A physiologically-based explanation of disparity attraction and repulsion

Samuel Mikaelian [1], Ning Qian *

Center for Neurobiology and Behavior, Columbia University, P.I. Annex Room 730, 722 W. 168th Street, New York, NY 10032, USA

Abstract

Westheimer and Levi [(1987) *Vision Research*, *27*, 1361–1368] found that when a few isolated features are viewed foveally, the perceived depth of a feature depends not only on its own disparity but also on those of its neighbors. The nature of this interaction is a function of the lateral separation between the features: When the distance is small the features appear to attract each other in depth but the interaction becomes repulsive at larger distances. Here we introduce a two-dimensional extension of our recent stereo model based on the physiological studies of Ohzawa, DeAngelis and Freeman [(1990) *Science*, *249*, 1037–1041] and demonstrate through analyses and simulations that these observations can be naturally explained without introducing ad hoc assumptions about the connectivity between disparity-tuned units. In particular, our model can explain the distance-dependent attraction/repulsion phenomena in both the vertical-line configuration used by Westheimer [(1986) *Journal for Neurophysiology*, *370*, 619–629], and the horizontal-line-and-point configuration used by Westheimer and Levi. Thus, the psychophysically observed disparity interaction may be viewed as a direct consequence of the known physiological organization of the binocular receptive fields. We also find that the transition distance at which the disparity interaction between features changes from attraction to repulsion is largely determined by the preferred spatial frequency and orientation distributions of the cells used in the disparity computation. This result may explain the observed variations of the transition distance among different subjects in the psychophysical experiments. Finally, our model can also reproduce the observed effect on the perceived disparity when the disparity magnitude of the neighboring features is changed. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: Disparity interaction; Depth illusion; Stereo vision; Orientation pooling; Spatial frequency pooling

1. Introduction

Many psychophysical experiments have demonstrated that nearby visual features interact with each other in generating our perception. Well-known examples include color constancy where interactions between neighboring color patches allow an observer to partially discount the spectral variations of the illumination (Land 1977), and motion repulsion where overlapping motion fields interact to create an exaggerated perceived angle between their directions of motion (Marshak & Sekuler, 1979). Such interactions among nearby visual features has also been documented in the domain of binocular disparity. In a series of psychophysical experiments, Westheimer (1986) and Westheimer and Levi (1987) showed that when the lateral separations (in the fronto-parallel plane) between the features are small, the depth separation between them appears smaller than the actual, as if the features were attracting each other in depth. As the separation increases, however, the strength of this attractive interaction decreases and finally it reverses sign to become repulsive. When the distance is very large there is no interaction between the features, as expected.

Lehky and Sejnowski (1990) have proposed a computational model to explain the disparity interaction described above. They started with a set of disparity tuning curves similar to those measured physiologically (Poggio & Fischer, 1977; Poggio & Poggio, 1984), and encoded disparity by population activities read off from these curves. The procedure was replicated to cover

---

* Corresponding author. Tel.: +1-212-5435213; fax: +1-212-5435161.

E-mail address: nq6@columbia.edu (N. Qian).

[1] Present address: Center for Neural Science, New York University, 4 Washington Place, New York, NY 10003, USA.

each spatial location. They found that the disparity attraction/repulsion phenomenon can be explained by implementing short-range excitatory and long-range inhibitory connections between disparity-tuned units at neighboring locations. We decided to model the same phenomenon at a lower level by starting with known binocular receptive field (RF) profiles instead of disparity tuning curves. This allowed us to consider the details of the stimulus configurations used by Westheimer (1986) and Westheimer and Levi (1987), which would not be possible for models starting with tuning curves. Specifically, we extended our recent stereo model for disparity computation (Qian, 1994; Zhu & Qian, 1996; Qian & Zhu, 1997a; Qian, 1997) to two spatial dimensions, and analyzed and simulated the psychophysical stimulus configurations. These models are based on known RF properties of binocular simple and complex cells in the cat primary visual cortex (Ohzawa, DeAngelis & Freeman, 1990; Freeman & Ohzawa, 1990; DeAngelis, Ohzawa & Freeman, 1991; Ohzawa, DeAngelis & Freeman, 1996, 1997). The population responses of complex cells at each location are used to compute the disparity for that location. Our results indicate that the reported distance-dependent disparity attraction and repulsion phenomenon can be naturally explained by our model without introducing ad hoc assumptions about the connectivity between disparity-tuned units at different locations. We also explored the parameters that influence the attraction-to-repulsion transition distance, and the consequences of changing the feature disparities at fixed separations. These results are contingent as to how the disparities reported by cell populations at different spatial scales are combined. A simple and plausible averaging scheme proves sufficient for our purposes. We also made new predictions that can be tested experimentally. Preliminary results have been reported previously in abstract forms (Mikaelian & Qian, 1997; Qian & Zhu, 1997b).
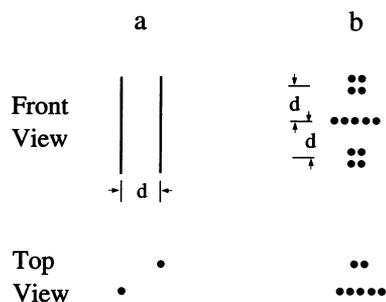

Fig. 1. Schematic stimulus configurations used in our analyses and simulations. (a) The front and top views of the vertical-line configuration containing two horizontally separated vertical lines. The left line is taken as the test line and right line as the inducing line. (b) The front and top views of the horizontal-line-and-point configuration consisting of 13 dots arranged into a horizontal line and two vertically separated points (squares of dots). The labels d in the frontal views indicate the lateral separations.

2. Methods

2.1. Stimulus configurations

2.1.1. Vertical-line configuration

We modeled disparity interactions with two very different stimulus patterns used in previous psychophysical experiments: the vertical-line configuration (Westheimer, 1986) and the horizontal-line-and-point configuration (Westheimer & Levi, 1987). The vertical-line pattern, also called the 1D pattern, consisted of two laterally separated vertical lines of equal luminance superimposed on a uniform background (Fig. 1a). The left line acts as the test line and the right line as the inducing line. The actual disparity of the test line is kept at zero, and its apparent disparity will be determined by the responses of the model cells (see below) centered on the line when the inducing line takes various lateral separations d and disparities D. Mathematically, the left and right images of this stereogram can be written as

$$I_l(x) = \delta(x) + \delta(x - d - D/2) + c \quad \text{and}$$

$$I_r(x) = \delta(x) + \delta(x - d + D/2) + c \quad (1)$$

where $\delta(\,)$ is the Dirac δ-function. Since an overall scaling factor for the images amounts to an overall scaling of the cell responses and does not change our results, we have set the intensity of the two lines to 1 and assigned an arbitrary intensity c to the background. We will compare the computed apparent disparities from our model with the perception determined from the psychophysical experiments (Westheimer, 1986).

Note that the actual experimental configuration used by Westheimer (1986) is slightly more complicated than our stimulus, for their practical purposes of making psychophysical measurements. It contained two sets of test and inducing lines located symmetrically around a central reference line. However, since the experimental results were reported to be similar for various large separations (>12 min) between the reference and the test lines, we did not include the reference line in our simulations to avoid undue mathematical complications. Eq. (1) also assumes infinite line lengths as a simplification.

2.1.2. Horizontal-line-and-point configuration

The horizontal-line-and-point configuration, also referred to as the 2D pattern, used in our simulation is identical to that used by Westheimer and Levi (1987) (see Fig. 1b). Although called a line with two points, the stimulus was actually composed of 13 individual dots — they used five equally spaced dots to represent the central horizontal 'line', and four dots arranged into a square to represent each of the two flanking 'points' (Fig. 1b). The line has a real disparity of zero and the flanking points have a real disparity of D. For

comparison with experimental data, the apparent disparity of the line is computed from the responses of model cells centered on it when the flanking points are at different vertical separations d from the line, and different disparities D. Mathematically, the left and right retinal images of the stereogram corresponding to this stimulus can be written as:

$$I_l(x, y) = I_0(x, y) + I_{1l}(x, y) + c \quad \text{and}$$

$$I_r(x, y) = I_0(x, y) + I_{1r}(x, y) + c \qquad (2)$$

where

$$I_0(\vec{r}) = \sum_{i=1}^{5} \delta(\vec{r} - \vec{X}_i), \quad I_{1l}(\vec{r}) = \varepsilon \sum_{i=1}^{2} \sum_{j=1}^{4} \delta(\vec{r} - \vec{\Delta}_{ij}^{(l)}),$$

$$I_{1r}(\vec{r}) = \varepsilon \sum_{i=1}^{2} \sum_{j=1}^{4} \delta(\vec{r} - \vec{\Delta}_{ij}^{(r)}), \qquad (3)$$

ε determines the luminance of the flanking points relative to the central line, and

$$\vec{r} = (x, y), \quad \vec{X}_i \equiv (x_i, 0), \quad \vec{\Delta}_{ij}^{(l)} \equiv (D_i, d_j),$$

$$\vec{\Delta}_{ij}^{(r)} \equiv (-D_i, d_j) \qquad (4)$$

$$\{x_i\} = \{-2, -1, 0, 1, 2\} \quad \text{for} \quad i = 1, ..., 5 \qquad (5)$$

$$\{d_j\} = \{d - 0.5, d + 0.5, -d - 0.5, -d + 0.5\} \quad \text{for}$$

$$j = 1, ..., 4 \qquad (6)$$

$$\{D_i\} = \{D/2 - 0.5, D/2 + 0.5\} \quad \text{for} \quad i = 1, 2 \qquad (7)$$

The units of all numerical values above are minutes of visual angle.

Westheimer and Levi (1987) also added a reference line solely for the purpose of making psychophysical measurements. We did not include it in our simulations because it was at a distance where its interaction with the test line would be negligible.

2.2. Models

We have applied three closely related stereo models of increasing levels of sophistication to the problem of disparity interaction. They are a one-dimensional (1D) stereo model, a primitive two-dimensional (2D) stereo model, and a fully elaborated 2D stereo model. Although our emphasis will be on the final elaborated model, we will also present results from the two simpler precursors because they are much easier to understand mathematically and help to paint an intuitive picture of how the models work.

2.2.1. 1D model

The details of the 1D stereo model have been presented elsewhere (Qian, 1994; Zhu & Qian, 1996; Qian, 1997; Qian & Zhu, 1997a). Briefly, the model is based on the quantitative physiological studies of Freeman and coworkers (Freeman & Ohzawa, 1990; Ohzzawa et al., 1990; DeAngelis et al., 1991), who found that the left and right RFs of a typical binocular simple cell in

the cat primary visual cortex can be described by two Gabor functions. When one is only interested in horizontal disparity computation and stimulus intensity mainly varies along the horizontal dimension, we may, as a first approximation, ignore the vertical dimension and use the following 1D Gabor filters to describe the left and right RF profiles of a binocular simple cell (centered at $x = 0$):

$$f_l(x) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{x^2}{2\sigma_1^2}\right)\cos(\omega_1 x + \phi_l) \quad \text{and}$$

$$f_r(x) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{x^2}{2\sigma_1^2}\right)\cos(\omega_1 x + \phi_r) \qquad (8)$$

where σ_1 and ω_1 are the Gaussian width and the preferred spatial frequency of the RFs, and ϕ_l and ϕ_r are the left and right phase parameters, respectively. The response of such a simple cell to a binocular stimulus with left and right retinal images $I_l(x)$ and $I_r(x)$ is determined according to (Ohzawa et al., 1990; Qian, 1994):

$$r_s = \int_{-\infty}^{+\infty} dx[f_l(x)I_l(x) + f_r(x)I_r(x)] \qquad (9)$$

It was found that the binocular simple cell response defined above is highly sensitive to the Fourier phases of input stimuli (Qian, 1994; Zhu & Qian, 1996). Simple cells are, therefore, not reliable disparity detectors (Ohzawa et al., 1990; Qian, 1994; Zhu & Qian, 1996). However, it has been shown (Qian, 1994) that an approximately phase-independent response can be obtained by summing the squared outputs of a quadrature pair of simple cells:

$$r_q = (r_{s,1})^2 + (r_{s,2})^2 \qquad (10)$$

where the two simple cells in the pair (labeled by the subscripts 1 and 2) have otherwise identical parameters except that their phases are related by

$$\phi_{l,2} = \phi_{l,1} + \pi/2 \quad \text{and} \quad \phi_{r,2} = \phi_{r,1} + \pi/2 \qquad (11)$$

These relations were first used by Freeman and coworkers to model responses of real binocular complex cells (Ohzawa et al., 1990), and are close relatives of the motion energy models (Adelson & Bergen, 1985; van Santen & Sperling, 1985; Watson & Ahumada, 1985).

A family of complex cells centered at a given location, all with the same preferred spatial frequency ω_1 and Gaussian width σ_1, but with different left–right phase-parameter differences

$$\phi_- \equiv \phi_l - \phi_r \qquad (12)$$

covering the range $[-\pi, \pi)$, can form a distributed representation of the stimulus disparity (Qian, 1994; Qian & Zhu, 1997a). The disparity of the stimulus can be estimated from this distributed representation according to

$$\tilde{D} = -\frac{\hat{\phi}_-}{\omega_1} \qquad (13)$$

where $\hat{\phi}_-$ is the phase-parameter difference of the most responsive cell in the family. In our previous studies (Qian, 1994; Qian & Zhu, 1997a), we considered a discrete set of cells in each simulation, and $\hat{\phi}_-$ was determined through a parabolic interpolation around the highest response to reduce the sampling error. In this study, however, because of the relative simplicity of the stimuli, we derived the closed-form, analytical expressions for $\hat{\phi}_-$, for both 1D and 2D models (see below and Appendix A). This is equivalent to an infinitely dense sampling of the cells and therefore no interpolation is necessary. Also note that Eq. (13) can be replaced by a population averaging model (Georgopoulos, Schwartz & Kettner, 1986). Very similar results should be obtained because the population response is typically well centered around the most responsive cell without strong skewing.

Thus, for a given stimulus, one can apply a family of complex cells all with the same preferred frequency ω_1 and bandwidth b (or equivalently, same ω_1 and σ_1) but with different phase-parameter differences, ϕ_-, to obtain a disparity estimation $\tilde{D}(\omega_1, b)$. Since real cells in the primary visual cortex have a wide range of preferred spatial scales (DeValois, Albrecht & Thorell, 1982; Shapley & Lennie, 1985), we will also compute the averaged estimation from all cell families with different ω_1 but the same b:

$$\bar{D}(b) = \int_0^\infty \tilde{D}(\omega_1, b) p(\omega_1) \, d\omega_1 \qquad (14)$$

where $p(\omega_1)$ is the probability density function for cells with differing ω_1s. We did not average across bandwidth because only cells with a relatively narrow bandwidth can be involved in disparity computation in the first place (Sanger, 1988; Qian, 1994; Qian & Zhu, 1997a). (The derivation of the algorithms relies on the assumption of small bandwidths.) Instead, for most simulations the bandwidth was fixed at a standard value of 1.5 octave; other values were also tried as indicated in Section 3. Since slight variations of $p(\omega_1)$ do not make qualitative difference in the results, we assume that it follows a normal distribution,

$$p(\omega_1) = \frac{1}{\sqrt{2\pi\sigma_{\omega_1}^2}} \exp\left[-\frac{(\omega_1 - \bar{\omega}_1)^2}{2\sigma_{\omega_1}^2}\right] \qquad (15)$$

where the parameter $\bar{\omega}_1$ represents the center of the preferred frequency distribution, and σ_{ω_1} determines the width of the distribution. The numerical values of these parameters used in our simulations were estimated from the reported macaque physiological data by DeValois et al. (1982). These investigators measured the spatial frequency tuning curves for foveal cells in the macaque striate cortex and determined from these measurements the distributions of the cells' preferred

frequencies and frequency tuning bandwidths. Based on their data, for our standard parameter set we let $\bar{\omega}_1$ and σ_{ω_1} of the preferred frequency distribution be 3.5 and 1.2 cycle/degree, respectively, and the bandwidth b was fixed at 1.5 octave. We also varied these parameters for many of our simulations as will be described in detail in Section 3.

It should be pointed out that in Eq. (14) we are not assuming that V1 cells perform frequency averaging. (Note that Eq. (14) is not an averaging of the cell responses.) Such an operation would render the cells insensitive to spatial frequency, contrary to experimental facts (DeValois et al., 1982). Instead, we propose a direct correspondence between the population activity of many families of cells tuned to different frequencies in the primary visual cortex and an overall percept given by Eq. (14). Alternatively, the averaging procedure could be explicitly performed at a stage beyond the striate cortex such as area MT.

We applied the above 1D model to explain the disparity interaction in the vertical-line configuration used by Westheimer (1986). This is reasonable because these patterns are essentially 1D with luminance variations along the horizontal axis only. However, to explain the results from the horizontal-line-and-point patterns used in Westheimer and Levi (1987) we have to use a 2D stereo model. We extend the 1D model to two spatial dimensions as follows.

2.2.2. Primitive 2D model

As a precursor to the elaborate 2D model, we first examine a simple extension of the above 1D algorithm. Specifically, we used the following 2D Gabor filters to describe the left and right RF profiles of a binocular simple cell:

$$f_l(x, y) = g(x, y) \cos(\omega_1 x + \omega_2 y + \phi_l) \quad \text{and}$$

$$f_r(x, y) = g(x, y) \cos(\omega_1 x + \omega_2 y + \phi_r) \qquad (16)$$

with

$$g(x, y) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left(-\frac{x^2}{2\sigma_1^2} - \frac{y^2}{2\sigma_2^2}\right) \qquad (17)$$

where the Gaussian widths σ_1 and σ_2 determine the RF dimensions along the x and y directions, and ω_1 and ω_2 are the horizontal and vertical components of the preferred spatial frequency of the cell, respectively. The ratio of the two preferred frequencies determines the preferred spatial orientation θ of the filters. And the ratio of the two Gaussian widths is the aspect ratio of the simple cell RFs. We used an aspect ratio of 1.7 as the standard parameter in our simulations. The subsequent steps of quadrature-pair construction for modeling complex cell responses, and of estimating apparent disparity by locating the maximum response among a population of complex cells, are identical to the 1D case. (See Appendix A for a more thorough discussion.)

However, the apparent disparity at this stage is a function of both ω_1 and ω_2, or equivalently, a function of total preferred frequency ω and orientation θ where

$$\omega = \sqrt{\omega_1^2 + \omega_2^2} \tag{18}$$

and

$$\tan \theta = \frac{\omega_2}{\omega_1} \tag{19}$$

so that $\theta = 0°$ corresponds to a vertical and $\theta = 90°$ to a horizontal preferred orientation. At this point, we may add a final pooling of disparities reported at various spatial scales and orientations. However, since this primitive version is only meant to provide insight for understanding the full elaborated version, we will omit the final pooling step.

2.2.3. Elaborated 2D model

The primitive 2D model is so called because, according to Eq. (16), the principal axes of the Gaussian envelopes are always along the horizontal and vertical directions regardless of the preferred orientations of the RFs. In reality, for simple cells the long axis of the envelope aligns with the preferred orientation of the cell (Jones & Palmer, 1987; Ohzawa et al., 1996; Anzai, Ohzawa & Freeman, 1999). This fact is taken into account in this fully elaborated 2D stereo model. Specifically, for the simple cell RFs, we replaced the $g(x, y)$ function in Eq. (16) by

$$g(x, y) = \frac{\sqrt{\det(\boldsymbol{\Sigma})}}{2\pi} \exp\left(-\frac{\mathbf{r}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{r}}{2} \right) \tag{20}$$

where $\mathbf{r}^{\mathrm{T}} \equiv (x \quad y)$ and

$$\boldsymbol{\Sigma} \equiv \begin{pmatrix} 1/\sigma_1^2 & 1/\sigma_{12} \\ 1/\sigma_{12} & 1/\sigma_2^2 \end{pmatrix} \tag{21}$$

To ensure that the long axis of the Gaussian envelope aligns with the preferred orientation θ of the cell, we require that the $\boldsymbol{\Sigma}$ matrix be transformed into the diagonal form

$$\boldsymbol{\Sigma}' \equiv \begin{pmatrix} 1/\sigma_1'^2 & 0 \\ 0 & 1/\sigma_2'^2 \end{pmatrix} \tag{22}$$

after a rotation of angle θ.

We used the same procedure for computing the simple cell and quadrature-pair responses as for the primitive 2D model. However, for the elaborated version we added one more step for obtaining complex cell responses. Previously, we pointed out that in order to account for the fact that real complex cell RFs are on average larger than those of simple cells, a spatial pooling step should be added after the quadrature-pair construction to simulate complex cell responses:

$$r_{\mathrm{c}} = r_{\mathrm{q}} * W \tag{23}$$

where W is a spatial weighting function, and $*$ denotes the spatial convolution operation (Qian & Zhu, 1995; Zhu & Qian, 1996; Qian & Zhu, 1997a). We showed that the pooling step is computationally important for certain types of stimuli such as random dot patterns because it greatly improves the phase-independence of the complex cell responses and consequently makes their disparity tuning curves much more reliable (Qian & Zhu, 1995; Zhu & Qian, 1996; Qian & Zhu, 1997a). Eq. (23) is the final complex cell response to be used in our simulations of the elaborated model. We have ignored the spatial pooling step in the 1D and the primitive 2D models for simplicity.

The final pooling of apparent disparities reported by different families of cells are similar to the 1D case, except that in addition to spatial scales, we now have to also pool over different orientations:

$$\bar{D}(b) = \int_{-\pi/2}^{\pi/2} \int_0^\infty \tilde{D}(\omega, \theta, b) p(\omega) p'(\theta) \, \mathrm{d}\omega \, \mathrm{d}\theta \tag{24}$$

where $p'(\theta)$ is the weighting function for the orientation pooling. Two factors should be considered when choosing the form of $p'(\theta)$. The first is the fact that more vertically oriented cells are better suited for horizontal disparity detection. $p'(\theta)$ should, therefore, be biased towards the vertical orientation. The second factor is that disparity estimations from the more active cell families should be weighted more in the pooling (Georgopoulos et al., 1986). Therefore, for vertical line patterns that activate vertically oriented cells more strongly, $p'(\theta)$ should be biased towards the vertical orientation; similarly, for horizontal line patterns $p'(\theta)$ should be biased towards the horizontal orientation. We used a Gaussian function centered at an orientation θ_0 and with a width σ_θ in our simulations. The values of the parameters were varied as explained in Section 3. (The broad-band nature of the patterns used in this paper renders it unnecessary to consider a similar stimulus dependence of $p(\omega)$.) Detailed derivations of how various stages of the elaborated model respond to the vertical-line and the horizontal-line-and-point stimuli are provided in Appendix A.

3. Results

3.1. 1D model

We first consider the simplest case by applying the 1D stereo model to the vertical-line configuration. We consider both $\tilde{D}$, the apparent disparity of the test line estimated from a single family of cells with a given preferred horizontal spatial frequency ω_1 whose RFs are centered on the test line, and $\bar{D}$, the averaged
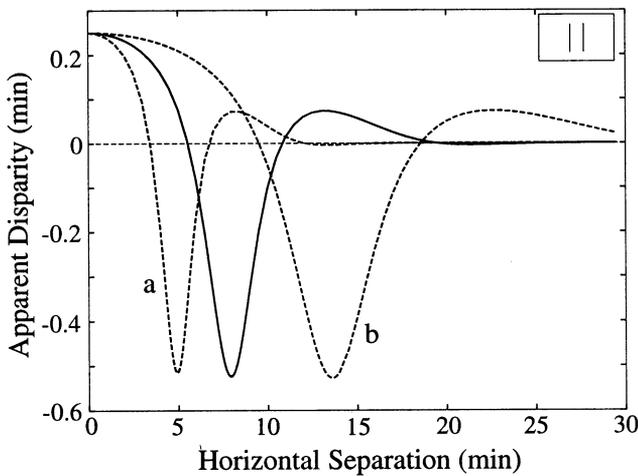
Fig. 2. 1D model applied to the vertical-line configuration. The apparent disparity of the test line, calculated from a single complex cell family with a given preferred spatial frequency and with RFs centered on the test line, is plotted as a function of the separation between the test and the inducing lines. The actual disparities of the test and the inducing lines are fixed at 0 and 0.5 min, respectively. Positive apparent disparities indicate attractive interaction while negative values indicate repulsion between the two lines. The three curves correspond to three different cell families with the preferred spatial frequencies equal to 5.6 (curve *a*), 3.5 (continuous curve) and 2.1 (curve *b*) cycle/degree. The cells in all three families have the same frequency bandwidth of 1.5 octave.

estimation from all cell families tuned to different scales and centered on the line. Since the actual disparities of the test and inducing lines are zero and D, respectively, if $\tilde{D}$ or $\bar{D}$ has the same sign as D, that means the apparent disparity of the test line is shifted towards the disparity of the inducing line indicating attractive interaction. On the other hand, an opposite sign between $\tilde{D}$ (or $\bar{D}$) and D indicates repulsive interaction.

For the special case of two equi-luminous lines on a dark background ($c = 0$), the conditions in the experiments of Westheimer (1986), $\tilde{D}$ is given by Eq. (40) in Appendix A. By averaging $\tilde{D}$ across all cells with differing ω_1s according to Eq. (14), we can obtain an expression for $\bar{D}$. This averaging will be performed numerically below. We now consider the dependence of the computed apparent disparities ($\tilde{D}$ and $\bar{D}$) of the test line on the distance d and disparity D of the inducing line.

The most prominent aspect of the disparity interaction revealed by Westheimer's experiments is the existence of attraction and repulsion zones as the distance d between the test and inducing lines is varied (Westheimer, 1986). First we examine whether the apparent disparity $\tilde{D}$ estimated from a single cell family can explain this experimental result. In the limit of large separation $d \gg \sigma_1$, the right hand side of Eq. (40) vanishes:

$$\lim_{d \to \infty} \tilde{D} = 0. \tag{25}$$

That is, the computed apparent disparity of the test line approaches its actual disparity when the separation between the test and inducing lines is large, consistent with the psychophysical experiments. This, of course, is due to the inducing line leaving the RF centered on the test line.

To simplify the analysis for the other d values, we may consider the limits

$$D \ll \sigma_1 \quad \text{and} \quad \omega_1 D \ll 1 \tag{26}$$

This is justified as the experimental value of D was set at 0.5 min and the cell population we are considering is centered at $\omega_1 = 3.5$ cycle/degree with a bandwidth of 1.5 octave. In this small-D limit, Eq. (40) can be expanded:

$$\lim_{D \to 0} \tilde{D}$$
$$= \frac{D}{\omega_1} \frac{e^{-d^2/2\sigma_1^2}\left[e^{-d^2/2\sigma_1^2}\omega_1 + \omega_1 \cos(\omega_1 d) - \frac{d}{\sigma_1^2}\sin(\omega_1 d)\right]}{1 + e^{-d^2/2\sigma_1^2}[e^{-d^2/2\sigma_1^2} + 2\cos(\omega_1 d)]} \tag{27}$$

At very small separations d, the above expression yields

$$\lim_{d, D \to 0} \tilde{D} = \frac{D}{2} \tag{28}$$

Thus, at small separation and inducing disparity, the apparent disparity of the test line is shifted half way towards the disparity D of the inducing line, indicating attractive interaction as observed experimentally. Intuitively, when both d and D are small, the images of the two lines on each retina are very close and consequently they should fall in the same excitatory/inhibitory subregions of the vertically-oriented simple cell RFs and contribute equally to the responses. It is thus not surprising that the disparities of the two lines are averaged at the complex cell stage which receives input from the simple cells. On the other hand, for sufficiently large values of the separation d, the denominator in Eq. (27) will stay positive while the numerator will oscillate in sign. There will, therefore, be intermediate regions where the interaction is repulsive.

To see the overall behavior of the apparent disparity $\tilde{D}$ estimated from a single family of complex cells, Eq. (40) is plotted in Fig. 2 as a function of the separation d between the two lines under three different values of preferred spatial frequency ω_1, while all other parameters are fixed. In these plots, the inducing disparity D is equal to 0.5 min of arc. The three curves in Fig. 2 correspond to three cell families which have preferred spatial frequencies of 2.1, 3.5, and 5.6 cycle/degree respectively and all have a bandwidth b of 1.5 octave. The figure clearly shows that for all three cell families, the interaction is attractive at small distance, and as the distance increases, the interaction oscillates between attraction and repulsion. When the separation is very large the interaction vanishes. The distance at which the

transition between attraction and repulsion occurs increases as the preferred spatial frequency of the cell family decreases.

These results reveal that the apparent disparity $\tilde{D}$ obtained from a single family of complex cells at a certain spatial scale exhibits distance-dependent disparity attraction and repulsion phenomenon. However, $\tilde{D}$ reverses its sign many times before it decays to zero while our perception reverses sign only once according to the experiments (Westheimer, 1986). (The multiple sign reversals are not specific to the particular choice of the Gabor filters, but rather, are a consequence of the multiple excitatory and inhibitory subregions in the underlying simple cells' RFs.) We now show that this discrepancy can be resolved by assuming that our perception corresponds to $\bar{D}$ of Eq. (14), the averaged estimation over all scales at a given location. It is easy to see that the behavior of $\bar{D}$ in the limiting cases of $d \to \infty$ and $d = 0$ are exactly same as that for $\tilde{D}$ shown in Eqs. (25) and (28) since under these two extremes, $\tilde{D}$ is not a function of ω_1.

To see the overall pattern of $\bar{D}$, we numerically integrated Eq. (14) and plotted the results in Fig. 3. The three curves in the figure correspond to preferred spatial frequency distributions centered at 2.1, 3.5, and 5.6 cycle/degree, from right to left respectively. The Gaussian widths of these frequency distributions are equal to 1.0, 1.2, and 1.9 cycle/degree, respectively. The continuous curve in the figure corresponds to the standard set
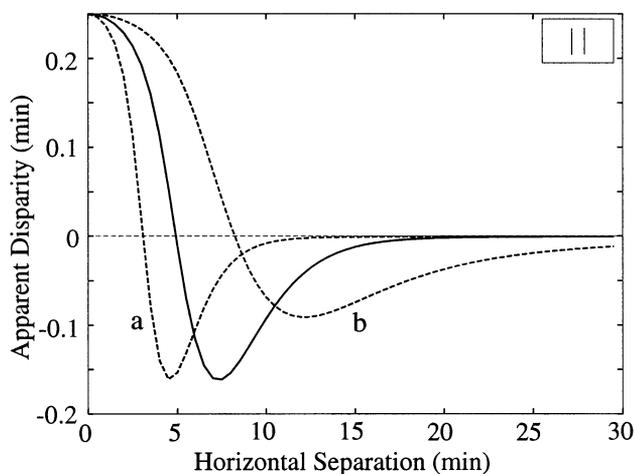

Fig. 3. 1D model applied to the vertical-line configuration with spatial-frequency pooling. The apparent disparity of the test line, averaged across complex cell families with different preferred spatial frequencies and with RFs centered on the test line, is plotted as a function of the separation between the test and the inducing lines. The actual disparities of the test and the inducing lines are fixed at 0 and 0.5 min, respectively. The bandwidth is 1.5 octave in all cases. The three curves were obtained from three different preferred spatial-frequency distributions of the cell population, with the distribution centers located at 5.6 (curve *a*), 3.5 (continuous curve) and 2.1 (curve *b*) cycle/degree, and corresponding Gaussian widths of 1.9, 1.2 and 1.0 cycle/degree.

of parameters based on the physiological data of De-Valois et al. (1982) (see Section 2). It is clear from Fig. 3 that with the averaging across the whole cell population there is only one transition from attraction to repulsion, closely resembling the main features of the psychophysical results (Westheimer, 1986).

Similar to Fig. 2, Fig. 3 shows that the distance at which this single transition occurs (the transition distance) increases with a decrease of the population's spatial-frequency distribution, characterized by the center of the distribution $\bar{\omega}_1$. This feature provides a possible explanation of the experimental observation that different subjects in the psychophysical experiments have different transition distances (see Section 4). Varying the other parameter of the 1D model, the Gaussian width of the frequency distribution, can account for variations in the strength of the repulsion effect among the subjects, as demonstrated in curve *b* of Fig. 3. We also found that in contrast to the above frequency pooling closely based on the physiologically obtained distribution (DeValois et al., 1982), a uniform pooling across all frequencies essentially wipes out the repulsion zones in the curves.

In the above simulations, we performed frequency pooling across cell families at different scales after extracting the apparent disparity at each scale. Alternatively one could first average complex cells' responses from different scales and then extract the apparent disparity (Fleet, Wagner & Heeger, 1996). The result using the standard set of parameter is shown as the continuous curve in Fig. 4. It does not fit the experimental data well because the repulsion occurred over too narrow a range of separation and the attractive tail at large distances was not observed. The curves labeled *a*, *b* and *c* were obtained by increasing the bandwidth, spatial frequency, and spatial frequency pooling width, respectively, from the standard set of parameters, as described in the caption. Like the continuous curve, none of these curves fit the psychophysical observations very well. It is possible to get a curve similar to the experimental observations by choosing parameter values significantly away from the standard set (curve *d* in the figure). However, those parameter values do not allow accurate disparity estimations when the input stimulus has a single, uniform disparity because the cells' bandwidth is too large.

3.2. Primitive 2D model

As mentioned earlier, since the horizontal-line-and-point patterns used by Westheimer and Levi (1987) are 2D, it can only be studied with a 2D model. In this subsection we demonstrate that the primitive 2D model described in Section 2 can be used to explain the illusion. We will take advantage of this simple model to provide a mathematically intuitive understanding of the
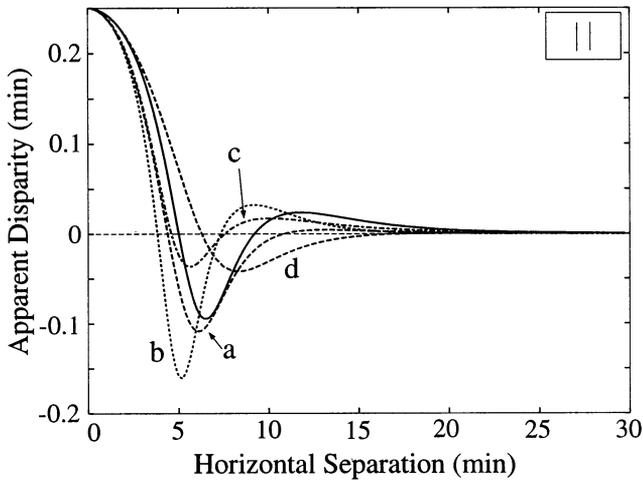
Fig. 4. 1D model applied to the vertical-line configuration using the frequency pooling at the level of complex cell responses (Fleet et al., 1996). The apparent disparity of the test line, calculated from spatial-frequency-averaged responses of cells centered on the test line, is plotted as a function of the separation between the test and the inducing lines. The actual disparities of the test and the inducing lines are fixed at 0 and 0.5 min, respectively. For the continuous curve, the central value and Gaussian width of the spatial-frequency pooling are 3.5 and 1.2 cycle/degree, respectively, and the bandwidth of all cells is 1.5 octave. For curves a–c, one (or a few) of the above parameters is changed while others are held constant, as follows. (a) The bandwidth is increased to 2.0 octave. (b) The central spatial frequency is increased to 5.0 cycle/degree. (c) The Gaussian width is increased to 1.9 cycle/degree. (d) The central spatial frequency is lowered to 2.0 cycle/degree, the Gaussian width raised to 1.4 cycle/degree and the bandwidth increased to 2.0 octave.

explanation. A more detailed explanation with a fully elaborated 2D model will be presented in the next subsection.

The primitive 2D model contains two new parameters, the RF Gaussian width σ_2 and the preferred spatial frequency ω_2 along the vertical dimension. The Appendix A shows that for the vertical-line configuration, the primitive 2D model yields a $\tilde{D}$ that is independent of σ_2 and ω_2. Hence, the results obtained for $\tilde{D}$ in the 1D model carry over to this primitive 2D model. As this model applied to the vertical-line configuration is a special case of the elaborated model (see the last section of the Appendix A), the effect of pooling $\tilde{D}$ across orientations as well as frequencies can be deduced from the results of the elaborated model. We would like to note the behavior of $\tilde{D}$ at very large RF sizes. For the 1D pattern, Eq. (40) gives

$$\lim_{\sigma_1 \to \infty} \tilde{D}(\omega_1) = \frac{D}{2} \tag{29}$$

independent of other parameters including d. This means that there will be a constant attraction at all d if infinitely large RFs are used. Therefore, for large enough RF size, the transition distance must increase with the size. As shown below, for the 2D pattern the behavior is different.

Eq. (41) of Appendix A describes the behavior of the estimated disparity $\tilde{D}$ from a single family of cells for the horizontal-line-and-point configuration. The term proportional to c in the expression for κ defined by Eq. (42) is the contribution of the uniform background whereas the other terms in that expression are contributions of the zero-disparity middle line. As with the 1D case, the apparent disparity $\tilde{D}$ of the test line approaches zero, its true disparity, with increasing separation from the inducing points. The transition points from the attraction to the repulsion zone can be identified from the zeroes of Eq. (41). The main configuration used by Westheimer and Levi (1987) can be simulated with the choices

$$D = 0.5 \text{ min}, \quad c = 0 \quad \text{and} \quad \varepsilon = 1 \tag{30}$$

for the stimuli. For separations d that are larger than a couple of minutes and with our standard choice of $\omega = 3.5$ cycle/degree, the zeroes of Eq. (41) are approximately given by:

$$\cos(\omega_2 d) \approx 0 \Rightarrow d \geq \frac{\pi}{2\omega_2} > \frac{\pi}{2\omega} = 4 \text{ (min)}. \tag{31}$$

Just as in the case of the vertical-line pattern, the estimated disparity $\tilde{D}$ oscillates between attraction and repulsion several times and the separation d corresponding to the first transition distance increases with decreasing ω. Here, though, the transition distance is inversely related to ω_2, whereas for the vertical-line case, it is inversely related to ω_1. This difference is caused by the fact that the lateral separations between elements of different disparities are along the vertical and horizontal dimensions, respectively, for the two cases. Moreover, Eq. (31) implies that the transition points are to a good approximation independent of the RF sizes, σ_1 and σ_2, in contrast to the case of the 1D pattern in Eq. (29). Results of pooling across cell families at different spatial scales would be very similar to the 1D case but we will leave the pooling steps to the fully elaborated model to be described next.

3.3. Elaborated 2D model

While the above primitive 2D model provides analytical insights into how disparity attraction and repulsion can arise in both the vertical-line and the horizontal-line-and-point configurations, we performed most of the simulations with the fully elaborated model because it is closest to physiological data (see Section 2). Eqs. (50) and (51) in Appendix A describe the estimated disparity $\tilde{D}$ from a single family of cells with a given preferred spatial frequency and preferred orientation for the horizontal-line-and-point configuration, according to the elaborated model. Simulation results based on these equations are shown in Fig. 5. Similarly to Fig. 2, although the curves show the basic attraction/re-

pulsion effect, they do not fit with the experimental data too well because of the exaggerated repulsion strength and the tail attraction at large separations. The strong repulsion is governed by our choice of preferred orientation, θ = 45° in this case. As indicated by Eq. (31) for the primitive model, for more horizontally-oriented filters the transition occurs at a smaller separation *d*, where the Gaussian falloff of the RF is not as significant as at larger separations, resulting in a more pronounced repulsion. As before, the exaggerated-repulsion and the attractive-tail problems are alleviated by pooling over populations of cells tuned to different spatial frequencies, and additionally, different orientations. (Here, as with the 1D and primitive 2D models, a non-physiological uniform pooling over spatial frequencies will not yield the correct behavior.) The repulsion is particularly reduced by considering an orientation distribution broadly centered around the vertical because vertically oriented cells should contribute more to horizontal disparity computation (DeAngelis et al., 1991). The results are shown in Fig. 6 where spatial pooling of quadrature pairs is implemented as well (see Eq. (23)). The continuous curve, in good agreement with psychophysical data, corresponds to the standard set of parameters.

The four dashed curves in the figure were obtained by changing one parameter at a time from the continuous curve while keeping all other parameters the same
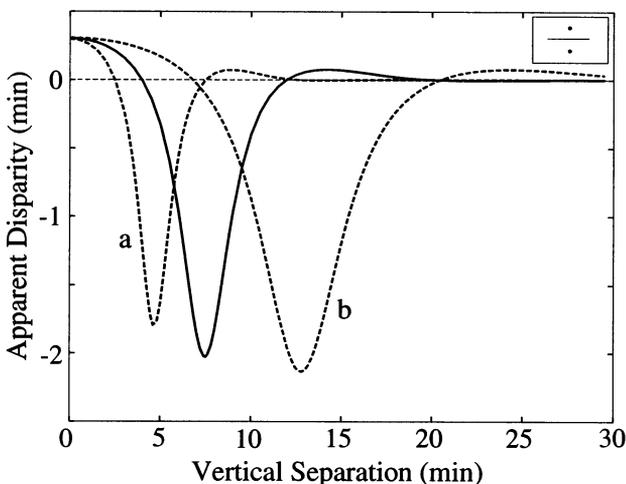

Fig. 5. Elaborated model applied to the horizontal-line-and-point configuration. The apparent disparity of the middle line, calculated from a single complex cell family with a given preferred spatial frequency and orientation and with RFs centered on the test line, is plotted as a function of the vertical separation between the line and the flanking points. No pooling is applied. The actual disparities of the line and the flanking points are fixed at 0 and 0.5 min, respectively. The three curves correspond to three different cell families with preferred spatial frequencies equal to 5.6 (curve *a*), 3.5 (continuous curve) and 2.1 (curve *b*) cycle/degree. The cells in all three families have the same frequency bandwidth of 1.5 octave, preferred orientation of 45° and RF aspect ratio of 1.7.
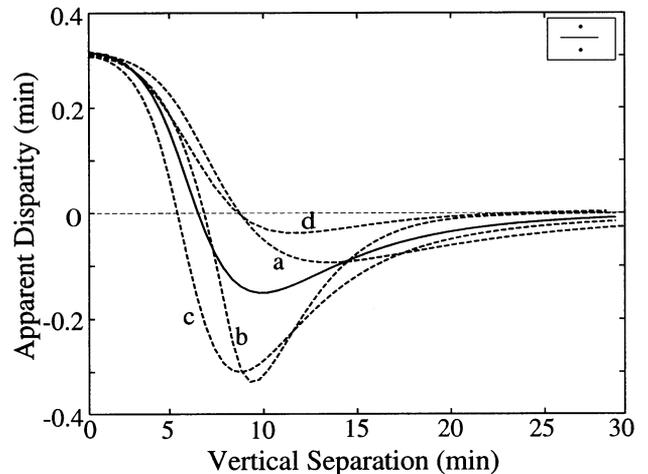

Fig. 6. Elaborated model applied to the horizontal-line-and-point configuration with pooling. The apparent disparity of the middle line, averaged across all cell families with different preferred spatial frequencies and orientations and with RFs centered at the middle of the line, is plotted as a function of the vertical separation between the line and the flanking points. The actual disparities of the line and the flanking points are fixed at 0 and 0.5 min, respectively. The continuous curve corresponds to the standard parameter set defined as follows: the spatial pooling of the quadrature pairs is centered on the middle of the line with the Gaussian width along each cell axis (η'_1 and η'_2 in Eq. (46)) equal to 0.25 times the corresponding Gaussian width of the underlying simple cells, the preferred spatial frequency distribution is centered at 3.5 cycle/degree with a Gaussian width of 1.2 cycle/degree, the preferred orientation distribution covers the range [−90°, 90°] and is centered at the vertical (0°) with a Gaussian width of 45°, the frequency bandwidth is 1.5 octave and the aspect ratio is 1.7 for all cells. Each dashed curve is obtained by varying one of the above parameters from the standard value while holding all the others fixed, as follows: (a) The central value of the spatial-frequency distribution is decreased to 2.5 cycle/degree. (b) No pooling of disparities across spatial frequency is performed. (c) A uniform distribution of the preferred orientations is assumed here. (d) The Gaussian width of the spatial pooling along each cell axis is increased to 0.5 times the corresponding Gaussian width of the underlying simple cells.

(see figure caption), for demonstrating the robustness of the results. Curve *a* is obtained by shifting the central value of the frequency distribution from the standard value of 3.5 to 2.5 cycle/degree. The rightward shift of the transition point is in agreement with Eq. (31) for the primitive 2D model, and the trend in Fig. 5. The remaining three curves show the effects of the various pooling schemes employed in the model. Curve *b* is obtained by turning off spatial frequency pooling. Since now only the orientation of the cells in the pool is varied, all cells have the same RF size and preferred spatial frequency. The repulsion is stronger under this condition because of the reduced spread of $\tilde{D}$ from individual cell families. Curve *c* is obtained by broadening the orientation pooling to a uniform distribution. The resulting leftward shift of the transition distance and strengthening of the repulsion are caused primarily by the involvement of more horizontally tilted filters.

Curve *d* corresponds to increasing the extent of spatial pooling. Consequently, the flanking points tend to be weighted more in computing complex cell responses, biasing apparent disparities towards the disparity of these points.

In order to check the consistency of our model, we also applied the elaborated 2D model to the vertical-line configuration. The simulation results are presented in Fig. 7, based on the expressions derived in Appendix A (Eqs. (52) and (53)). As expected, the 2D model works well on the 1D pattern, as demonstrated by the continuous curve, corresponding to the standard parameter values. Notice that the asymptotic approach of this curve to zero at large separations is slower than the continuous curve in Fig. 3 and, therefore, more in accord with the results of Westheimer (1986). The dashed curves exhibit the effects of varying parameters from the standard values as detailed in the figure caption. Curve *a* implements the orientation pooling with the Gaussian width halved, i.e. with RFs more narrowly centered around the vertical. Including more vertically oriented filters at a fixed spatial frequency effectively increases the horizontal component of the frequency, ω_1, thereby shifting the curve to the left. Unlike the horizontal-line-and-point configuration in Fig. 6, the repulsion for the vertical-line configuration
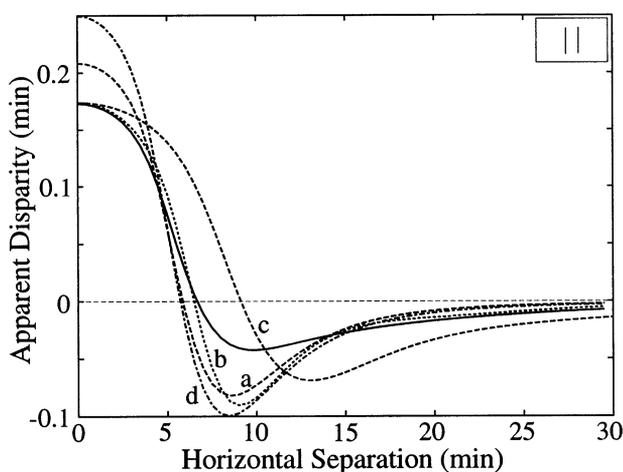

Fig. 7. Elaborated model applied to the vertical-line configuration with pooling. The apparent disparity of the test line, averaged across all cell families with different preferred spatial frequencies and orientations and with RFs centered on the test line, is plotted as a function of the separation between the test and the inducing lines. The actual disparities of the test and the inducing lines are fixed at 0 and 0.5 min, respectively. The continuous curve corresponds to standard parameter values as given for Fig. 6. The dashed curves are obtained by changing one or two parameter values from the standard ones, as follows. (a) The Gaussian width of the orientation pooling is reduced to 22.9°. (b) The Gaussian width of the frequency pooling is reduced to 0.6 cycle/degree. (c) The center of the spatial-frequency distribution is reduced to 2.5 cycle/degree and the corresponding Gaussian width is reduced to 0.6 cycle/degree. (d) The aspect ratio is reduced to 1.0.

increases with more vertically oriented cells in the population because the two lines are separated horizontally, and thus their interactions can be better sensed by cells oriented more vertically. Curve *b* is obtained by reducing the Gaussian width of the frequency pooling by half, resulting mainly in stronger repulsion. This is because of the reduced spread of $\tilde{D}$ from individual cell families, similar to curve *b* in Fig. 6. The rightward shift of the transition distance in curve *c* is expected as it is obtained by decreasing the central value of the spatial frequency used to generate curve *b*. Curve *d* corresponds to a reduction of the aspect ratio to 1. This effectively increases ω_o per Eq. (53) and thus decreases the resulting transition distance.

The above discussions concentrate on the distance (*d*) dependence of disparity interaction while the inducing disparity (*D*) is kept constant. We now investigate the dependence of disparity interaction on the inducing disparity at fixed distances. This is interesting because Westheimer and Levi (1987) have done the corresponding psychophysical experiments using the horizontal-line-and-point pattern. Specifically, they fixed separation distance *d* at three different values in the attraction, repulsion, and transition zones, respectively. They found that when the distance is chosen to be in the attraction or repulsion zones, increasing the inducing disparity *D* increases the attraction or repulsion, respectively. However, if the separation is picked in the transition zone but on the attraction side, increasing the inducing disparity *D* first increases the attraction but then the curve turns around to show strong repulsion. We have picked three separations accordingly in our simulations and the results shown in Fig. 8 are similar to the psychophysical observations. When *D* is very large (not shown in the figure), the model predicts no interaction. The parameter values for this figure (see caption) deviate slightly from the standard ones for the continuous curve in Fig. 6, but are within the physiological range. When the standard parameter set is used, the two left-panel curves in Fig. 8 remain similar but the transition-zone curve in the right panel will have the zero-crossing point shifted further to the right. The zero-crossing location appears to be very sensitive to the model parameters. As we mentioned before, the basic *d*-dependent attraction and repulsion phenomenon for both the vertical-line and the horizontal-line-and-point configurations is highly robust and can be readily reproduced with the parameter set in Fig. 8 (results not shown).

We would also like to mention that we made several simulations with a log normal distribution of the preferred frequency for the elaborated 2D model, in order to consider the effects of a positively skewed distribution. The zero-crossing in Fig. 8 becomes much less sensitive to the parameters, while the transition distances for the attraction/repulsion curves are somewhat
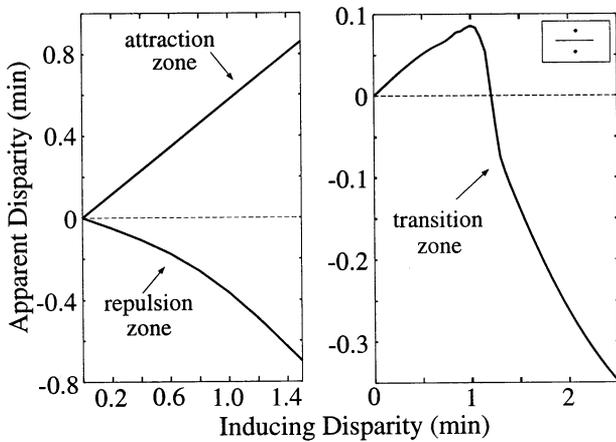
Fig. 8. Elaborated model applied to the horizontal-line-and-point configuration for inducing-disparity dependence. The apparent disparity of the middle line, averaged across all cell families with different preferred spatial frequencies and orientations and with RFs centered at the middle of the line, is plotted as a function of the disparity of the flanking points at three fixed separations between the line and the points. The actual disparity of the line is fixed at zero. The parameter values are only slightly modified from the standard set. Spatial pooling is centered at the middle of the line with the Gaussian width along each cell axis equal to 0.43 times the corresponding Gaussian width of the underlying simple cells; the preferred spatial frequency distribution is centered at 4.0 cycle/degree with a Gaussian width of 1.2 cycle/degree; the preferred orientation distribution is centered at the vertical with a Gaussian width of 57°; the frequency bandwidth is 1.3 octave and the simple-cell aspect ratio is 1.7 for all cells. The left-panel graphs correspond to line-to-point separations of 2.0 (attraction zone) and 9.0 (repulsion zone) min. The right-panel graph is plotted for a line-to-point separation of 5.5 min.
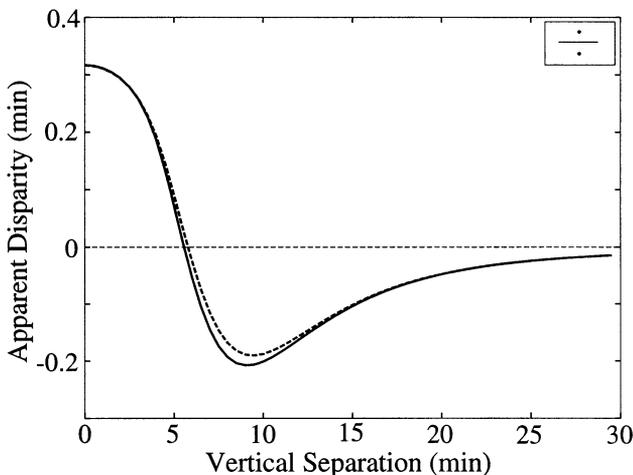

Fig. 9. Elaborated model applied to the horizontal-line-and-point configuration at different points on the line. The apparent disparity of the middle line, averaged across all cell families with different preferred spatial frequencies and orientations, is plotted as a function of the separation between the line and the flanking points. The actual disparities of the line and the flanking points are fixed at 0 and 0.5 min, respectively. Both curves correspond to the standard parameter values used for generating the continuous curve in Fig. 6, but without the spatial pooling step. The continuous curve is the result from filters centered at the middle of the line, while the dashed curve is from filters centered at one end of the line.

shifted towards a smaller separation d due to the inclusion of higher frequency cells (results not shown).

For the horizontal-line-and-point stimulus configuration, we have been using RFs centered at the center of the horizontal line, thus determining the disparity of that point in our simulations. How much variation would ensue from estimating the disparity of other points along the line? Fig. 9 addresses this concern by repeating the simulations corresponding to the continuous curve in Fig. 6 for RFs centered at the center (as before) and at one end of the line. The spatial pooling step has been turned off in generating these two curves in order to estimate the maximum variation of the apparent disparity along the line; the presence of spatial pooling helps reduce the variation by averaging over a larger area. As evident in Fig. 9, the discrepancy between the two curves is not significant. This is not particularly surprising given that for our choice of bandwidth and population preferred spatial frequency, a filter size is typically more than five times larger than the length of the middle line (4 min).

As a prediction of the model, we varied the ratio of the contrast magnitudes of the middle line to the flanking points in the horizontal-line-and-point configuration. As expected, increasing this ratio pulls the curves towards zero disparity for all separations d (results not shown). The effect of decreasing the ratio from 1 is displayed in Fig. 10. If the ratio is small the apparent disparity of the line moves towards the disparity of the points over a range of separations d, as in curves b and c, which correspond to 70 and 50% contrast ratios, respectively, and standard cell parameters. In fact, there is no repulsion zone for a small enough contrast ratio. What is a 'small enough' contrast ratio to observe such behavior is mainly dependent on the extent of spatial pooling in the model. Curves a and d demonstrate that dependence; both curves correspond to an 80% contrast ratio, but curve d was obtained with a spatial-pooling width twice that for curve a, which uses the standard value. (Indeed, in the absence of spatial pooling, at 80% contrast ratio the repulsion can be significantly more pronounced than at equal contrasts.)

We also suggest a physiological test of the model by examining disparity tuning curves for a long vertical line stimulus (the test line) in the presence of another long vertical line (the inducing line) in the RF. We consider a complex cell whose preferred orientation matches that of the stimulus orientation, with a preferred spatial frequency of $\omega = 3.5$ cycle/degree, a bandwidth of 1.5 octave, and a phase-parameter difference of $\phi_- = -0.37$ rad. Such a cell has a preferred disparity of $(-\phi_-/\omega) = 1.0$ min (Qian, 1994). The expressions for the cell's response, as a function of the test-line disparity D', the inducing-line disparity D, and the lateral separation between them d, are derived at the end of the Appendix A. The disparity tuning curves

of the cell (its response plotted against D') under various conditions are shown in Fig. 11. The continuous curve in the figure is the tuning curve for the test line
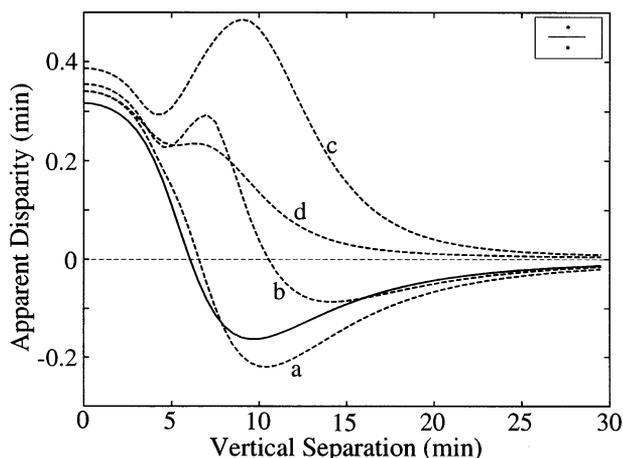

Fig. 10. Elaborated model applied to the horizontal-line-and-point configuration with variations in relative contrast between the line and points. The apparent disparity of the test line, averaged across all cell families with different preferred spatial frequencies and orientations and with RFs centered on the test line, is plotted as a function of the separation between the middle line and the flanking points. The continuous curve is the same as in Fig. 6. The dashed curves are obtained by varying parameters as follows. For curves *a*, *b* and *c* the line-to-point contrast ratios are 0.8, 0.7 and 0.5, respectively, with all other parameters same as for the continuous curve. For curve *d* this contrast ratio is 0.8 and the Gaussian width of the spatial pooling along each cell axis is increased to 0.5 times the corresponding Gaussian width of the underlying simple cells.
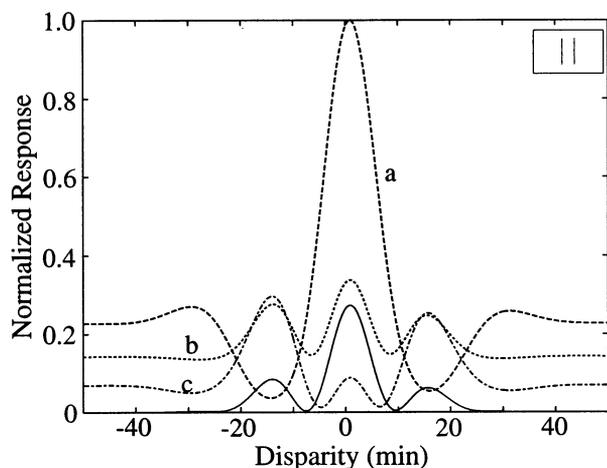

Fig. 11. Disparity tuning curves for a complex cell of the elaborate model applied to vertical-line configurations. Instantaneous firing rate is plotted as a function of the disparity of the test line. The complex cell is vertically oriented with a preferred spatial frequency of 3.5 cycle/degree, bandwidth of 1.5 octave and a binocular phase-parameter difference of -0.37 rad. The continuous curve is obtained by having a single long vertical line as the stimulus. The dashed curves are obtained by adding another long vertical line (the inducing line) at a fixed disparity of 3.0 min as part of the stimulus. For curves a, b and c the lateral separations of the two lines in the fronto-parallel plane are 0, 5.0 and 8.0 min, respectively.

alone. As expected, the main peak is at $D' = 1.0$ min, with two smaller 'side peaks' (Zhu & Qian, 1996). (The number of noticeable peaks on each side of the main one depends on the cell parameters.) The dashed curves are the tuning curves when the inducing line is introduced at a fixed disparity of $D = 3$ min and various lateral separations d. Curves a, b and c correspond to separations of 0, 5 and 8 min, respectively. Although the peak at $D' = 1.0$ min remains in the presence of the inducing line, it is no longer an absolute maximum for certain choices of d. As d increases, the main peak shrinks while the side peaks grow and eventually become more prominent. Although we only present here the results from one set of arbitrary stimulus parameters, we have performed many other simulations to reach the general prediction that the main peak in the disparity tuning curve for the test line alone may not remain so with the introduction of the inducing line. The main-peak location with the inducing line present is always at either the main- or one of the side-peak locations of the tuning curve for the test line alone. Note that it is the smooth variation of the peak locations (with d) among a population of cells that underlies our explanation of disparity attraction and repulsion.

4. Discussion

In this paper we applied our previous stereo model (Qian, 1994; Qian & Zhu, 1997a) and its 2D extensions to the stimulus configurations used by Westheimer (1986) and Westheimer and Levi (1987), and demonstrated that the psychophysically observed distance-dependent disparity attraction/repulsion phenomenon can be naturally explained by the known physiological properties of binocular cells in the primary visual cortex. Although intuitively one may speculate that the attractive and repulsive disparity interactions between nearby visual features may have something to do with the excitatory and inhibitory subregions of simple cells' binocular RFs, a clear understanding of the phenomenon cannot be obtained without quantitative modeling. In fact, as we have shown elsewhere (Qian, 1994; Zhu & Qian, 1996; Qian & Zhu, 1997a), simple cells do not always have reliable disparity tuning curves and therefore cannot be directly used to compute disparity and explain the illusion. Although the disparity interaction in our models originates from the simultaneous presence of stimulus features in simple cells' excitatory and inhibitory subfields, the final explanation comes from the responses of complex cells that are non-linear functions of simple cell responses and do not have separate excitatory and inhibitory subregions within their RFs. Furthermore, without the quantitative modeling shown here, it would not be possible to

assert that the disparity interactions in the vertical-line configuration *and* the horizontal-line-and-point configuration can both be explained by the 2D RF structure of binocular cells in the brain. Our results suggest that it may not be necessary to propose explicit excitatory and inhibitory connections between disparity-tuned units (Lehky & Sejnowski, 1990) to explain the observed disparity attraction and repulsion. However, they are not against such proposals either. It is possible that both the RF organization and the explicit connections contribute to the disparity interaction.

The simulation results indicate that the transition distance at which disparity interaction changes from attraction to repulsion strongly depends on the orientation pooling distribution and the spatial-frequency pooling distribution functions for the cell population. The dependence on the spatial-frequency distribution is particularly strong with the effect that when the distribution is shifted towards the high frequency, the transition distance becomes smaller. Interestingly, in the psychophysical experiments by Westheimer (1986), it was found that different individual subjects have different transition distances ranging from 3 to 8 min of arc. One possible explanation of this experimental result is therefore that different subjects may have different preferred spatial-frequency and/or orientation distributions for the cells in their visual cortex. In addition, even for subjects with identical pooling distributions, those with less perfect front-end focusing system may pre-filter out more high frequency components in the stimulus and obtain an effectively different preferred frequency distribution among the activated cells in the brain. A prediction based on this argument is that those subjects with higher visual acuity should have smaller transition distances.

We used the phase-difference RF description proposed by Freeman and coworkers (Freeman & Ohzawa, 1990; Ohzawa et al., 1990, 1996, 1997; DeAngelis et al., 1991) for modeling binocular simple cells in this paper. An alternative RF model is the traditional position-shift model (Bishop, Henry & Smith, 1971; Maske, Yamane & Bishop, 1984). We have previously investigated the similarities and differences between these two RF models (Zhu & Qian, 1996; Qian & Andersen, 1997; Qian & Zhu, 1997a) and found that when stimulus disparities are considerably smaller than RF sizes these two models and their hybrid often produce similar results except under some special cases such as narrow-band stimulation. Since both types of RF models have alternating excitatory and inhibitory subfields and since the origin of disparity interaction is the simultaneous presence of stimulus features in these subfields, both models should produce similar attraction/repulsion phenomena. We have verified this conclusion by repeating the simulation shown in Fig. 3 but with the position-shift RFs for the front-end simple cells. The results

(not shown) are nearly identical to those shown in Fig. 3 except that at small separations the attraction increases slightly before turning around towards repulsion.

Despite the success of our model in explaining the basic findings of the distance (d) dependence and the inducing disparity (D) dependence of the disparity interactions in both the vertical-line (Westheimer, 1986) and the horizontal-line-and-point configurations (Westheimer & Levi, 1987), an unresolved issue is how the model should be modified to explain the results observed when the horizontal line and the flanking points had opposite signs of contrast and when the horizontal-line length was varied over a wide range (Westheimer & Levi, 1987). In the case of opposite contrast signs, our simulation results differ from the actual observations in that the simulated curves do not cross over to the attraction side early enough and the attraction is not as strong. In the case of line-length variations, the attractive interactions in our simulations do not decay with the line length as fast as in the actual experiment. Nevertheless, it should be emphasized that the most robust feature of the observed disparity interaction is the distance-dependent attraction and repulsion phenomenon since it has been demonstrated in two separate experiments with two very different stimulus configurations (Westheimer, 1986; Westheimer & Levi, 1987). Our model is similarly robust in explaining this phenomenon: the simulated attraction and repulsion effects under the two different stimulus configurations correspond closely to the observations even under large variations in the key parameters of the model. To our knowledge, previous models have neither taken into account the details of these configurations nor provided as quantitative an explanation of the observed disparity interaction.

A different, but closely related disparity attraction and repulsion phenomenon has been reported by Stevenson, Cormack and Schor (1991). Using transparent random dots stereograms with two completely overlapping surfaces, these authors found that when the separation between the two surfaces in depth (instead of laterally) is small, the two surfaces attract each other while when the separation in depth is large, they repel each other. We currently are unable to model any phenomena involving transparent stereograms because our present models cannot solve the problem of stereo transparency. Indeed, to our knowledge, no physiologically realistic model of stereo transparency has been proposed although there are several computationally or psychophysically inspired algorithms available (Prazdny, 1985; Pollard, Mayhew & Frisby, 1985; Nishihara, 1987; Qian & Sejnowski, 1989; Marshall, Kalarickal & Graves, 1996). We hope to extend our physiological models to solve stereo transparency and then apply them to this experimental observation in the future.

A distinctive feature of our models is that they are based on physiological properties of binocular cells in the brain while at the same time they allow a significant degree of mathematical analysis to be carried out to gain insights into the problem. Our results suggest that the perceived interaction in the disparity domain may be viewed as a direct consequence of the underlying organization of the binocular RFs. The work therefore provides a link between physiology and perception, and between seemingly unrelated experimental data. It would be interesting to investigate whether similar approaches can be used to explain the attraction and repulsion phenomena in motion perception (Marshak & Sekuler, 1979; Qian & Geesaman, 1995) and orientation discrimination (Blakemore, Carpenter & Georgeson, 1970; Bouma & Andriessen, 1970; Carandini & Ringach, 1997). It would also be worthwhile to examine whether the model could explain the related phenomenon of disparity interpolation (Wurger & Landy, 1989), which could be viewed as another form of disparity interaction.

$$+ \left| \int_{-\infty}^{\infty} g(\vec{r} - \vec{r}_0)\{\sin[\vec{\omega}\cdot(\vec{r} - \vec{r}_0) + \phi_1]\, I_l(\vec{r}) \right.$$
$$\left. + \sin[\vec{\omega}\cdot(\vec{r} - \vec{r}_0) + \phi_r]\, I_r(\vec{r})\}\mathrm{d}^2 r \right|^2 \tag{32}$$

where $\vec{r} = (x, y)$, and $\mathrm{d}^2 r \equiv \mathrm{d}x\,\mathrm{d}y$. The complex–cell response is generated through spatial pooling of the above quadrature-pair response with a weighting function $W(\vec{r}_0)$, and in complex notation is given by

$$r_c = \int_{-\infty}^{\infty} \left| \int_{-8}^{\infty} g(\vec{r} - \vec{r}_0)e^{i\vec{\omega}\cdot\vec{r}}[e^{i\phi_-}\, I_l(\vec{r}) + I_r(\vec{r})]\, \mathrm{d}^2 r \right|^2$$
$$W(\vec{r}_0)\,\mathrm{d}^2 r_0 \tag{33}$$

where

$$\phi_- \equiv \phi_1 - \phi_r \tag{34}$$

As only the cross term in Eq. (33) will depend on ϕ_-, maximizing the complex-cell response is equivalent to maximizing

$$\mathscr{I}_{lr} \equiv \int_{-\infty}^{\infty} g(\vec{r} - \vec{r}_0)g(\vec{s} - \vec{r}_0)I_l(\vec{r})I_r(\vec{s})\cos[\vec{\omega}\cdot(\vec{r} - \vec{s}) + \phi_-]$$
$$\times\ W(\vec{r}_0)\,\mathrm{d}^2 r\,\mathrm{d}^2 s\,\mathrm{d}^2 r_0 \tag{35}$$

The primitive and elaborate versions of the model are defined by specific choices of the envelope and weighting functions $g(\vec{r})$ and $W(\vec{r})$, as described in the following sections.

When the left and right images have same contrast polarity, the maximum response of Eq. (35) is obtained when ϕ_- is given by

$$\hat{\phi}_- = -\arctan\left\{ \frac{\displaystyle\int_{-\infty}^{\infty} g(\vec{r} - \vec{r}_0)g(\vec{s} - \vec{r}_0)\, I_l(\vec{r})I_r(\vec{s})\, \sin[\vec{\omega}\cdot(\vec{r} - \vec{s})]\, W(\vec{r}_0)\,\mathrm{d}^2 r\,\mathrm{d}^2 s\,\mathrm{d}^2 r_0}{\displaystyle\int_{-\infty}^{\infty} g(\vec{r} - \vec{r}_0)g(\vec{s} - \vec{r}_0)\, I_l(\vec{r})I_r(\vec{s})\, \cos[\vec{\omega}\cdot(\vec{r} - \vec{s})]\, W(\vec{r}_0)\,\mathrm{d}^2 r\,\mathrm{d}^2 s\,\mathrm{d}^2 r_0} \right\} \tag{36}$$

The disparity $\tilde{D}$ estimated from a family of complex cells all with the same parameter values but covering the full range of ϕ_- is taken to be (Qian, 1994)

$$\tilde{D} = -\frac{\hat{\phi}_-}{\omega_1} \tag{37}$$

A.1. Primitive 2D model

This version of the model is defined by choosing $W(\vec{r}_0) = \delta(\vec{r}_0)$, i.e. no spatial pooling, and the envelope functions of Eq. (17), when evaluating Eqs. (33) and (36). From Eq. (33):

$$r_c = \left| e^{i\phi_-} \int_{-\infty}^{\infty} g(\vec{r})\, I_l(\vec{r})\, e^{i\vec{\omega}\cdot\vec{r}}\, \mathrm{d}^2 r \right.$$

Appendix A. Complex-cell response and disparity estimation

In general, for a given RF envelope function $g(\vec{r})$ centered at $\vec{r} = 0$ the quadrature-pair response at the location $\vec{r}_0$ is given by

$$r_q(\vec{r}_0) = \left| \int_{-\infty}^{\infty} g(\vec{r} - \vec{r}_0)\{\cos[\vec{\omega}\cdot(\vec{r} - \vec{r}_0) + \phi_1]\, I_l(\vec{r}) \right.$$
$$\left. + \cos[\vec{\omega}\cdot(\vec{r} - \vec{r}_0) + \phi_r]\, I_r(\vec{r})\}\mathrm{d}^2 r \right|^2$$

$$+ \left| \int_{-\infty}^{\infty} g(\vec{r}) I_r(\vec{r}) \, e^{i\vec{\omega}\cdot\vec{r}} \, d^2r \right|^2 = \left| e^{i\phi} - \rho_1 \, e^{i\theta_1} + \rho_r \, e^{i\theta_r} \right|^2 \quad (38)$$

where ρ and θ signify the magnitude and phase of each integral. The maximum of r_c now corresponds to the choice

$$\hat{\phi}_- = \theta_r - \theta_1$$

$$= \arctan \left[\frac{\int_{-\infty}^{\infty} g(\vec{r}) I_r(\vec{r}) \sin(\vec{\omega}\cdot\vec{r}) \, d^2r}{\int_{-\infty}^{\infty} g(\vec{r}) I_r(\vec{r}) \cos(\vec{\omega}\cdot\vec{r}) \, d^2r} \right]$$

$$- \arctan \left[\frac{\int_{-\infty}^{\infty} g(\vec{r}) I_l(\vec{r}) \sin(\vec{\omega}\cdot\vec{r}) \, d^2r}{\int_{-\infty}^{\infty} g(\vec{r}) I_l(\vec{r}) \cos(\vec{\omega}\cdot\vec{r}) \, d^2r} \right] \quad (39)$$

Eqs. (36) and (39) are equivalent in this case, though the latter is computationally more convenient. Note that θ_r and θ_1 are simply the Fourier phases of the envelope-weighted right and left retinal images.

A.1.1. 1D and primitive 2D models applied to the vertical-line configuration

For the vertical-line pattern, it is shown here that the 1D and primitive 2D models result in identical expressions for the estimated disparity by a given cell family. For the latter model, substituting from Eqs. (1), (17) and (39) into Eq. (37), yields

$$\tilde{D} = \frac{1}{\omega_1} \left\{ \arctan \left[R\left(\omega_1 \left(d + \frac{D}{2} \right) \right) \right] \right.$$

$$\left. - \arctan \left[R\left(\omega_1 \left(d - \frac{D}{2} \right) \right) \right] \right\}$$

where

$$R(x) \equiv \frac{\exp\left(-\frac{x^2}{2a_1^2} \right) \sin x}{1 + \exp\left(-\frac{x^2}{2a_1^2} \right) \cos x + c \sqrt{2\pi\sigma_1^2} \, e^{-a_1^2/2}}$$

and

$$a_1 \equiv \sigma_1 \omega_1 \quad (40)$$

This expression for $\tilde{D}$ does not depend on the vertical component of the parameters, ω_2 and σ_2. Hence, for fixed ω_1 and σ_1, the apparent disparity is independent of the orientation of the subfield and the vertical extent of the envelope. Essentially, this expression for $\tilde{D}$ is that of a 1D filter, as the contribution of the other dimension to the complex-cell response factors out as an overall scaling of that response. This amplitude scaling does not affect the peak locations of the disparity tuning curves.

A.1.2. Primitive 2D model applied to the horizontal-line-and-point configuration

The estimated disparity is obtained from Eqs. (37) and (39) after substitutions from Eqs. (2)–(7) and (17):

$$\tilde{D} = \frac{2}{\omega_1} \arctan$$

$$\left[\frac{2\varepsilon \sum\limits_{i=1}^{2} \exp\left(-\frac{u_i^2}{2a_1^2} \right) \sin u_i \sum\limits_{j=1}^{2} \exp\left(-\frac{v_j^2}{2a_2^2} \right) \cos v_j}{\kappa + 2\varepsilon \sum\limits_{i=1}^{2} \exp\left(-\frac{u_i^2}{2a_1^2} \right) \cos u_i \sum\limits_{j=1}^{2} \exp\left(-\frac{v_j^2}{2a_2^2} \right) \cos v_j} \right]$$

$$(41)$$

where

$$\kappa \equiv 2\pi c \sigma_1 \sigma_2 \exp\left(-\frac{1}{2} \sum_{i=1}^{2} \omega_i^2 \sigma_i^2 \right) + 1$$

$$+ 2 \sum_{j=1}^{2} \exp\left(-\frac{j^2}{2\sigma_1^2} \right) \cos(\omega_1 j), \quad a_i \equiv \sigma_i \omega_i$$

$$u_1 \equiv \frac{(D-1)\omega_1}{2} \quad u_2 \equiv \frac{(D+1)\omega_1}{2} \quad v_1 \equiv (d-0.5)\omega_2$$

$$v_2 \equiv (d+0.5)\omega_2 \quad (42)$$

In evaluating Eq. (39) in this case, the Fourier phases of the two weighted retinal images, θ_1 and θ_r, turn out equal and opposite in sign. Hence, the transition between the attraction and repulsion zones, i.e. $\tilde{D} = 0$, occurs when these Fourier phases vanish.

A.2. Elaborated 2D model

For this model, the computed disparity is best expressed in terms of the parameters in two coordinate systems. One coordinate system (the unprimed) is identical to the one used above, with its origin at the center of the RFs (the location where disparity is being computed) and its horizontal axis along the interocular axis. We refer to this system as the 'standard coordinates'. The other coordinate system (the primed) is rotated relative to the standard one such that its horizontal axis is perpendicular to the preferred orientation of the cell under consideration, and is dubbed the 'cell coordinates'. They are related to each other by

$$\mathbf{r}' = \mathbf{R}\mathbf{r} \quad \text{with} \quad \mathbf{R} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \quad (43)$$

($\theta = 0°$ for vertically and $\theta = 90°$ for horizontally oriented cells). By definition, in the cell coordinates the components of the preferred spatial frequency are given by

$$(\omega_1', \omega_2') = (\omega, 0) \quad (44)$$

so that in the standard coordinates they become

$$(\omega_1, \omega_2) = \omega(\cos\theta, \sin\theta) \quad (45)$$

The elaborated model is defined by a correctly aligned envelope function shown in Eqs. (20)–(22), so that $\mathbf{\Sigma} = \mathbf{R}^T \mathbf{\Sigma}' \mathbf{R}$, and by the choice

$$W(\vec{r}_0) = \frac{\sqrt{\det(\mathbf{N})}}{2\pi} \exp\left(-\frac{\mathbf{r}_0^{\mathrm{T}}\mathbf{N}\mathbf{r}_0}{2}\right) \quad \text{with}$$

$$\mathbf{N} \equiv \mathbf{R}^{\mathrm{T}}\begin{pmatrix} 1/\eta_1'^2 & 0 \\ 0 & 1/\eta_2'^2 \end{pmatrix}\mathbf{R} \tag{46}$$

Here $\mathbf{N}$ is taken to be diagonal in the cell coordinates so that W is oriented along the preferred orientation of the cells. Two relations that connect our parameter values to physiological data are

$$\sigma_1'\omega = \sqrt{2\ln 2}\,\frac{2^b+1}{2^b-1} \quad \text{and} \quad \sigma_2' = q\sigma_1' \tag{47}$$

where b and q are the bandwidth and aspect ratio of the simple cell, respectively.

With these parameterizations we can perform the spatial pooling in Eq. (36) to obtain

$$\hat{\phi}_- = -\arctan \frac{\displaystyle\int_{-\infty}^{\infty} F(\vec{r},\vec{s})\,I_l(\vec{r})\,I_r(\vec{s})\sin[\vec{\omega}\cdot(\vec{r}-\vec{s})]\,\mathrm{d}^2r\,\mathrm{d}^2s}{\displaystyle\int_{-\infty}^{\infty} F(\vec{r},\vec{s})\,I_l(\vec{r})\,I_r(\vec{s})\cos[\vec{\omega}\cdot(\vec{r}-\vec{s})]\,\mathrm{d}^2r\,\mathrm{d}^2s} \tag{48}$$

where

$$F(\vec{r},\vec{s}) \equiv \exp\left[-\frac{\mathbf{r}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{r}}{2} - \frac{\mathbf{s}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{s}}{2} + \frac{(\mathbf{r}+\mathbf{s})^{\mathrm{T}}\mathbf{B}(\mathbf{r}+\mathbf{s})}{2}\right]$$

and

$$\mathbf{B} \equiv \begin{pmatrix} 1/\beta_1^2 & 1/\beta_{12} \\ 1/\beta_{21} & 1/\beta_2^2 \end{pmatrix} = \mathbf{R}^{\mathrm{T}}\begin{pmatrix} 1/\beta_1'^2 & 0 \\ 0 & 1/\beta_2'^2 \end{pmatrix}\mathbf{R},$$

$$\beta_i'^2 \equiv 2\sigma_i'^2 + \frac{\sigma_i'^4}{\eta_i'^2} \quad (i=1,2) \tag{49}$$

As expected, Eqs. (48) and (49) indicate that spatial pooling has broadened the Gaussian weighting of the images.

A.2.1. Elaborated 2D model applied to horizontal-line-and-point configuration

The complete expression of the computed disparity, Eq. (37), for this stimulus configuration is derived by substituting Eqs. (2)–(7) in Eq. (48), to arrive at

$$\tilde{D} = \frac{1}{\omega_1}\arctan(\mathcal{N}/\mathcal{D})$$

where

$$\mathcal{N} \equiv \sum_{i,i'=1}^{2}\sum_{j,j'=1}^{4} \mathscr{I}_{1ii'jj'}\sin[\vec{\omega}\cdot(\vec{\Delta}_{ij}^{(l)}-\vec{\Delta}_{ij}^{(r)})]$$

$$+ \sum_{i=1}^{2}\sum_{j=1}^{4}\sum_{k=1}^{2N+1} \{\mathscr{I}_{2ijk}^{(l)}\sin[\vec{\omega}\cdot(\vec{\Delta}_{ij}^{(l)}-\vec{X}_k)]$$

$$+ \mathscr{I}_{2ijk}^{(r)}\sin[\vec{\omega}\cdot(\vec{X}_k-\vec{\Delta}_{ij}^{(r)})]\}$$

$$+ \sum_{i=1}^{2}\sum_{j=1}^{4} \left\{\mathscr{I}_{3ij}^{(l)}\sin\left(2\frac{\alpha_1'^2}{\gamma_1'^2}\vec{\omega}\cdot\vec{\Delta}_{ij}^{(l)}\right)\right.$$

$$\left.+ \mathscr{I}_{3ij}^{(r)}\sin\left(-2\frac{\alpha_1'^2}{\gamma_1'^2}\vec{\omega}\cdot\vec{\Delta}_{ij}^{(r)}\right)\right\}$$

and

$$\mathcal{D} \equiv \sum_{i,j=1}^{2N+1} \exp\left[-\frac{1}{2}\mathbf{X}_i^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{X}_i - \frac{1}{2}\mathbf{X}_j^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{X}_j\right.$$

$$\left.+ \frac{1}{2}(\mathbf{X}_i+\mathbf{X}_j)^{\mathrm{T}}\mathbf{B}(\mathbf{X}_i+\mathbf{X}_j)\right]\cos[\vec{\omega}\cdot(\vec{X}_i-\vec{X}_j)]$$

$$+ 4\pi c\alpha_1'\alpha_2'\,\mathrm{e}^{-\omega^2\alpha_1'^2/2}\sum_{i=1}^{2N+1}\exp\left[-\frac{1}{2}\mathbf{X}_i^{\mathrm{T}}(\boldsymbol{\Sigma}-\boldsymbol{\Lambda})\mathbf{X}_i\right]$$

$$\times \cos\left(2\frac{\alpha_1'^2}{\gamma_1'^2}\vec{\omega}\cdot\vec{X}_i\right)$$

$$+ \sum_{i,i'=1}^{2}\sum_{j,j'=1}^{4} \mathscr{I}_{1ii'jj'}\cos[\vec{\omega}\cdot(\vec{\Delta}_{ij}^{(l)}-\vec{\Delta}_{ij}^{(r)})]$$

$$+ 2\pi^2 c^2\sigma_1'\sigma_2'\gamma_1'\gamma_2'\,\mathrm{e}^{-\omega^2\sigma_1'^2}$$

$$+ \sum_{i=1}^{2}\sum_{j=1}^{4}\sum_{k=1}^{2N+1} \{\mathscr{I}_{2ijk}^{(l)}\cos[\vec{\omega}\cdot(\vec{\Delta}_{ij}^{(l)}-\vec{X}_k)]$$

$$+ \mathscr{I}_{2ijk}^{(r)}\cos[\vec{\omega}\cdot(\vec{X}_k-\vec{\Delta}_{ij}^{(r)})]\}$$

$$+ \sum_{i=1}^{2}\sum_{j=1}^{4} \left\{\mathscr{I}_{3ij}^{(l)}\cos\left(2\frac{\alpha_1'^2}{\gamma_1'^2}\vec{\omega}\cdot\vec{\Delta}_{ij}^{(l)}\right)\right.$$

$$\left.+ \mathscr{I}_{3ij}^{(r)}\cos\left(2\frac{\alpha_1'^2}{\gamma_1'^2}\vec{\omega}\cdot\vec{\Delta}_{ij}^{(r)}\right)\right\} \tag{50}$$

with the definitions

$$\mathscr{I}_{1ii'jj'} \equiv \varepsilon^2 \exp\left[-\frac{1}{2}\boldsymbol{\Delta}_{ij}^{(l)\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{\Delta}_{ij}^{(l)} - \frac{1}{2}\boldsymbol{\Delta}_{i'j'}^{(r)\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{\Delta}_{i'j'}^{(r)}\right.$$

$$\left.+ \frac{1}{2}(\boldsymbol{\Delta}_{ij}^{(l)}+\boldsymbol{\Delta}_{i'j'}^{(r)})^{\mathrm{T}}\mathbf{B}(\boldsymbol{\Delta}_{ij}^{(l)}+\boldsymbol{\Delta}_{i'j'}^{(r)})\right]$$

$$\mathscr{I}_{2ijk}^{(l)} \equiv \varepsilon \exp\left[-\frac{1}{2}\mathbf{X}_k^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{X}_k - \frac{1}{2}\boldsymbol{\Delta}_{ij}^{(l)\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{\Delta}_{ij}^{(l)}\right.$$

$$\left.+ \frac{1}{2}(\mathbf{X}_k+\boldsymbol{\Delta}_{ij}^{(l)})^{\mathrm{T}}\mathbf{B}(\mathbf{X}_k+\boldsymbol{\Delta}_{ij}^{(l)})\right]$$

$$\mathscr{I}_{2ijk}^{(r)} \equiv \varepsilon \exp\left[-\frac{1}{2}\mathbf{X}_k^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{X}_k - \frac{1}{2}\boldsymbol{\Delta}_{ij}^{(r)\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{\Delta}_{ij}^{(r)}\right.$$

$$\left.+ \frac{1}{2}(\mathbf{X}_k+\boldsymbol{\Delta}_{ij}^{(r)})^{\mathrm{T}}\mathbf{B}(\mathbf{X}_k+\boldsymbol{\Delta}_{ij}^{(r)})\right]$$

$$\mathscr{I}_{3ij}^{(l)} \equiv 2\pi c\varepsilon\alpha_1'\alpha_2'\,\mathrm{e}^{-\omega^2\alpha_1^2/2}\exp\left[-\frac{1}{2}\boldsymbol{\Delta}_{ij}^{(l)\mathrm{T}}(\boldsymbol{\Sigma}-\boldsymbol{\Lambda})\boldsymbol{\Delta}_{ij}^{(l)}\right]$$

$$\mathscr{I}_{3ij}^{(r)} \equiv 2\pi c\varepsilon\alpha_1'\alpha_2'\,\mathrm{e}^{-\omega^2\alpha_1^2/2}\exp\left[-\frac{1}{2}\boldsymbol{\Delta}_{ij}^{(r)\mathrm{T}}(\boldsymbol{\Sigma}-\boldsymbol{\Lambda})\boldsymbol{\Delta}_{ij}^{(r)}\right]$$

$$\gamma_n'^{-2} \equiv \frac{1}{2}\sigma_n'^{-2} - \beta_n'^{-2} \quad \text{and} \quad \alpha_n'^{-2} \equiv \sigma_n'^{-2} - \beta_n'^{-2}$$

$$\text{for} \quad n=1,2$$

$$\boldsymbol{\Lambda} \equiv \mathbf{R}^{\mathrm{T}}\begin{pmatrix} (\beta_1'^2-\sigma_1'^2)^{-1} & 0 \\ 0 & (\beta_2'^2-\sigma_2'^2)^{-1} \end{pmatrix}\mathbf{R} \tag{51}$$

A.2.2. Elaborated 2D model applied to vertical-line configuration

To simplify expressions, we take $c = 0$ in Eq. (1) as there was no background illumination in the experiments of Westheimer (1986). The result for the computed disparity is

$$\tilde{D} = \frac{1}{\omega_1} \arctan(\mathcal{N}/\mathcal{D})$$

where

$$\mathcal{N} \equiv \mathscr{I}_1^{(l)} \sin\left[\omega_o\left(\frac{D}{2} + d\right)\right] + \mathscr{I}_1^{(r)} \sin\left[\omega_o\left(\frac{D}{2} - d\right)\right] + \mathscr{I}_2 \sin(\omega_o D)$$

and

$$\mathcal{D} \equiv 1 + \mathscr{I}_1^{(l)} \cos\left[\omega_o\left(\frac{D}{2} + d\right)\right] + \mathscr{I}_1^{(r)} \cos\left[\omega_o\left(\frac{D}{2} - d\right)\right] + \mathscr{I}_2 \cos(\omega_o D) \tag{52}$$

with the definitions

$$\mathscr{I}_1^{(l)} \equiv \varepsilon \exp\left[-\left(\frac{D}{2} + d\right)^2\left(\frac{1}{4\alpha^2} + \frac{1}{2\gamma^2}\right)\right]$$

$$\mathscr{I}_1^{(r)} \equiv \varepsilon \exp\left[-\left(\frac{D}{2} - d\right)^2\left(\frac{1}{4\alpha^2} + \frac{1}{2\gamma^2}\right)\right]$$

$$\mathscr{I}_2 \equiv \varepsilon^2 \exp\left(-\frac{D^2}{4\alpha^2} - \frac{2d^2}{\gamma^2}\right)$$

$$\omega_o \equiv \omega_1 - \omega_2 \frac{\sigma_2^2}{\sigma_{12}} \quad \alpha^2 \equiv \frac{\sigma_1'^2 \sigma_2'^2}{\sigma_2^2}$$

$$\gamma^2 \equiv \left(\frac{1}{2\sigma_2^2} - \frac{1}{\beta_2^2}\right)\Big/\det\left(\frac{\Sigma}{2} - B\right) \tag{53}$$

In the absence of spatial pooling ($W(\vec{r}_0) = \delta(\vec{r}_0)$), $\mathbf{B} = 0$ and $\gamma^2 = 2\alpha^2$. It can then be shown from Eqs. (52) and (53) that $\tilde{D}$ satisfies the same expression as that for the primitive 2D model, Eq. (40), with the substitutions

$$\omega_1 \to \omega_o \quad \text{and} \quad \sigma_1^2 \to \alpha^2 \tag{54}$$

in the latter equation. With circular RFs, i.e. $q = 1$ in Eq. (47), $\omega_1 = \omega_o$ and $\sigma_1^2 = \alpha^2$. In that case, the elaborated model without spatial pooling and the primitive 2D model yield identical results for the vertical-line configuration.

Finally, we present the expressions used to generate the disparity tuning curves of a vertically oriented model complex cell in response to a vertical-line stimulus, the test line, presented either by itself or with another vertical line, the inducing line, in the RF. The stimuli are given by

$$I_l(x) = \delta(x - D'/2) + A\delta(x - d - D/2)$$

$$I_r(x) = \delta(x + D'/2) + A\delta(x - d + D/2) \tag{55}$$

where $A = 1$ when the inducing line is present and $A = 0$ when it is absent, D' and D are the disparities of the test and inducing lines, respectively, and d is the lateral separation between the two lines. The complex–cell response Eq. (33) can now be reduced to

$$r_c = \frac{1}{2\pi\sigma_1\eta_1\varsigma_1}\Bigg\{\sum_{i=l,r}\Bigg[\exp\left(-\frac{a_i^2}{\sigma_1^2} + \frac{2a_i^2}{\sigma_1^2\varsigma_1^2}\right)$$

$$+ 2A \exp\left[-\frac{a_i^2 + b_i^2}{2\sigma_1^2} + \frac{(a_i + b_i)^2}{2\sigma_1^2\varsigma_1^2}\right]\cos[\omega(a_i - b_i)]$$

$$+ A^2 \exp\left(-\frac{b_i^2}{\sigma_1^2} + \frac{2b_i^2}{\sigma_1^2\varsigma_1^2}\right)\Bigg]$$

$$+ 2A^2 \exp\left(-\frac{2d^2 + D^2/2}{2\sigma_1^2} + \frac{2d^2}{\sigma_1^2\varsigma_1^2}\right)\cos(\phi_- + \omega D)$$

$$+ 2 \exp\left(-\frac{D'^2}{4\sigma_1^2}\right)\cos(\phi_- + \omega D')$$

$$+ 2A \exp\left[-\frac{a_l^2 + b_r^2}{2\sigma_1^2} + \frac{(a_l + b_r)^2}{2\sigma_1^2\varsigma_1^2}\right]\cos[\phi_- + \omega(a_l - b_r)]$$

$$+ 2A \exp\left[-\frac{a_r^2 + b_l^2}{2\sigma_1^2} + \frac{(a_r + b_l)^2}{2\sigma_1^2\varsigma_1^2}\right]\cos[\phi_- + \omega(b_l - a_r)]\Bigg\} \tag{56}$$

where

$$a_l \equiv \frac{D'}{2} \equiv -a_r \quad b_l \equiv d + \frac{D}{2} \quad b_r \equiv d - \frac{D}{2} \quad \text{and}$$

$$\varsigma_1 \equiv \frac{\sqrt{2\eta_1^2 + \sigma_1^2}}{\eta_1} \tag{57}$$

References

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

Anzai, A., Ohzawa, I., & Freeman, R. D. (1999). Neural mechanisms for processing binocular information: I. Simple cells. *Journal of Neurophysiology*, *82*, 891–908.

Bishop, P. O., Henry, G. H., & Smith, C. J. (1971). Binocular interaction fields of single units in the cat striate cortex. *Journal of Physiology*, *216*, 39–68.

Blakemore, C., Carpenter, R. H. S., & Georgeson, M. A. (1970). Lateral inhibition between orientation detectors in the human visual system. *Nature*, *228*, 37–39.

Bouma, H., & Andriessen, J. J. (1970). Induced changes in the perceived orientation of line segments. *Vision Research*, *10*, 333–349.

Carandini, M., & Ringach, D. L. (1997). Prediction of a recurrent model of orientation selectivity. *Vision Research*, *37*, 3061–3071.

DeAngelis, G. C., Ohzawa, I., & Freeman, R. D. (1991). Depth is encoded in the visual cortex by a specialized receptive field structure. *Nature*, *352*, 156–159.

DeValois, R. L., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, *22*, 545–559.

Fleet, D. J., Wagner, H., & Heeger, D. J. (1996). Encoding of binocular disparity: energy models, position shifts and phase shifts. *Vision Research*, *36*, 1839–1858.

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

Georgopoulos, A. P., Schwartz, A. B., & Kettner, R. E. (1986). Neuronal population coding of movement direction. *Science*, *233*, 1416–1419.

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

Land, E. H. (1977). The retinex theory of color vision. *Science America*, *237*, 108–128.

Lehky, S. R., & Sejnowski, T. J. (1990). Neural model of stereoacuity and depth interpolation based on a distributed representation of stereo disparity. *Journal of Neuroscience*, *10*, 2281–2299.

Marshak, W. M., & Sekuler, R. (1979). Mutual repulsion between moving visual targets. *Science*, *205*, 1399–1401.

Marshall, J. A., Kalarickal, G. J., & Graves, E. B. (1996). Neural model of visual stereomatching: slant, transparency, and clouds. *Network: Computerised Neural Systems*, *7*, 635–669.

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

Mikaelian, S., & Qian, N. (1997). Disparity attraction and repulsion in a two-dimensional stereo model. *Society of Neuroscience Abstracts*, *23*, 569.

Nishihara, H. K. (1987). Hidden information in transparent stereograms, *Proceedings of the twenty-first asilomar conference on signals, systems and computers* (pp. 695–700), Pacific Grove, CA.

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

Ohzawa, I., DeAngelis, G. C., & Freeman, R. D. (1996). Encoding of binocular disparity by simple cells in the cat's visual cortex. *Journal of Neurophysiology*, *75*, 1779–1805.

Ohzawa, I., DeAngelis, G. C., & Freeman, R. D. (1997). Encoding of binocular disparity by complex cells in the cat's visual cortex. *Journal of Neurophysiology*, *77*, 2879–2909.

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

Poggio, G. F., & Poggio, T. (1984). The analysis of stereopsis. *Annual Review in Neuroscience*, *7*, 379–412.

Pollard, S. B., Mayhew, J. E., & Frisby, J. P. (1985). PMF: a stereo correspondence algorithm using a disparity gradient limit. *Perception*, *14*, 449–470.

Prazdny, K. (1985). Detection of binocular disparities. *Biological Cybernetics*, *52*, 93–99.

Qian, N. (1994). Computing stereo disparity and motion with known binocular cell properties. *Neural Computations*, *6*, 390–404.

Qian, N. (1997). Binocular disparity and the perception of depth. *Neuron*, *18*, 359–368.

Qian, N., & Andersen, R. A. (1997). A physiological model for motion-stereo integration and a unified explanation of Pulfrich-like phenomena. *Vision Research*, *37*, 1683–1698.

Qian, N., & Geesaman, B. J. (1995). Motion repulsion depends on the distance between the moving elements. *Investegative Ophthalmology and Visual Science Supplement (ARVO)*, *36*, 50.

Qian, N., & Sejnowski, T. J. (1989). Learning to solve random-dot stereograms of dense and transparent surfaces with recurrent backpropagation. In D. S. Touretzky, G. E. Hinton, & T. J. Sejnowski, *Proceedings of the 1988 connectionist models summer school* (pp. 435–443), San Mateo, CA.

Qian, N., & Zhu, Y. (1995). Physiological computation of binocular disparity. *Society of Neuroscience Abstracts*, *21*, 1507.

Qian, N., & Zhu, Y. (1997). Physiological computation of binocular disparity. *Vision Research*, *37*, 1811–1827.

Qian, N., & Zhu, Y. (1997). A physiological explanation of disparity attraction and repulsion. *Investegative Ophthalmology and Visual Science Supplement (ARVO)*, *38*, 906.

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

Shapley, R., & Lennie, P. (1985). Spatial frequency analysis in the visual system. *Annual Review in Neuroscience*, *8*, 547–583.

Stevenson, S. B., Cormack, L. K., & Schor, C. M. (1991). Depth attraction and repulsion in random dot stereograms. *Vision Research*, *31*, 805–813.

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

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

Westheimer, G. (1986). Spatial interaction in the domain of disparity signals in human stereoscopic vision. *Journal of Physiology*, *370*, 619–629.

Westheimer, G., & Levi, D. M. (1987). Depth attraction and repulsion of disparate foveal stimuli. *Vision Research*, *27*, 1361–1368.

Wurger, S. M., & Landy, M. S. (1989). Depth interpolation with sparse disparity cues. *Perception*, *18*, 39–54.

Zhu, Y., & Qian, N. (1996). Binocular receptive fields, disparity tuning, and characteristic disparity. *Neural Computation*, *8*, 1611–1641.