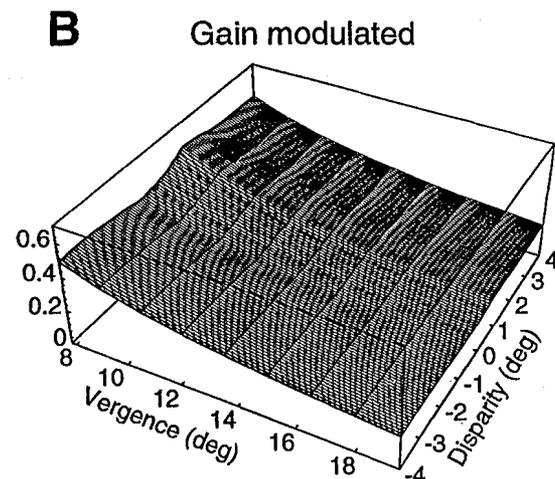
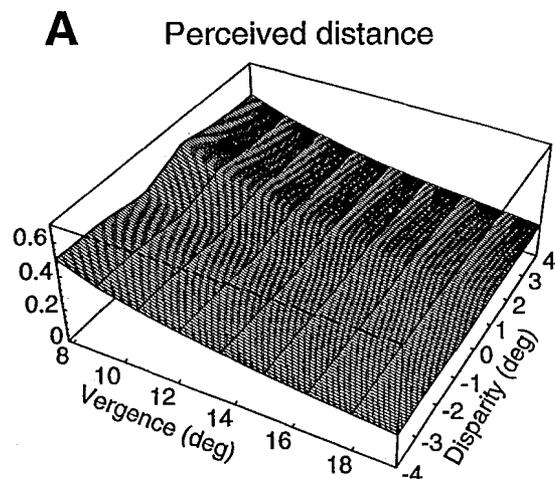


Figure 9. Network approximations to perceived distance as a function of vergence and disparity. *A*, Perceived distance. *B*, Approximation by a network of gain-modulated units. *C*, Approximation by a network of distance-tuned units.

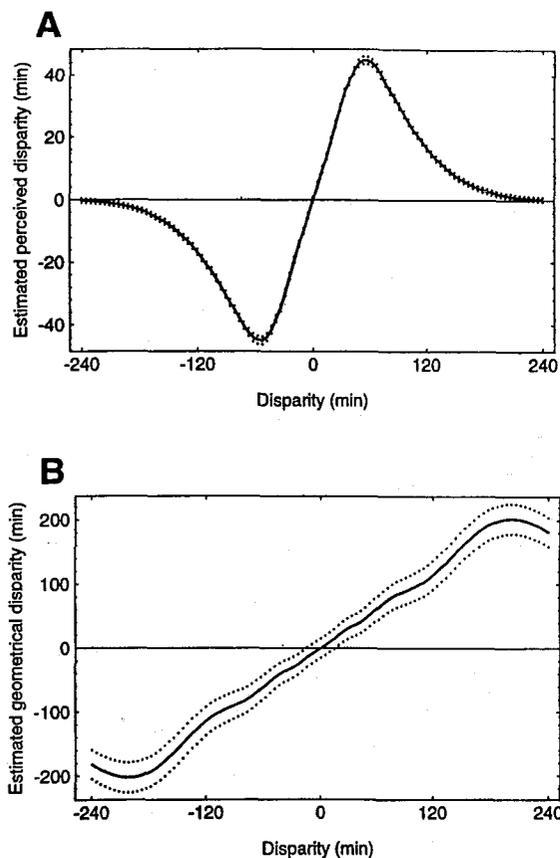


Figure 10. Estimation of the perceived (A) and geometrical (B) disparities from a distributed representation of disparity. The *solid lines* are the means of the estimations and the *dotted lines* are one standard deviation away. Note different scales on the two graphs: the standard deviation when estimating geometrical disparity is 20 times greater than the standard deviation for estimating perceived disparity.

functions that can be used for accurately approximating geometrical distance.

We then repeated the procedure with the set of functions made of distance-tuned units. As shown in Figure 8C, the best approximation had an overall error of only 0.76%, so distance-tuned units were as good for estimating distance as the basis functions.

Perceived Distance

Distance-tuned units have not been reported in the visual cortex even though they are adequate for representing distance and are conceptually simpler than a distributed representation. It is possible that such neurons do exist, but have not yet been found; alternatively, they may not be necessary or might even be a disadvantage. We provide a computational explanation for their absence.

It has been implicitly assumed that humans accurately perceive geometrical distance when only disparity and vergence are available as cues. In fact, humans tend to overestimate close viewing distances and underestimate far distances and the ratios are a function of the experimental paradigm (Foley and Held, 1972; Foley, 1980). This would not affect our previous results since both sets of functions can approximate this new distance function by simply multiplying the distance function by a constant (the gain).

A more serious problem arises from nonlinearities in the psychometric curve of perceived disparity as a function of geometrical disparity (Richards, 1971) (Fig. 7A). The perception of disparity appears to be accurate within the range of $[-40', 40']$ but reverses and returns to zero as the disparity increases. As a consequence, a stimulus with 4° of disparity is perceived at nearly the plane of fixation. As explained in Methods above, we derived the function giving the perceived distance as a function of disparity and vergence by replacing disparity in Equation 1 by perceived disparity (Fig. 9A). Using the same optimization technique as before, we found the best possible approximation of perceived distance for the disparity-tuned base units and the distance-tuned units.

Although the gain-modulated basis functions achieved 0.47% error (Fig. 9B), it was not possible to get a reasonable approximation of perceived distance when using the distance-tuned units (23.43% error) (Fig. 9C). This was because the distance-tuned units (Fig. 6D) were all oriented along the diagonals in the (ν, δ) plane, so they could only approximate functions whose gradients were perpendicular to this diagonal. In particular, they cannot be used to approximate perceived distance, which does not have this property. So even if distance-tuned units were present in the cortex, they could not be used for estimating perceived distance for some ranges of parameters.

If gain-modulated neurons could be used to approximate geometrical distance as well as perceived distance, and such neurons have been found in the cortex, then why does the brain not compute the actual geometrical distance? An answer to this question, given below, depends on noise. Cortical neurons fire action potentials with a Poisson distribution and the variance in the spike rate in response to a stimulus is proportional to the mean value of the response (Tolhurst et al., 1982; Vogels et al., 1989). If distance is estimated by pooling the responses of a neuronal population, then it is important to characterize the variance of the estimation, that is, the reliability of the estimated distance. We show in the next section that the variance depends on whether geometrical or perceived distance is being estimated from the response of the gain-modulated neurons.

Reliability of Distance Estimation

Bias-Variance Trade-off

In estimating the value of a quantity by averaging measurements from noisy data, there is a trade-off between the bias of the estimate—the difference between the mean of the estimator and the true value of the quantity—and the variance of the estimate. Estimators with smaller bias have a higher variance and low variance estimators have greater bias. A compromise is choosing an estimator between minimizing the bias and reducing the variance (Scott, 1992).

Estimating distance from the responses of noisy neurons is subject to this bias-variance trade-off. The bias corresponds to the difference between the esti-

mated and the "true" geometrical distance. Recovery of the true geometrical distance without bias leads to a high variance. We show here that the perceived distance, a highly biased estimate of geometrical distance especially for large positive and negative disparities, produces a smaller average variance.

Because disparity is the only important variable for the bias-variance trade-off, the analysis described in this section is limited to disparity estimation. The reasons for which vergence is excluded from the analysis are discussed in Results below.

Methods

We assume that there are pools of noisy neurons whose disparity selectivities are of the type shown in Figure 2. The goal is to use the information in this neural population to estimate the disparity.

Noise Model

The variance of an estimator depends on the type of noise present in the data. Single-cell recordings from the cortex of cats and monkeys (Tolhurst et al., 1982; Vogels et al., 1989) suggest that the variance, σ^2 , of the firing rate of a typical neuron is

$$\sigma^2 = Ka^b, \quad (10)$$

where K and b are constants and a is the mean activity (firing rate or total number of spikes) of a neuron for a given stimulus. Measurements of K vary from 1 to 3 and b is usually close to 1. We used $K = 2$ and $b = 1$ for most of our analysis, which are the average values for these parameters reported in awake animals (Vogels et al., 1989). Our results were not sensitive to the exact values of these constants. It is essential for our analysis only that σ^2 be proportional to the activity.

For a pool of neurons, i , with the same disparity tuning and independent noise for each neuron, the variance of the average firing rate in the pool, σ_i^2 , is inversely proportional to the number of cells, N_i , in the pool:

$$\sigma_i^2(\delta) = \frac{\sigma^2(\delta)}{N_i} = \frac{K}{N_i} a_i(\delta) \quad (11)$$

Note that σ_i^2 is a function of horizontal disparity, δ , because the activities of the neurons in the pool, a_i , are functions of disparity. We assume that the number of cells with a given selectivity for horizontal disparity is a gaussian function of the peak of the tuning curve, centered at zero disparity (LeVay and Voigt, 1988):

$$N_i = N_0 e^{-\delta^2/2\alpha^2}, \quad (12)$$

where N_i is the number of cells in the pool with preferred disparity δ_i , N_0 is the number of cells in the pool with a preferred disparity of zero, and α^2 is the variance of the distribution. The influence of N_0 and α^2 on the results is discussed in the next section.

It is important for our results that the disparity tuning curves shown in Figure 2 correspond to pools containing a variable number of neurons. In particular, there were many fewer neurons tuned to large disparities than to small disparities.

Our noise model does not capture all aspects of neuronal noise. It describes only how spike rates vary around their mean value while ignoring variations in interspike intervals. These variations would have been relevant if we had considered neural codes involving the temporal distribution of the spike train. Evidence for such code in the cortex exists (McClurkin et al., 1991), but is still tentative. Consequently, we assumed that the idealized neurons in our model encode information through their mean firing rates.

Computing the Estimator and the Associated Variance

Our goal was to find a way to combine the responses of disparity-selective neurons to recover either true disparity or perceived disparity. This was accomplished by creating networks from weighted linear combination of the curves shown in Figure 2 that give a mean output either to the solid line shown in Figure 7B, corresponding to perceived disparity, or to the broken line in Figure 7B, corresponding to geometrical disparity. The estimated disparity is

$$\Delta(\delta) = \sum_{i=1}^n w_i a_i(\delta), \quad (13)$$

where $\{w_i\}$ is the set of weights for a given estimator (geometrical or perceived), n is the number of disparity pools (41 in all cases), and $a_i(\delta)$ is the response of the disparity selective pool i to the disparity δ .

Instead of applying gradient descent, as before, to find the optimum set of weights in the network, a more efficient technique based on matrix inversion was used (see Appendix) (Casdagli, 1989); this was feasible because the network had only 41 weights, compared to 410 weights in the previous network. If all the pools of neurons were independent, the variance of the estimate is the weighted sum of variances:

$$\sigma_{\Delta(\delta)}^2 = \sum_{i=1}^n w_i^2 \sigma_i^2(\delta) = \sum_{i=1}^n w_i^2 \frac{K}{N_i} a_i(\delta), \quad (14)$$

where $\sigma_{\Delta(\delta)}^2$ is the variance of the estimator for disparity δ .

Consequently, if a pool of neurons contributes strongly to the estimation of disparity, $\Delta(\delta)$ (i.e., its weight w_i is high), it will also strongly influence the total variance, $\sigma_{\Delta(\delta)}^2$. Notice, too, that $\sigma_{\Delta(\delta)}^2$ is a function of δ since σ_i^2 are functions of δ .

Results

Variance of Estimation for Geometrical and Perceived Disparity

Figure 10 shows the standard deviation (square root of the variance) as a function of disparity for perceived (Fig. 10A) and geometrical (Fig. 10B) disparity for $N_0 = 1000$ and $\alpha^2 = 1.15$. With these values for the distribution in Equation 5, there is only a single neuron in the pools tuned to $+4^\circ$ and -4° , compared to 1000 neurons for the pool centered at 0° . On average, the standard deviation is about 20 times smaller ($20.97/0.92 = 22.74$) when perceived disparity is be-

ing computed than for the geometrical disparity. For perceived disparity, the large bias for large disparities is compensated by a low standard deviation, particularly in the central region, where the bias is also quite small. For geometric disparity, the bias of the estimation is zero but the standard error is high for all disparities.

An explanation for why the variance increases as the bias goes down derives from the sizes of the pools and the weights in the network. For large disparities, for example 4° , only five or six neuronal pools will be significantly active, the ones with peaks of disparity selectivity between 2° and 4° . The weights on these pools must be quite large to compute the true disparity and they will therefore make a large contribution to the total variance. To reduce the total variance, the weights from the noisiest pools must be reduced, and these are from the pools with the largest preferred disparities. The consequence is, however, a poor estimate for large values of disparity.

The trade-off between bias and variance arises from noise that is not homogeneous over the representation of disparity. More cells are tuned to zero disparity than to large values of disparity. We do not have enough data to decide whether vergence is similarly inhomogeneous, which is why we have restricted our analysis to disparity. If the noise is homogeneous over the whole vergence axis, then this would only scale the variance but would not change our conclusions. If more neurons were selective for some range of vergence, then an analysis similar to that of disparity would apply. For example, if grasping distances were overrepresented in the cortex, then estimates of distance within this range could be optimized at the expense of accuracy for more distant locations.

This result provides a tentative answer to the question of why we perceive distance the way we do when it would be theoretically possible to recover geometric distance. Estimating geometric distance would lead to an extremely high variance over all disparities. Fortunately, many other cues can be used to assess the distance of objects when disparity values are large, so our visual system can afford to optimize distance estimation for small disparities.

Sensitivity to Distribution of Neurons per Disparity Pool

Changing the values of N_0 and α^2 affects the quantitative predictions but not the conclusions. Fitting a gaussian to the data in LeVay and Voigt (1988) for anesthetized cats leads to a value of $\alpha^2 = 2.0$. Recordings from behaving monkeys rarely extend beyond the disparity range $[-1^\circ, 1^\circ]$, but the value obtained for α^2 in the cat probably provides an upper bound. For $\alpha^2 = 2.0$, the ratio of the standard deviation of the geometric disparity over the standard deviation of the perceived disparity was around 15, and around 30 for $\alpha^2 = 0.6$. There is also uncertainty about the sizes of the disparity pools. For $N_0 = 100$, the ratio was 15. Over all feasible values of N_0 , [50, 10,000], and α^2 , [0.5, 4.0], the ratios of the variances was at least 10.

Discussion

This study of how distance is coded in populations of cortical neurons has three conclusions. First, even though none of the gain-modulated units in the model was explicitly tuned to distance, together they formed a distributed representation from which egocentric distance could be estimated. This distributed representation can be described as a set of basis functions with the tuning curve of each neuron providing a single basis function. Second, a representation based on distance-tuned units is not consistent with the way humans perceive distance. It is therefore not surprising that distance-tuned neurons have not been reported in the cortex. Finally, perceived disparity corresponds to a low variance, though biased, estimate of disparity based on the known disparity-selective neurons observed in the visual cortex. This could account for the fact that we do not perceive geometrical distance even though, in principle, we should be able to compute it.

Posterior Parietal Cortex

Lesions in posterior parietal cortex lead to spatial deficits that include problems with using depth cues (Holmes, 1918; Holmes and Horrax, 1919; Godwin-Austen, 1965). Single-unit studies in LIP and area 7a of the posterior parietal cortex have revealed a distributed representation of egocentric space (see Andersen, 1989, and Stein, 1992, for reviews). These studies have focused on horizontal and vertical eye position and have shown that retinotopic receptive fields are modulated by eye position. Recent results suggest that the parietal cortex might also encode distance in a similar way. Gnadt has shown that many neurons in LIP respond to accommodation, disjunctive eye movements (Gnadt and Mays, 1991), and disparity. Among those neurons tuned to disparity, some are modulated by the distance to fixation in a manner similar to that found in area V1 (Gnadt and Mays, 1991). This finding is consistent with our model and strongly supports the hypothesis that the parietal cortex represents egocentric distance as part of a spatial representation.

The distributed representation of depth in this study is similar to the way that position is represented in head-centered space along the vertical and horizontal axes (Zipser and Andersen, 1988). Tuning for disparity is eye-centered and the responses are modulated by eye position, though not by conjugate eye movements as in previous models but by vergence eye movements. Neurons with response fields fixed in egocentric space, responding to a fixed distance or a fixed location in head-centered space independent of eye position, have not been reported in LIP or area 7a.

Early Egocentric Representation

Neurons in the foveal region of area V1 have small retinal receptive fields and are widely believed to encode the positions of objects in eye-centered coordinates. Our model, however, along with the data

from Trotter et al., challenges this belief and suggests instead that neurons in V1 encode the egocentric distance of objects with a distributed representation. The limited size of the receptive fields of V1 neurons might seem a serious problem, but a recent hierarchical model demonstrates that a "retino-spatiotopic" representation of spatial location is feasible (Pouget et al., 1993). By retino-spatiotopic we simply mean that cells encode the spatial location of an object falling on a limited region of the retina, just as many cells code orientation or color.

Further evidence for retino-spatiotopic maps was found in area 17 in cats (Weyand and Malpeli, 1989) and area V3A in monkeys (Galletti and Battaglini, 1989), where the visual response amplitudes of neurons are modulated by the vertical and horizontal eye position. This suggests that the egocentric position of an object is already encoded in the early visual cortex in a way similar to that found in the parietal cortex (Andersen et al., 1985; Zipser and Andersen, 1988). Such egocentric representations would be useful reference frames for planning actions.

Early spatial representations can be reconciled with physiological and psychophysical experiments claiming the contrary (Pouget et al., 1993). Lesion experiments in area V1 cannot be used as a guide since such lesions typically result in blindness, a nonspecific impairment that would certainly mask other more specific disorders. For example, specific deficits for orientation discrimination have been found following cortical damage in higher visual areas (Goodale and Milner, 1990) but not following lesions of V1, even though V1 represents orientation.

Do Gain-modulated Neurons Form a Basis Set?

A set of N vectors is called a basis for a vector space V of dimension N if (1) any arbitrary vector in V can be expressed as a linear combination of the vectors in the set and (2) the vectors in the set are linearly independent. The vector space used for function approximation has an infinite number of dimensions, so an infinite number of basis functions is required to span this space. Obviously, our basis set does not meet this requirement and therefore cannot be called a basis in the strict sense.

The functions we used were formed from products of gaussian functions of disparity and sigmoid functions of vergence. A set of gaussians with all possible widths and peak positions is known to form a basis. This is also the case for a set of sigmoids including all possible gains (temperature parameters) and threshold positions. If two sets individually form bases, the set made by the product of all possible pairs of functions from these two sets is also a basis. Consequently, the set of functions used here, products of gaussians and sigmoids, is a subset from a larger set which does form a basis.

In contrast, the set formed by all distance-tuned units does not form a basis. The response function of a distance-tuned unit is intrinsically a function of only one variable—distance. Consequently, this set cannot

form a basis for all possible functions of vergence and disparity, which depend on two variables. This is the reason why it was not possible to find a set of weights to approximate perceived distance using distance-tuned units.

Advantages of Basis Functions Representation over an Explicit Representation

Basis functions greatly simplify learning because only one layer of weights needs to be determined, those from the basis units to the output-units (Moody and Darken, 1989; Poggio, 1990). However, a serious problem with using basis functions is that many are required to estimate an arbitrary function accurately. In our simulations we used 410 basis functions, which is a large but manageable number. This number could have been reduced using an appropriate optimization procedure, but in a nonsystematic attempt to do so we found that at least 100 basis functions were needed to get reasonably accurate estimates of perceived distance.

In the brain, each basis function can be identified with a single neuron or a group of related neurons. We have assumed that each basis function contributes independently to the approximation. However, recent measurements from pairs of nearby cortical neurons in the visual system have shown that the noise in their responses typically has a correlation of 0.15 (Zohary et al., 1992; Gawne and Richmond, 1993). This correlation might seem low but in fact it severely limits the information that can be obtained by averaging over many neurons (Britten et al., 1992). A thousand correlated neurons carry little more information than around 50 neurons. A quantitative application of our model to the cortex awaits measurements of the correlations in the relevant neural populations carrying information about disparity and vergence.

We have shown that distance can be recovered from the responses of gain-modulated neurons; the same basis functions can be used to approximate other functions of vergence and disparity—perceived distance is only one function. Take, for example, the visually guided control of an arm. The angles between the arm segments required to reach a particular object are a different function of the object disparity and viewing distance and could be approximated by the same set of gain-modulated neurons used for estimating distance. Gain-modulated neurons could be used for many different purposes through multiple projections to different areas.

Even though we have shown that gain-modulated neurons in a distributed representation of disparity can represent egocentric distance, it would be misleading to call this a representation of egocentric distance since other functions could be equally well approximated by the same neural population. The actual functions explicitly represented are determined by the outputs of the neurons, or their projective fields, in addition to their receptive fields (Lehky and Sejnowski, 1990). The same argument applies as well to motor cortex, which could encode many func-

tions of reaching in addition to the direction of hand movement (Georgopoulos et al., 1986). In this view, the particular function attributed to the representation depends as much on the task selected by the experimenter as by the intrinsic parameters coded by the cortex (Poggio, 1990; Sanger, 1994).

Comparison to Backpropagation Networks

In a previous study, we trained a feedforward neural network to estimate distance by backpropagation of errors using the same input coding schemes for disparity and vergence studied here (Lehky et al., 1990). The weights between the inputs and the hidden layer of units were optimized along with the weights from the hidden to the output layer of units. Thus, the hidden units were not prespecified, as in our basis function network, but were optimized for the desired goal. With this approach it was possible to produce networks with nearly perfect performance using a minimal number of hidden units. In the case of geometrical distance from disparity and vergence, only 20 hidden units were required.

There are three disadvantages with this earlier model of distance estimation. First, each hidden unit took on a combination of properties whose function was difficult to decipher. This is a consequence of the compactness of the transformation produced by backpropagation and the complete connectivity between layers of units. Second, the backpropagation network took longer to optimize because two layers of weights were being simultaneously determined. Third, the representation in the hidden layer was highly specialized for the particular task the network was trained to perform and would not necessarily be useful for other purposes. If several functions of the same input units were required, it would be necessary to create a separate network for each function.

There are several advantages to having compact, dedicated representations, but it is an open question whether the cortex uses them. We suspect that cortical learning mechanisms produce more versatile basis functions that span the dimensions of the space within which functions are synthesized for each cortical area. In summary, a representation using basis functions requires more hidden units than a backpropagation network for any single task, but is more versatile and easier to train.

Extraretinal versus Retinal Cues for Viewing Distance

Vertical disparity could also be used to recover viewing distance (Longuet-Higgins, 1982; Mayhew and Longuet-Higgins, 1982) and recent evidence from humans shows conditions under which it can be used for absolute distance judgments and disparity scaling (Rogers and Bradshaw, 1993). However, vertical disparity is not used by the visual system for objects smaller than 20° of visual angle (Cumming et al., 1991; Rogers and Bradshaw, 1993), probably because the vertical disparity of smaller objects is too small for the visual system to detect. Subjects can, nonetheless, judge the distance or size of objects well below the

20° required for the vertical disparity system. In other experiments, manipulation of the vergence angle shows that it can influence the assessment of viewing distance (von Hofsten 1976, 1979; Ritter, 1977; Cumming et al., 1991). Therefore, both vergence and vertical disparity can be used when they provide reliable information.

This raises the issue of whether experimental evidence for vergence modulation could be attributed to vertical disparity instead (Gnadt and Mays, 1991; Trotter et al., 1991, 1992). The receptive fields of the neurons recorded by Trotter et al. were within the central 5° of the visual field (Y. Trotter, personal communication). Vertical disparity differences across small objects in this central region were, at most, 3'. It would seem highly unlikely that the gain modulation of the disparity-sensitive neurons observed by Trotter et al. could be fully accounted for by such small changes in vertical disparity. This remains, however, an empirical question whose resolution would require independent manipulations of vertical disparity and vergence.

Our model could be expanded to include vertical disparity in the input representation. In addition to vergence and disparity, we would expand the basis set to represent arbitrary functions of three variables. If the tuning of neurons in the visual cortex to vertical disparity were found to be a gaussian function or a sigmoid function, we would then be able to approximate any function of these three variables. The problem with continuing to increase the number of input variables in this way is that the number of required basis functions increases exponentially with the number of dimensions.

Distance versus Scaling

Vergence provides information for solving a problem that is closely related to distance estimation. When a rigid object moves away from a viewer, its size is usually perceived as constant even though its angular extent on the retina and the relative disparities between the parts of the object change with distance. The invariance of the perceived size of the object with viewing distance is called size constancy.

Several investigators have suggested that size constancy could be achieved by scaling the retinal size to compensate for viewing distance and have reported experimental results supporting this hypothesis (Cumming et al., 1991; Rogers and Bradshaw, 1993) (W. Richards, personal communication). The extent of an object in depth, S , as a function of the disparity of its front and back edges, respectively δ_{front} and δ_{back} , and vergence angle, ν , is given by

$$S = \frac{I}{2 \tan\left(\frac{\nu - \delta_{\text{back}}}{2}\right)} - \frac{I}{2 \tan\left(\frac{\nu - \delta_{\text{front}}}{2}\right)}, \quad (15)$$

where I is the interocular distance.

This equation is similar in form to Equation 1, which gives egocentric distance as a function of vergence and disparity. Thus, the neural mechanisms

underlying shape and size constancy may be similar to the ones described in this report. In particular, the modulation of disparity-selective neurons by vergence could be responsible for several perceptual constancies.

Appendix

We describe here two methods that were used to compute the optimal weights given a set of input units and the function to be approximated.

Estimating Distance from Basis Functions

Estimating geometrical or perceived distance with a linear combination of basis functions is equivalent to adjusting the hidden to output weights of a three-layer network whose hidden units correspond to the basis functions (Fig. 5). The delta rule provides a way of finding the optimum set of weights by iteratively presenting a set of examples of input-output pairs. The optimum weights are the ones that minimize the mean square error:

$$e = \frac{1}{2} \sum_{p=1}^m \left(y_p^* - \sum_{i=1}^n w_i B_i(\delta_p, \nu_p) \right)^2, \quad (15)$$

where m is the number of examples; y_p^* is the value of the target for an input vergence and disparity pair, (δ_p, ν_p) , w_i is the weight or coefficient associated with basis function i , and $B_i(\delta_p, \nu_p)$ is the value of the basis function i for example p . The corresponding values of the basis function, $B_i(\delta_p, \nu_p)$, were described in the first Methods section and the targets associated were either geometrical or perceived distance depending on the simulation.

The delta rule changes the weights by a small amount in the direction opposite to the gradient of increasing error:

$$\Delta w_i = -\alpha \frac{\partial e}{\partial w_i}, \quad (16)$$

$$\Delta w_i = \alpha \sum_{p=1}^m \left(y_p^* - \sum_{i=1}^n w_i B_i(\delta_p, \nu_p) \right) B_i(\delta_p, \nu_p), \quad (17)$$

where α is a parameter that controls the rate at which the weights change. The delta rule is guaranteed to converge to the minimum of the error after a sufficient number of presentations of the training set if α is sufficiently small (Widrow and Stearns, 1985).

Estimating Geometrical and Perceived Disparity

In the presence of noise, it is more convenient to treat the output of the network and the activity of the neuronal pools as random variables. Let Y^* be a random variable corresponding to the target disparities (either perceived or geometrical). Let Z be a vector of random variables whose components are the activities of the neuronal pools. Z is the sum of two vectors, $Z = A + N$, where A is the vector of the mean activities for a given disparity and N is the noise vector. Our goal is to find the linear estimation of Y^* from Z , $W^T Z$, which minimizes the variance of the error ($Y^* - W^T Z$):

$$W^* = \underset{W}{\operatorname{argmin}} [E[(Y^* - W^T Z)(Y^* - W^T Z)^T]] \quad (18)$$

The solution of this estimation problem is well known (Anderson and Moore, 1979) and has the form

$$W^* = C_{ZZ}^{-1} C_{ZY^*}, \quad (19)$$

where C_{ZZ} and C_{ZY^*} are, respectively, the covariance matrices of Z with itself and Z and Y^* . Substituting $Z = A + N$, we have

$$C_{ZZ} = E[(A + N)(A + N)^T], \quad (20)$$

$$C_{ZY^*} = E[AA^T + NN^T + AN^T + NA^T], \quad (21)$$

$$C_{ZZ} = C_{AA} + C_{NN}. \quad (22)$$

$C_{AN} = C_{NA} = 0$ because the noise around the mean is independent of the mean activity. At first, this statement may sound counterintuitive since the variance of the noise was chosen to be proportional to the mean activity (see Eq. 11). However, it can be easily shown that these two random variables are nevertheless independent. Similarly,

$$C_{ZY^*} = E[(A + N)Y^{*T}] = C_{AY^*}. \quad (23)$$

$C_{NY^*} = 0$ because N and Y^* are independent random variables. Substituting Equations 22 and 23 in Equation 19 leads to

$$W^* = [C_{AA} + C_{NN}]^{-1} C_{AY^*}. \quad (24)$$

The entries of the covariance matrices can be obtained from, for C_{AA} ,

$$c_{ij} = \frac{1}{m} \sum_{p=1}^m a_i(\delta_p) a_j(\delta_p), \quad (25)$$

and for C_{AY^*} ,

$$c_i = \frac{1}{m} \sum_{p=1}^m a_i(\delta_p) y(\delta_p), \quad (26)$$

where $a_i(\delta_p)$ is the activity of pool i for input p and $y(\delta_p)$ is the disparity, either perceived or geometrical, for input p .

The covariance matrix of the noise, C_{NN} , can be directly computed from the probability distribution of the noise. Nondiagonal terms are all zero since we assumed independence between neuronal pools. The diagonal terms are obtained from

$$\sigma_i^2 = \int_{-\infty}^{+\infty} n_i^2 p(n_i) dn_i. \quad (27)$$

Although we do not know the noise distribution, $p(n_i)$, directly, we have assumed that, given a mean level of activity of the pool, the noise has a normal distribution with a variance proportional to the mean (Eq. 11). Consequently, we need to introduce conditional distribution in Equation 27. In general, for any random variables, X and Y , we can write

$$p(x) = \int_{-\infty}^{+\infty} p_i(x) p(y) dy, \quad (28)$$

or equivalently, when Y is a discrete random variable,

$$p(x) = \sum_{p=1}^m p_V(x) P(y_p), \quad (29)$$

where $p(x)$ and $p(y)$ are the probability density functions of X and Y and $p_V(x)$ is the probability density function of X conditioned on Y .

Applying Equation 29 to Equation 27 leads to

$$\sigma_i^2 = \sum_{p=1}^m \left(\int_{-\infty}^{+\infty} n_i^2 p_{a_i(\delta_p)}(n_i) dn_i \right) \frac{1}{m}, \quad (30)$$

where $p_{a_i(\delta_p)}$ is the probability density of the noise conditional on $a_i(\delta_p)$, the mean activity of pool i for disparity δ_p .

The term inside the parentheses in Equation 30 is the conditional variance of the noise, which we know is proportional to the mean activity (Eq. 11):

$$\sigma_i^2(\delta_p) = \int_{-\infty}^{+\infty} n_i^2 p_{a_i(\delta_p)}(n_i) dn_i = \frac{K}{N_i} a_i(\delta_p); \quad (31)$$

therefore,

$$\sigma_i^2 = \frac{K}{N_i m} \sum_{p=1}^m a_i(\delta_p) = \frac{K}{N_i} \bar{a}. \quad (32)$$

Equation 24 provides a one-shot method for computing the optimal set of weights and can be easily implemented in standard symbolic mathematics programs such as *MATHEMATICA*, the one we used here. However, this method is impractical when the matrices have more than 100 dimensions. Matrix inversion was not used, for example, in the optimization problem posed in the first section of this appendix because the size of the matrix C_{AA} was 410×410 .

Notes

This research was supported by grants from the Office of Naval Research, the Howard Hughes Medical Institute, and the McDonnell-Pew Center for Cognitive Neuroscience in San Diego. We thank Drs. Yves Trotter, Simon Thorpe, Whitman Richards, and Sidney Lehky for valuable discussions.

Correspondence should be addressed to the authors at the Howard Hughes Medical Institute, Computational Neurobiology Laboratory, The Salk Institute, 10010 North Torrey Pines Road, La Jolla, CA 92037.

References

- Andersen RA (1989) Visual and eye movement functions of the posterior parietal cortex. *Annu Rev Neurosci* 12: 377-403.
- Andersen RA, Essick GK, Siegel RM (1985) Encoding of spatial location by posterior parietal neurons. *Science* 230:456-458.
- Anderson BDO, Moore JB (1979) Optimal filtering. Englewood Cliffs, NJ: Prentice-Hall.
- Britten KH, Shadlen MN, Newsome WT, Movshon JA (1992) The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J Neurosci* 12:4745-4765.
- Casdagli M (1989) Non-linear prediction of time series. *Physica* 35D:335-356.
- Cormack LK, Stevenson SB, Schor CM (1993) Disparity-tuned channels of the human visual system. *Vis Neurosci* 10:585-596.
- Cumming BG, Johnston EB, Parker AJ (1991) Vertical disparity and perception of three dimensional shape. *Nature* 349:411-413.
- Ferster D (1981) A comparison of binocular depth mechanisms in areas 17 and 18 of the cat visual cortex. *J Physiol (Lond)* 311:623-655.
- Foley JM (1980) Binocular distance perception. *Psychol Rev* 87:411-434.
- Foley JM, Held R (1972) Visually directed pointing as a function of target distance, direction and available cues. *Percept Psychophys* 12:263-268.
- Galletti C, Battaglini PP (1989) Gaze-dependent visual neurons in area V3A of monkey prestriate cortex. *J Neurosci* 9:1112-1125.
- Gawne TJ, Richmond BJ (1993) How independent are the messages carried by adjacent inferior temporal cortical neurons? *J Neurosci* 13:2758-2771.
- Georgopoulos A, Schwartz A, Kettner R (1986) Neuronal population coding of movement direction. *Science* 233: 1416-1419.
- Gnadt JW, Mays LE (1991) Depth tuning in area LIP by disparity and accommodative cues. *Soc Neurosci Abstr* 17:1113.
- Godwin-Austen RB (1965) A case of visual disorientation. *J Neurol Neurosurg Psychiatry* 28:448-453.
- Goodale MA, Milner AD (1990) Separate visual pathways for perception and action. *Trends Neurosci* 15:20-25.
- Holmes G (1918) Disturbances of visual orientation. *Br J Ophthalmol* 2:449-468.
- Holmes G, Horrax G (1919) Disturbances of spatial orientation and visual attention. *Arch Neurol Psychiatry* 1:385-407.
- Hornik K, Stinchcombe M, White H (1989) Multilayer feed-forward network are universal approximators. *Neural Networks* 2:359-366.
- Hubel D, Wiesel T (1970) Cells sensitive to binocular depth in area 18 of the macaque monkey cortex. *Nature* 225: 41-42.
- Lapedes A, Farber R (1988) How neural nets work. In: *Neural information processing system* (Anderson DN, ed), pp 442-456. New York: American Institute of Physics.
- Lehky SR, Sejnowski TJ (1990) Neural network model of visual cortex for determining surface curvature from images of shaded surfaces. *Proc R Soc Lond [Biol]* 240:251-278.
- Lehky SR, Pouget A, Sejnowski TJ (1990) Neural models of binocular depth perception. *Cold Spring Harbor Symp Quant Biol* 55:765-777.
- LeVay S, Voigt T (1988) Ocular dominance and disparity coding in cat visual cortex. *Vis Neurosci* 1:395-414.
- Longuet-Higgins HC (1982) The role of vertical dimension in stereoscopic vision. *Perception* 11:377-386.
- Maunsell JHR, van Essen DC (1983) Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. *J Neurophysiol* 49:1148-1167.
- Mayhew JEW, Longuet-Higgins HC (1982) A computational model of binocular depth perception. *Nature* 297:377-378.
- McClurkin JW, Optican LM, Richmond BJ, Gawne TJ (1991) Concurrent processing and complexity of temporally encoded neuronal messages in visual perception. *Science* 253:675-677.
- Moody J, Darken CJ (1989) Fast learning in networks of locally-tuned processing units. *Neural Comput* 1:281-294.
- Nikara T, Bishop PO, Pettigrew JD (1968) Analysis of retinal correspondence by studying receptive fields of binocular single units in cat striate cortex. *Exp Brain Res* 6:353-372.
- Pettigrew JD, Nikara T, Bishop PQ (1968) Binocular interaction on single units in cat striate cortex: simultaneous stimulation by single moving slit with receptive fields in correspondence. *Exp Brain Res* 6:391-410.
- Poggio G (1984) Processing of stereoscopic information in primate visual cortex. In: *Dynamic aspects of neocortical function* (Edelman GM, Gall WE, Cowan WM, eds), p. 613-635. New York: Wiley.

- Poggio GF, Fischer B (1977) Binocular interactions and depth sensitivity in striate and prestriate cortex of behaving rhesus monkeys. *J Neurophysiol* 40:1392-1405.
- Poggio GF, Motter BC, Squatrito S, Trotter Y (1985) Responses of neurons in visual cortex (V1 and V2) of the alert macaque to dynamic random-dot stereograms. *Vision Res* 25:397-406.
- Poggio GF, Gonzalez F, Krause F (1988) Stereoscopic mechanism in monkey visual cortex: binocular correlation and disparity selectivity. *J Neurosci* 8:4531-4550.
- Poggio T (1990) A theory of how the brain might work. *Cold Spring Harbor Symp Quant Biol* 55:899-910.
- Pouget A, Fisher SA, Sejnowski TJ (1993) Egocentric spatial representation in early vision. *J Cognit Neurosci* 5:150-161.
- Richards W (1971) Anomalous stereoscopic perception of depth. *J Opt Soc Am* 61:410-414.
- Richards W, Kaye MG (1974) Local versus global stereopsis. *Vision Res* 14:1345-1347.
- Ritter M (1977) Effect of disparity and viewing distance on perceived depth. *Percept Psychophys* 22:400-407.
- Rogers BJ, Bradshaw MF (1993) Vertical disparities, differential perspective and binocular stereopsis. *Nature* 361:253-255.
- Rumelhart DE, Hinton GE, Williams RJ (1986) Learning internal representations by error propagation. In: *Parallel distributed processing: explorations in the microstructure of cognition* (Rumelhart DE, McClelland JL, PDP Research Group, eds), Vol 1, Foundations. Cambridge, MA: MIT Press.
- Sanger T (1994) Theoretical consideration for the analysis of population coding in motor cortex. *Neural Comput* 6:29-37.
- Scott DW (1992) *Multivariate density estimation*. New York: Wiley Interscience.
- Stein J (1992) A distributed reference frame for egocentric space in the posterior parietal cortex. *Behav Brain Sci* 15(4):691-700.
- Stevenson SB, Cormack LK, Schor CM, Tyler CM (1992) Disparity-tuned mechanisms of human stereopsis. *Vision Res* 32:1685-1689.
- Tolhurst DJ, Movshon JA, Dean AD (1982) The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Res* 23:775-785.
- Trotter Y, Celebrini S, Thorpe SJ, Imbert M (1991) Modulation of stereoscopic processing in primate visual cortex V1 by the distance fixation. *Soc Neurosci Abstr* 17:1016.
- Trotter Y, Celebrini S, Stricanne B, Thorpe SJ, Imbert M (1992) Modulation of neural stereoscopic processing in primate area V1 by the viewing distance. *Science* 257:1279-1281.
- Vogels R, Spilleers W, Orban GA (1989) The response variability of striate cortical neurons in the behaving monkey. *Exp Brain Res* 77:432-436.
- von Hofsten C (1976) The role of convergence in space perception. *Vision Res* 16:193-198.
- von Hofsten C (1977) Binocular convergence as a determinant of reaching behavior in infancy. *Perception* 6:139-144.
- von Hofsten C (1979) Recalibration of the convergence system. *Perception* 8:37-42.
- Westheimer G, Tanzman I (1956) Qualitative depth localization with diplopic images. *J Opt Soc Am* 46:116-117.
- Weyand TG, Malpel JG (1989) Responses of neurons in primary visual cortex are influenced by eye position. *Soc Neurosci Abstr* 15.
- Widrow B, Stearns SD (1985) *Adaptive signal filtering*. Englewood Cliffs, NJ: Prentice Hall.
- Widrow G, Hoff ME (1960) Adaptive switching circuits. *IRE WESCON Convention Rec* 4:96-104.
- Zipser D, Andersen R (1988) Back propagation learning simulates response properties of a subset of posterior parietal neurons. *Nature* 331:679-684.
- Zohary E, Shadlen MN, Newsome WT (1992) Correlated activity of neuron in area MT. *Soc Neurosci Abstr* 18:1101.