Self-Organization of Binocular Disparity Tuning by Reciprocal Corticogeniculate Interactions

Alexander Grunewald and Stephen Grossberg
Boston University

Abstract

■ This article develops a neural model of how sharp disparity tuning can arise through experience-dependent development of cortical complex cells. This learning process clarifies how complex cells can binocularly match left and right eye image features with the same contrast polarity, yet also pool signals with opposite contrast polarities. Antagonistic rebounds between LGN ON and OFF cells and cortical simple cells sensitive to opposite contrast polarities enable anticorrelated simple cells to learn to activate a shared set of complex cells. Feedback from binocularly tuned cortical cells to monocular LGN cells is proposed to carry out a matching process that dynamically stabilizes the learning process. This feedback represents a type of matching process that is elaborated at higher visual processing areas into a volitionally controllable type of attention. We show stable learning when both of these properties hold. Learning adjusts the initially coarsely tuned disparity preference to match the disparities present in the environment, and the tuning width decreases to yield high disparity selectivity, which enables the model to quickly detect image disparities. Learning is impaired in the absence of either antagonistic rebounds or corticogeniculate feedback. The model also helps to explain psychophysical and neurobiological data about adult 3-D vision. ■

INTRODUCTION

The rapid processing of binocular disparity information requires highly tuned disparity-selective neural responses, yet at birth infants show only a coarse level of stereopsis (Birch, Gwiazda, & Held, 1983; Blakemore, Hawken, & Mark, 1982; Blakemore & van Sluyters, 1974; Daw, 1994; Daw & Wyatt, 1976; Freeman & Ohzawa, 1992; Held, Birch, & Gwiazda, 1980; Leventhal & Hirsch, 1980; Movshon & Düsteler, 1977; Shimojo, Bauer, O'Connell, & Held, 1986) Here we present a neural model that simulates how cortical complex cells can develop fine disparity tuning starting from coarse tuning. This binocular circuit forms part of a larger theory of binocular vision that has previously been derived to explain data concerning both static (Grossberg, 1994, 1997; Grossberg & McLoughlin, 1997; McLoughlin & Grossberg, 1998: Ohzawa, DeAngelis, & Freeman, 1990) and dynamic (Grossberg & Grunewald, 1995) properties of adult binocular vision. The present model hereby suggests how binocular development helps to select the parameters that are used to explain both behavioral and neural data about adult binocular vision.

Competition across cortical complex cells (Sillito, 1977, 1979) determines a local winner, which can learn the pattern of simple cell activities that feed into the complex cell (Singer, 1983). ON and OFF cells at the retina and LGN (Schiller, 1992) help to ensure that whenever learning links simple cells that are sensitive to one contrast polarity with a complex cell, subsequent learning also links simple cells sensitive to the opposite polarity to the same complex cell. Antagonistic rebounds of activity between ON and OFF cells play a key role in this process. The antagonistic rebound mechanism clarifies how anticorrelated simple cells can become associated with the same complex cell. Complex cells hereby develop to pool opposite polarities of image contrast. This defining characteristic of complex cells (Gilbert, 1977; Hubel & Wiesel, 1962; Skottun et al., 1991) helps the visual system to form object boundaries against textured backgrounds whose contrast relative to that of the object may reverse along the length of the boundary (Grossberg & Mingolla, 1985a, 1985b).

On the other hand, it is known that observers have difficulty when they try to binocularly fuse features from the two eyes that have opposite contrast polarities (Anderson & Nakayama, 1994; von Helmholtz, 1925). The model clarifies how the simple-to-complex cell filter learns to binocularly fuse only stimuli for which both eyes process the same contrast polarity, even as it pools together fused signals from both contrast polarities.

Whenever a complex cell emerges as a winner, it sends a top-down matching, or confirmation, signal to the LGN (Sillito, Jones, Gerstein, & West, 1994; Varela &

Singer, 1987). When the confirmation signal matches the LGN activity pattern, the matched LGN activities are selected. A mismatch between the confirmation signal pattern and the LGN input pattern leads to a reduction of LGN activity. This happens as follows: When there is activity at the cortex, it sends back spatially distributed inhibition. If the excitatory cortical feedback matches the excitatory activity pattern of the retinal input to the LGN, the two activities summate, thus causing a higher level of excitation that does not get quenched by the inhibitory feedback from the cortex. If, on the other hand, the retinal feedforward and the cortical excitatory feedback signals do not match, the broad inhibitory cortical feedback quenches them both. This selective attenuation of mismatched lateral geniculate nucleus (LGN) cells helps to stabilize the learning process (Carpenter & Grossberg, 1991; Grossberg, 1976b) and to trigger selection of a new complex cell winner if the match is bad enough. Thus the model suggests a possible role for corticogeniculate pathways in self-organization of cortical disparity tuning during the developmental critical period. In particular, the simulated adaptive tuning of binocular feature detectors such as complex cells in V1 can be explained by synaptic plasticity that is stabilized by the interaction between feedforward signals from LGN to V1, and feedback signals from V1 to LGN, which in turn may affect the feedforward signals to V1 in a reverberating cycle (Figure 1).

Most of the physiological data cited in this study originates from cat studies. We are interested in modeling development of cortical disparity detectors in humans, which would mean that a better source of neurophysiological data is the monkey, rather than the cat. Indeed there are striking differences at the geniculate and cortical level between cats and monkeys. The main difference is found in layer 4, where in the cat there is orientation selectivity, whereas in the macaque, although layer 4c is populated by oriented cells, the layer 4cβ cells that receive LGN afferents and feed layer 4c are less oriented (Blasdel & Fitzpatrick, 1984; Hawken & Parker, 1984; Leventhal, Thompson, Liu, Zhou, & Ault, 1995). The present study develops a lumped model of cortical development that does not attempt to distinguish these cortical sublaminae.

Learning in the model occurs in both feedforward and feedback pathways when activities converge within the circuit. Such an interaction between feedforward and feedback signals has been shown to control stable learning under rather general circumstances in Adaptive Resonance Theory, or ART (Carpenter & Grossberg, 1991; Grossberg, 1976b, 1980). This analysis provides a rationale for the widespread occurrence of reciprocal thalamocortical and corticocortical connections throughout the brain (Felleman & Van Essen, 1991; Macchi & Rinvik, 1976; Tsumoto, Creutzfeldt, & Legendy, 1978; van Essen & Maunsell, 1983). In ART, top-down feedback stabilizes

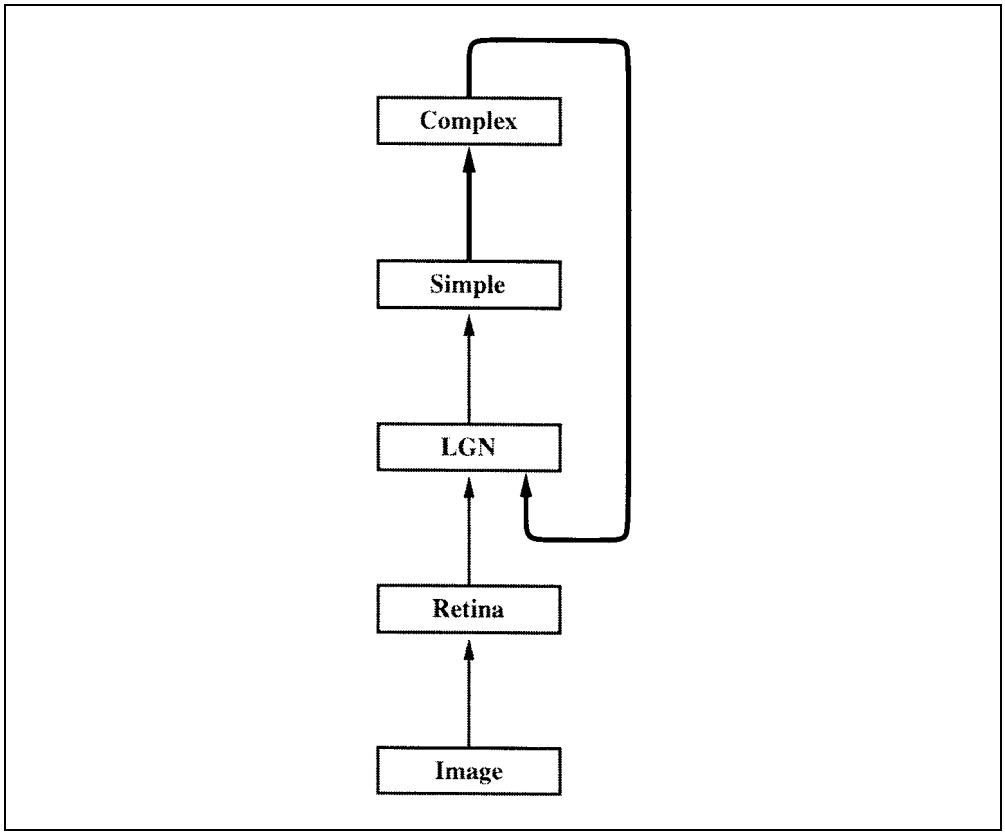

Figure 1. Processing stages of the model. The thicker paths indicate adaptive pathways that contribute to binocular disparity tuning.

adaptive synapses by regulating the gain of these circuits: Bottom-up processing by itself can activate its target circuits, whereas top-down processing by itself can subliminally prime these circuits. When both bottom-up and top-down signals are simultaneously active, cell activities at which these signals converge are preserved and possibly amplified, whereas cell activities that receive only small top-down signals are attenuated. Thus top-down processing can be thought of as a matching, verification, or hypothesis testing operation because the combination of bottom-up and top-down processing selects those bottom-up activations that are consistent with top-down processing, while suppressing those that are not.

Grossberg (1976b) suggested that corticogeniculate feedback carries out such a matching function in order to stabilize the development of cortical binocular tuning during the visual cortical period by selectively matching monocular LGN cell activities that are consistent with binocular cortical activities. Several experiments have supported the existence of a matching process with these properties (Sillito et al., 1994; Varela & Singer, 1987). In addition, Gove, Grossberg and Mingolla (1995) have used a monocular version of the present model to simulate how corticogeniculate feedback influences brightness percepts and the formation of perceptual groupings, such as illusory contours. The neural network model herein simulates how such feedback can stabilize the disparity tuning that occurs at cortical complex cells during the critical period, and it illustrates what can go wrong when it is absent.

The matching process that is proposed to occur in corticogeniculate pathways may be interpreted as a type of automatic attentional focusing. Similar top-down circuits have been incorporated into models of attentive recognition learning in the inferotemporal cortex, where it is suggested how attention may be volitionally controlled in a task-selective way by auxiliary circuits (Carpenter & Grossberg, 1993; Grossberg, 1995; Grossberg & Merrill, 1996). Neurophysiological data on feature-selective attentive matching has been reported in area V4 (Motter, 1994a, 1994b). Reynolds, Chelazzi, Luck, and Desimone (1994) and Reynolds, Nicholas, Chelazzi, and Desimone (1995) have reported neurophysiological evidence that spatial attention in areas V2 and V4 may also use similar attentive mechanisms. Thus the mechanism that is proposed to help stabilize disparity-selective learning in area V1 seems to be used at multiple levels of the visual cortex and beyond.

Although a lot of energy has been expended to understand how orientation selectivity arises, comparatively little effort has gone into understanding how disparity tuning can arise as an infant grows up. In this article, we show how a general theory of neural self-organization can be brought to bear on this question and how the known physiology of the visual system can be understood to contribute to the development of disparity-selective detectors.

A SELF-ORGANIZING MODEL OF DISPARITY TUNING

The model is summarized in Figure 2. It uses the same processing stages, equations, and parameters as in Grossberg and Grunewald (1995), where dynamic properties of adult binocular vision were simulated. The only exception is the breadth of the disparity-sensitive kernels from simple cells to complex cells and from complex cells to the LGN, which start out broad but become sharply tuned through learning.

The model depends upon antagonistic rebound responses between ON and OFF cells at the LGN and simple cell stages to enable anticorrelated cells to converge onto a common complex cell. These responses are driven by antagonistic rebound responses at the retinal stage. The rebounds arise from the organization of model retinal, LGN, and simple cells into opponent pairs of ON and OFF cells in circuits called gated dipoles (Carpenter & Grossberg, 1981; Gaudiano, 1994; Grossberg, 1972, 1976b). A key property of the gated dipole circuit is that it generates a transient antagonistic rebound of activity at an OFF cell after a sustained response in the corresponding ON cell terminates. Ringach, Carandini, Sapiro, and Shapley (1996) have used reverse correlation techniques to detect rebounds in V1 cells that occur after visual stimulation.

Given such opponent circuitry, when a complex cell is first activated, its adaptive weights change in the pathways to it from the simple cells that are active, say dark-light simple cells. When a rebound response occurs, as it does in both retinas, simple cells at the same positions as before, but of opposite polarity, say light-dark, are activated. The circuitry that resets a complex cell after its activating simple cells shut off permits it to remain on for a while after the antagonistic rebound occurs. Thus the same complex cell can adjust its weights to become tuned to the simple cells of opposite polarity that are activated by the rebound. The complex cell hereby learns to be associated with both polarities of simple cell each time it gets activated. As a consequence, the complex cell pools activities from both polarities of simple cells at the same position. This model property provides a new rationale for the experimental observation that oppositely polarized simple cells compete before their half-wave rectified outputs are pooled at complex cells (Ferster, 1988; Liu, Gaska, Jacobson, & Pollen, 1992; Ohzawa et al., 1990).

The equations and parameters of the model will now be given. A 1-D simulation of the model was used. Each model neuron was modeled as a single voltage compartment in which the membrane potential, $V(t)$, was given by

$$C_m \frac{dV(t)}{dt} = -(V(t) - E_{LEAK})g_{LEAK} - (V(t) - E_{EXCIT})g_{EXCIT}(t) - (V(t) - E_{INHIB})g_{INHIB}(t), \quad (1)$$

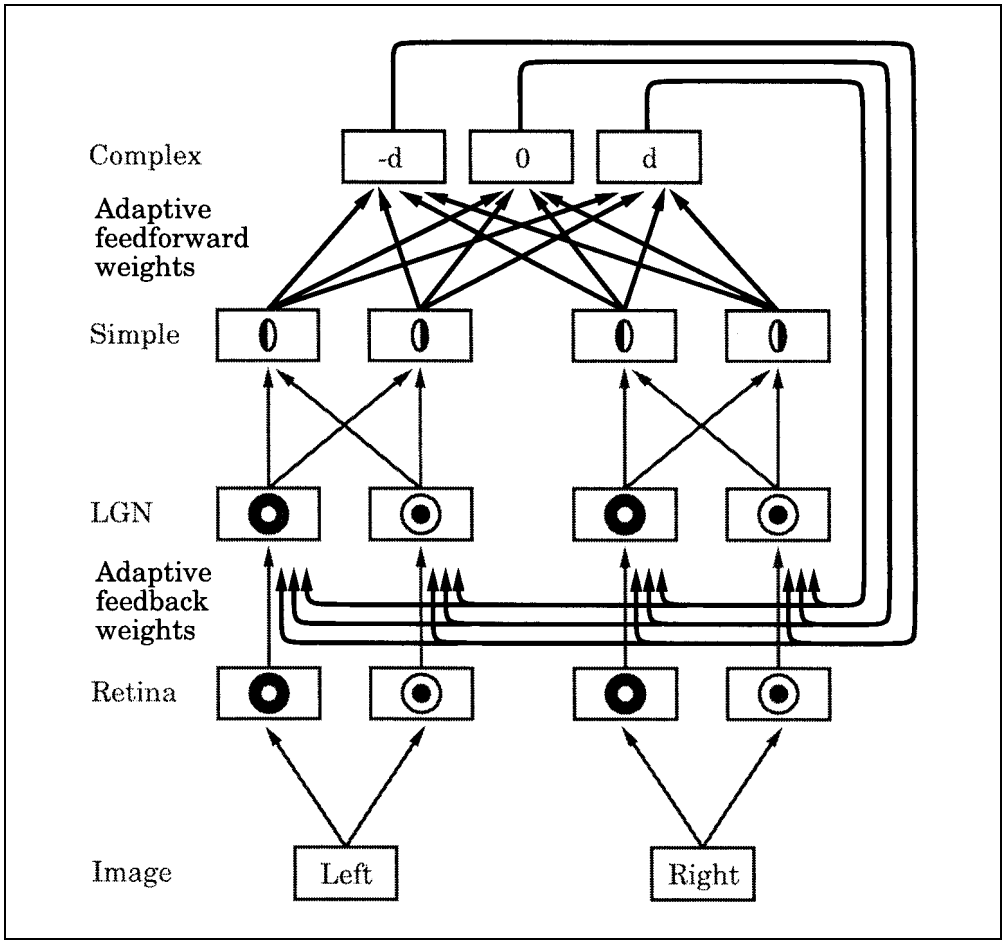

Figure 2. Detailed model architecture. The dark lines indicate pathways that are adaptive and that contribute to the development of binocular disparity tuning.

where the parameters E represent reversal potentials, g_{LEAK} is a constant leakage conductance, and the time-varying conductances $g_{EXCIT}(t)$ and $g_{INHIB}(t)$ represent the total inputs to the cell (Grossberg, 1973; Hodgkin, 1964). Transient after hyperpolarization terms (AHP) were not incorporated because the dynamics of interest operate on the slower time scale of learning. The capacitance term C_m was set equal to 1 by rescaling time t. The leakage reversal potential E_{LEAK} was set equal to 0 by shifting the definition of $V(t)$. With this convention, the inhibitory reversal potential E_{INHIB} is nonpositive. Then Equation (1) can be written in the form

$$\frac{dV}{dt} = -DV + (U - V)g_{EXCIT} - (V + L)g_{INHIB}, \qquad (2)$$

where $D = g_{INHIB}$ is a constant decay rate, $U = E_{EXCIT}$, and $L = |E_{INHIB}|$. Half-wave rectified activities $\max(V, 0)$ are typically passed on as output signals.

Kernels

Dendritic and cell body input weights are modeled using convolution kernels. Most kernels used in this model are Gaussians (except when otherwise indicated); namely,

$$G(y) = k \exp\left(-\frac{y^2}{2\sigma^2}\right), \qquad (3)$$

where σ specifies the size of the kernel. All kernels are normalized so that $\Sigma_y G(y) = 1$, and k is chosen accordingly. The sizes of the kernels have been chosen so that the resulting receptive fields increase in size from the retina to the complex cells. However, we made no attempt to match receptive field sizes with physiological findings. In this study, we consider the development of kernels, not of receptive fields, because the receptive fields are the result of the entire cascade of processing stages and dynamics.

Image

There are two images, left and right. The activities of the left and right retinal images are denoted by I_i^l and I_i^r, respectively. The superscripts are suppressed when both eyes use the same processing.

Retinal Stage

Two stages of retinal processing are modeled. At each stage, there are four fields of neurons at the retinal level

corresponding to the two eyes and the two types of cell, ON cells and OFF cells. The activities r_i at the first level of processing obey

$$\frac{dr_i}{dt} = -Dr_i + (U - r_i)F_i^+ - (L + r_i)F_i^-, \tag{4}$$

where the on-center (F_i^+) and off-surround (F_i^-) feedforward inputs are defined by

$$F_i^+ = \sum_k G_I^+(k - i)I_k \tag{5}$$

and

$$F_i^- = \sum_k G_I^-(k - i)I_k, \tag{6}$$

with $\sigma^+ = 0.3$ and $\sigma^- = 0.9$. This process normalizes retinal responses and computes a measure of image contrast that is Weber-law modulated (Grossberg, 1980, 1983). The output that is passed on to the next level of retinal processing is defined by $P_i = M_p \max(r_i, 0)$. The output of an ON cell is denoted by P^+, and then G_I^+ is a narrow "center" Gaussian kernel, and G_I^- is a wider "surround" Gaussian kernel. The kernels are flipped for the OFF cells, whose outputs are denoted by P^-. The parameters are $D = U = L = 1$.

The second processing stage defines the opponent interaction of a gated dipole. First, a chemical transmitter multiplies, or gates, the output on its way to the next level (Abbott, Varela, Sen, & Nelson, 1997; Francis, Grossberg, & Mingolla, 1994). For each position, there is a transmitter gate g_i that obeys the equation

$$\frac{dg_i}{dt} = R(B - g_i) - D(P_i + T)g_i. \tag{7}$$

In Equation (7), R defines the rate of transmitter accumulation, B defines the maximal level of accumulated transmitter, and $D(P_i + T)$ defines the rate at which the transmitter is inactivated, or habituated, by an input signal P_i. Term $P_i g_i$ says that such inactivation occurs by mass action. Parameter T denotes a background level of activity. This background level of activity may be generated at the previous stage. Its source can be envisioned as intrinsic noise within a neuron. The parameters used are $R = 0.2$, $B = 1$, $D = 2$, and $T = 0.3$. The ON and OFF cells compete to generate the final outputs of the retinal stage:

$$R_i^+ = M_r \max[(P_i^+ + T)g_i^+ - (P_i^- + T)g_i^-, 0] \tag{8}$$

$$R_i^- = M_r \max[(P_i^- + T)g_i^- - (P_i^+ + T)g_i^+, 0] \tag{9}$$

A second upper index indicates which retina a cell belongs to (left or right); thus there are the following variables at this level: R_i^{l+}, R_i^{l-}, R_i^{r+}, and R_i^{r-}, where R_i^{l+} denotes (for example) the output at location i of the ON cell channel in the left eye.

LGN

There are four fields of neurons at the LGN level: two eyes × two polarities (ON or OFF). The equations defining LGN activities are as follows:

$$\frac{dl_i}{dt} = -Dl_i + (U - l_i)(R_i + B_i^+) - (L + l_i)B_i^-, \tag{10}$$

where $D = U = L = 1$. The retinal input R_i has the same polarity and ocularity as the LGN cell in question. The excitatory (B_i^+) and inhibitory (B_i^-) feedback signals from cortical complex cells are given by

$$B_i^+ = M_b^+ \sum_{k,d} w(i, d, k)C_{kd} \tag{11}$$

$$B_i^- = M_b^- \sum_{k,d} C_{kd}. \tag{12}$$

The signals C_{kd} are derived from the complex cell stage. Expression $w(i, d, k)$ denotes the adaptive weight between the complex cell at location k and of disparity d that projects to an LGN cell at position i. There are four such weights for each i, d, and k, one for each of the four LGN fields.

The output signal that is passed on to the simple cells is defined as follows:

$$L_i^+ = \max(l_i^+, 0) \tag{13}$$

$$L_i^- = \max(l_i^-, 0) \tag{14}$$

A second upper index indicates which LGN a cell belongs to (left or right); thus there are the following variables at this level: L_i^{l+}, L_i^{l-}, L_i^{r+}, L_i^{r-}.

Simple Cells

There are four fields of neurons at the simple cell level: two eyes × two contrast polarities (light-dark or dark-light). The responses of simple cells are convolutions of the LGN cell responses with odd-symmetric kernels such that

$$s_i^+ = \sum_k K_s(i - k)L_k^+, \tag{15}$$

and similarly for s^- except that the sign of K_s is reversed.

In this expression, K_s is an odd-symmetric kernel defined as follows:

$$K_s(y) = k \sin(y + 0.5) \exp\left[-\frac{(y+0.5)^2}{2\sigma^2} \right], \quad (16)$$

where $\sigma = 0.3$ gives the width of the kernel, and k normalizes the kernel. In this kernel, y is shifted by 0.5 so that the simple cell is positioned between a pair of LGN cells, thereby achieving good edge localization. Oppositely polarized simple cells compete before their net activity is half-wave rectified to generate an output signal (Ferster, 1988, Gove et al., 1995, Liu et al., 1992), as follows:

$$S_i^+ = M_s \max[s_i^+ + s_i^- - \alpha Abs(s_i^+ - s_i^-), 0] \quad (17)$$

$$S_i^- = M_s \max[s_i^- + s_i^+ - \alpha Abs(s_i^+ - s_i^-), 0] \quad (18)$$

In this expression, the activities from the ON and OFF subregions are added ($s_i^+ + s_i^-$), and a correction term is subtracted [$\alpha Abs(s_i^+ - s_i^-)$] that prevents isolated ON or OFF signals from triggering a simple cell. In Equations (17) and (18), the indices stand for dark-light (+) and light-dark (–) contrast polarities, $M_s = 2$ and $\alpha = 1.3$. Another upper index is added to denote the eye of origin (l or r). Thus at this level there are the following activities: S_i^{l+}, S_i^{l-}, S_i^{r+}, and S_i^{r-}.

Complex Cells

At the complex cell stage, there are three fields of complex cells: one each for zero uncrossed (far) and crossed (near) disparities. The disparities of the immature model were 0, −1, 1. A disparity of −1 means that the left image has been shifted by −1 (1 to the left), and the right image by 1 (1 to the right). Associated with each complex cell is also an inhibitory interneuron (Hirsch & Gilbert, 1991; McGuire, Gilbert, Rivlin, & Wiesel, 1991). The equations for the excitatory complex cells c_{id}^+ and the inhibitory interneurons c_{id}^- are as follows:

$$\frac{dc_{id}^+}{dt} = -Dc_{id}^+ + (U - c_{id}^+)(F_{id}^+ + B_{id}^+) - (L + c_{id}^+)(F_i^- + B_{id}^- + \beta c_{id}^-) \quad (19)$$

$$\frac{1}{\delta}\frac{dc_{id}^-}{dt} = -Dc_{id}^- + (U - c_{id}^-) f(c_{id}^+) - (L + c_{id}^-) F_i^- \quad (20)$$

where $D = U = L = 1$. The parameter $\beta = 20$ denotes the interneuron strength. It is chosen so that activation of the inhibitory interneuron in the absence of simple cell activity leads to inhibition of complex cells. This prevents undue persistence of complex cell activation, as a result of the positive feedback B^+, after inputs shut off. The parameter $\delta = 0.5$ ensures that the inhibitory interneuron is slower than the complex cell.

The feedforward signals from the simple cells are given by

$$F_{id}^+ = M_c^f Abs\left[\sum_k b_{l-}(i, d, k) S_k^{l-} + \sum_k b_{r-}(i, d, k) S_k^{r-} - \sum_k b_{l+}(i, d, k) S_k^{l+} - \sum_k b_{r+}(i, d, k) S_k^{r+} \right] \quad (21)$$

$$F_i^- = M_c^f Abs\left[\sum_k G_{s_{l-}}^-(k - i) S_k^{l-} + \sum_k G_{s_{r-}}^-(k - i) S_k^{r-} - \sum_k Gsubs_{l+}^-(k - i) S_k^{l+} - \sum_k G_{s_{r+}}^-(k - i) S_k^{r+} \right], \quad (22)$$

where Abs denotes the absolute value and $M_c^f = 2$ is the strength of feedforward activities. For complex cells at zero disparity ($d = 0$), the feedforward weight is scaled by a factor of 1.05. This scaling factor models a bias toward zero disparities; in other words, if the input is ambiguous, the model will choose a zero-disparity complex cell.

The feedforward inhibitory kernels G^- are defined by $\sigma^- = 5$. They are not disparity tuned. The feedforward excitatory kernels b are disparity tuned by developmental learning. Expression $b_{l-}(i, d, k)$ denotes the kernel between simple cells of type S^{l-} at location k and a complex cell of disparity d at position i and similarly for the other kernels b_{r-}, b_{l+}, and b_{r+}. The difference within the absolute value expression ensures that the argument of the absolute value operation is maximal only when simple cells of the same polarity are active in the two eyes (Grossberg & McLoughlin, 1997; Ohzawa et al., 1990). If the polarities differ, the difference ensures a weak signal. The absolute value operation, on the other hand, ensures that the feedforward signal does not depend on which polarities the simple cells have. In other words, feedforward activities are designed so that only simple cell activities of the same polarities can fuse, but at the same time the complex cell pools signals from opposite contrast polarities.

The feedback activities in Equation (19) are given by

$$B_{id}^+ = M_c^b \sum_{j,e} G_c^+(j - i) f(c_{je}^+) \quad (23)$$

$$B_{id}^- = M_c^b \sum_{j,e} G_c^-(j - i) f(c_{je}^+), \quad (24)$$

where $M_c^b = 300$ is the strength of feedback interactions. Feedback activities are also not disparity tuned.

It is important to note that here we are only interested in the development of the disparity tuning of complex cells, and therefore it suffices to study their input kernels. As pointed out above, the kernel in the model is not the receptive field. The receptive field that neurophysiologists measure is the aggregate of all processing stages and dynamics that lie between the stimulus and the cell under study. With this in mind, it is easy to understand that although the kernels we show have similar shapes, independent of position (i.e., disparity preference), the resulting receptive fields are not the same. A zero-disparity cell, for example, will respond for any stimulus of disparity 1, 0, and −1, whereas a far disparity cell of preference −3 will respond to any stimulus of disparities −2 and smaller (and analogously for a near disparity cell of preference +3). Thus the disparity tuning curves for zero-disparity cells are much narrower than for nonzero disparities, as has been found physiologically (Poggio & Fischer, 1977). This is discussed further in related studies using this model of daVinci stereopsis (Grossberg & McLoughlin, 1997), dynamic disparity processing (Grossberg & Grunewald, 1995), and binocular fusion (McLoughlin & Grossberg, 1998).

In a two-dimensional implementation, spatial competition sharpens cell responses close to the ends of lines before the resulting end-stopped complex cells generate feedback to the LGN (Gove et al., 1995; Murphy & Sillito, 1987; Sillito et al., 1994; Weber, Kalil, & Behan, 1989). Because the present model is only a one-dimensional implementation, this kind of competition has no effect on simulated values. Here, end-stopped complex cell responses are a replica of complex cell responses. The feedback signals to the LGN in Equations (10) and (11) are thus defined by

$$C_{id} = f(c_{id}^{+}). \tag{25}$$

Bottom-Up Learning

Four different excitatory feedforward kernels b that are disparity selective are associated with each complex cell in Equation (19), one for each of the four simple cell types from which it receives input (S_{l+}, S_{r+}, S_{l-}, and S_{r-}). These kernels are b_{l+}, b_{r+}, b_{l-}, and b_{r-}. Each of these kernels is convolved with the corresponding simple cell type. By virtue of a disparity difference coded within those kernels, the input to the complex cell field is disparity tuned.

For each complex cell, each of these kernels has to develop from coarse initial conditions to a sharper level of disparity selectivity. This can be achieved if learning occurs. Whenever a complex cell has emerged as a winner, as in self-organizing feature maps (Grossberg, 1976a; Kohonen, 1984; von der Malsburg, 1973), the adaptive weights that make up the kernel are adjusted through a learning law to mimic the activities at the simple cell stage. Instar learning (Grossberg, 1976a; Singer, 1983) is such a learning rule:

$$\frac{db(i, d, k)}{dt} = \varepsilon[C_{id} - \Gamma]^{+}[S_k - b(i, d, k)]. \tag{26}$$

Here $b(i, d, k)$ denotes the adaptive weight, S_k is the signal from the simple cell to the complex cell, and $[C_{id} - \Gamma]^{+} = max(C_{id} - \Gamma, 0)$ is an activity-dependent complex cell signal that gates learning on and off. Parameter Γ acts as a learning threshold. Only when complex cells exceed it does learning occur. Its value is 0.3. The factor ε ensures that learning occurs at a slower pace than the integration of the cell activations. Its value is 0.05.

As mentioned above, there are four populations of simple cells that converge upon each complex cell, and for each population there is a separate set of weights $b(i, d, k)$. Due to the absolute value operation in Equation (21), the complex cell is insensitive to the polarity of simple cell responses. This means that the complex cell responses are symmetric with respect to the polarity, and therefore the same learning rule in Equation (26) can be used for both polarities.

Instar learning is an appropriate rule in the present context because complex cells combine signals from several different populations of simple cells. Although a complex cell might not know which simple cell activated it, because there are separate kernels for each simple cell, instar learning enables each kernel to be selectively enhanced by activation of the corresponding simple cell.

Top-Down Learning

Each complex cell in Equation (11) projects to four different populations of LGN cells, the ON and the OFF cells in the left and the right LGNs with output signals L^{l+}, L^{r+}, L^{l-}, and L^{r-}. A different set of weights is associated with each LGN cell type and with each complex cell. After learning, complex cells are supposed to selectively project to the LGN so as to strengthen those LGN cells that led to its firing. Hence a good learning rule is capable of learning what LGN cells are active during learning. Such a learning rule is called outstar learning (Grossberg, 1968, 1974). In the present instance,

$$\frac{dw(i, d, k)}{dt} = \varepsilon[C_{id} - \Gamma]^{+}[L_k - w(i, d, k)], \tag{27}$$

where $w(i, d, k)$ is the top-down adaptive weight in Equation (11). The parameters Γ and ε are described above.

As noted above, learning in response to a complex

environment is stabilized if both top-down and bottom-up learning occur in parallel. The simulations in the next section show that stable learning occurs when both bottom-up and top-down kernels learn in parallel. The subsequent section describes how learning can break down if there are no rebound responses and if corticogeniculate feedback is abolished.

SIMULATION RESULTS

ART was initially developed to explain how stable learning could take place in a changing environment. In the simplest ART circuit, there are two levels, Level 1 and Level 2. Level 1 receives exogenous inputs but also endogenous feedback from Level 2. Level 1 sends its output to Level 2, where the best match is determined through a winner-take-all operation that selects the best-matched category representation of the Level 1 output pattern. The winning Level 2 category then sends a confirmation signal pattern to Level 1. At Level 1 a matching operation takes place, such that activity at Level 1 only persists if feedback from Level 2 and the exogenous input match.

ART relies on the exclusive use of local processing (Grossberg, 1976b). One of the key elements of local processing is that all neural operations be performed by individual neurons, or neuron populations, that have access only to gated inputs from other neurons that can excite or inhibit them via membrane equation dynamics. All sophisticated model capabilities are emergent properties of the resulting neural network. It can be shown that a relatively simple on-center off-surround feedback neural network acts like a winner-take-all circuit using only local operations (Grossberg, 1973). In that study, it was also demonstrated how uniform activity at a stage could be quenched. When two input patterns converging on the same network do not match, a uniform distribution of activity is the result, which then is quenched. Based on those results, it appeared likely that a more complex theory, such as ART, would be able to use these properties to achieve stable learning.

As described above, ART has feedback between Levels 2 and 1 (to achieve the matching operation), and there is feedback within Level 2 (to determine the winner). It is possible that feedback between different processing levels might interfere with the ability to pick a global winner, which is an important part of the stabilization of learning process. In many previous studies, ART was simulated by assuming that some of the model equations could be solved at equilibrium, thereby removing the problem of the interaction between multiple types of feedback. Here we show that an entirely dynamically simulated system can stably learn sharp disparity tuning. This shows that, in spite of multiple levels of feedback and similar time courses at each processing stage, the network can currently perform both the ART matching operation at Level 1, and the winner-take-all operation at Level 2, to stabilize the disparity learning process. Although we cannot yet provide a proof to establish under which conditions stable learning occurs, it is clear from our simulations that these properties are robust with respect to modest parameter changes.

Normal Development of Complex Cells

Learning in the present context solves two problems. The first problem is to obtain disparity preference from initial conditions in which only a minor preference is present. This includes development of the ability to fuse bigger disparities, starting from the ability to fuse small disparities. The second problem is to sharpen the tuning of the kernels starting from very broadly tuned kernels. This leads to sharpening of disparity selectivity starting from initial conditions with coarse selectivity. This latter property allows quicker fusion to occur, because more of the selection process is taken over by the feedforward kernels and is not dependent upon the recurrent dynamics of the complex cell field.

The same integration procedure was used as in the model described previously (Grossberg & Grunewald, 1995; Grunewald, 1995). Each field of neurons has 100 cells. The units were arranged in a ring so that no problems occur due to edge effects. All differential equations were integrated using the fourth-order Runge–Kutta method, with a step size of $H = 0.01$. Update of the network was performed so that only values from the previous processing time step were used in calculations. Simulations were implemented as a *C* program running on Sun and SGI workstations. Table 1 summarizes all the parameters that were used in the simulations. The parameter $\varepsilon = 0.05$ in equations (26) and (27) ensures that learning proceeds at a slower rate than cell activation, which has rate 1.

There are $3 \times 100 = 300$ complex cells, and each kernel has 17 units. Each complex cell has a separate kernel for each of four simple cell fields, and each complex cell has a separate kernel for each of four LGN cell fields. Thus in total there are $2 \times 4 \times 17 \times 300 = 40{,}800$ weights that are learned.

To speed up the learning process, a stimulus of size 20 was used. The stimulus was presented to each location for 8.00 time units. Then the stimulus was shifted by 40 neurons (modulo 100, which is the size of the network), and the next disparity was picked by the complex cell stage. In that way the network cycled through five locations and three disparities. After $5 \times 3 \times 8 = 120$ time units, each location had been visited twice, once for each polarity (dark-light and light-dark). In the simulations discussed in this section, this procedure was repeated 10 times, thus taking 1200.00 time units, which consumed about 4 to 5 hr of computation time on a SGI indigo 2 workstation.

Table 1. The Parameters Used in the Binocular Model before Learning

Parameter	*Value*	*Description*
D	1	Passive decay constant of cell activity
U	1	Upper limit of cell activity
L	1	Lower limit of cell activity
R	0.2	Rate of gate recovery
B	1	Baseline gate activity
D	2	Active gate decay
T	0.3	Background activity
σ_l^+, s_r^-	0.3, 0.9	Width of retinal kernels
σ_l^-	100	Width of corticogeniculate feedback kernels
σ_s	0.3	Width of simple cell kernels
σ_c^{f-}	5	Width of feedforward complex cell kernels
σ_c^{b+}, σ_c^{b-}	0.3, 4	Width of feedback complex cell kernels
M_p	10	Multiplicative factor of gate activities
M_r	200	Multiplicative factor of retinal activities
M_s	2	Multiplicative factor of simple cell activities
M_c^f	2	Multiplicative factor of feedforward complex cell activities
M_c^b	300	Multiplicative factor of feedback complex cell activities
M_b^+, M_b^-	10, 1	Multiplicative factor of corticogeniculate feedback
α	1.3	Simple cell threshold
β	20	Weight of complex cell inhibitory interneuron
Γ	0.3	Learning threshold
ε	0.05	Learning rate

Initial Conditions

The disparity preference of the bottom-up kernels b in Equation (21) of the three complex cell fields was put to −1, 0, and 1, respectively, at the beginning of the program by calculating shifted kernels for each complex cell field. These disparities are significantly smaller than the ones used previously to simulate adult disparity processing data (Grossberg & Grunewald, 1995), where preferences of −3, 0, and 3 were used. The same is also the case for the top-down kernels.

The tuning width of the bottom-up kernels was initially set to $\sigma = 2.5$, which is substantially wider than the $\sigma = 0.3$ that was used to simulate adult data. The tuning width for the top-down kernels was set to $\sigma = 6$. In other words, before learning, the top-down kernels are untuned and the bottom-up kernels are broadly tuned. We show here how the process of learning of the bottom-up kernels proceeds in parallel with learning of the top-down kernels without requiring that the two types of kernel have any a priori similarities. These initial conditions were chosen so that the system was initially coarsely tuned, but some residual capability was available for symmetry-breaking purposes. Without it, all complex cells would respond equally. No differentiation between different disparities can occur if there are no initial biases present.

Analysis of Learning

The large number of adaptive weights restricts the way in which learning can be analyzed. Three methods were used: analyzing the development of individual kernels over time, and tracking the mean $\mu = \sum_x x g(x)$ and standard deviation $\sigma^2 = \sum_x x^2 g(x) - \mu^2$ of all kernels across the network. Here, $g(x)$ is the strength of a particular kernel value divided by the sum of all kernel values for normalization purposes. The mean gives information about the disparity preference, and the standard deviation provides information about the tuning width.

Figure 3 shows the development of individual bottom-up kernels of a complex cell that initially has a slight preference for far, or uncrossed, disparities. The kernels develop to strengthen the initial disparity preference and to match it to the numerical disparities of the model's environment, in the present case a disparity of −3. This learning process manifests itself in two ways. First, the peaks of the kernels move in opposite directions, depending on whether a given kernel is convolved with left or with right simple cells. Second, the width of the kernels decreases over time. Both polarities of kernels (dark-light, or DL, and light-dark, or LD) learn in parallel. The complex cell continues to pool signals from simple cells of opposite polarity and thus, in this sense, remains insensitive to image contrast polarity.

At the same time that feedforward learning occurs, feedback learning is also taking place. This is shown in Figure 4. Once again the kernel develops a disparity preference for the stimuli with which the model is presented, and the tuning narrows. Each edge in the image causes spatially offset, and temporally simultaneous, ON and OFF responses. The rebound responses mean that each stimulus presentation causes ON and OFF responses at the same location in close temporal succession. From this it follows that after each presentation of an edge, both ON and OFF responses will have occurred at each of the two locations adjacent to the edge. In other words, both sides of the edge need to be learned by the top-down kernel. As a result, the kernel has no single peak; it has a narrow plateau in the middle. This is in fact the case, as is shown in Figure 4.

Figure 5 shows that not only individual bottom-up kernels but all kernels that were stimulated in the present simulation developed a strong disparity preference. Initially, the kernels were mildly tuned for far, zero, and near disparities with means of −1, 0, and 1. Over the course of learning these preferences move toward −3,0, and 3. Two observations need to be made. First, the initial disparity preference (far, zero, near) is never altered. Second, the emerging preferred disparities are precisely those disparities that are used in the stimuli. In other words, the disparities of the neural network adapt to those that are present in the environment. The same is also true for top-down disparities. Figure 6 shows learning by the top-down kernels.

Figure 5 also shows how the width of the bottom-up kernels changes over time. Note that from the very beginning of learning, the kernels become narrower. Note also that the kernels corresponding to the zero-disparity case get narrower quicker than the other kernels. This is mainly because those kernels do not need to move before narrowing takes place. Thus, as a result of self-organization, zero-disparity complex cells in the model have more narrowly tuned kernels than far or near tuned complex cells. Although this is not the only factor affecting disparity tuning curves, the model suggests that zero-disparity complex cells are more narrowly tuned throughout development. Figure 6 shows

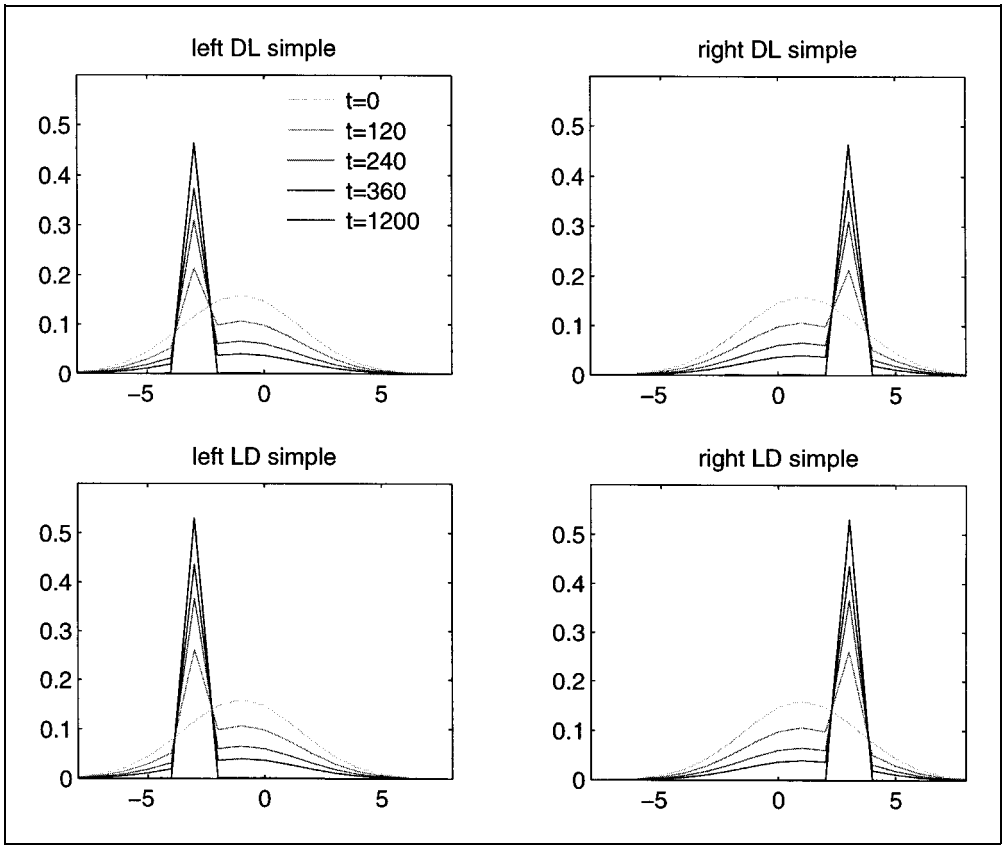

Figure 3. The development of individual bottom-up kernels to a complex cells of far, or uncrossed, disparity preference. The top panels show the kernels between DL simple cells and the complex cell. On the left is shown the kernel that connects the left DL simple cells with the complex cell, and on the right is shown the kernel that connects the right DL simple cells with the complex cell. The bottom panel shows the kernels between LD simple cells and the same complex cell. Note that the DL and LD kernels are indistinguishable. Over time the kernels become narrower, and their preference shifts away from the central, zero-disparity location.

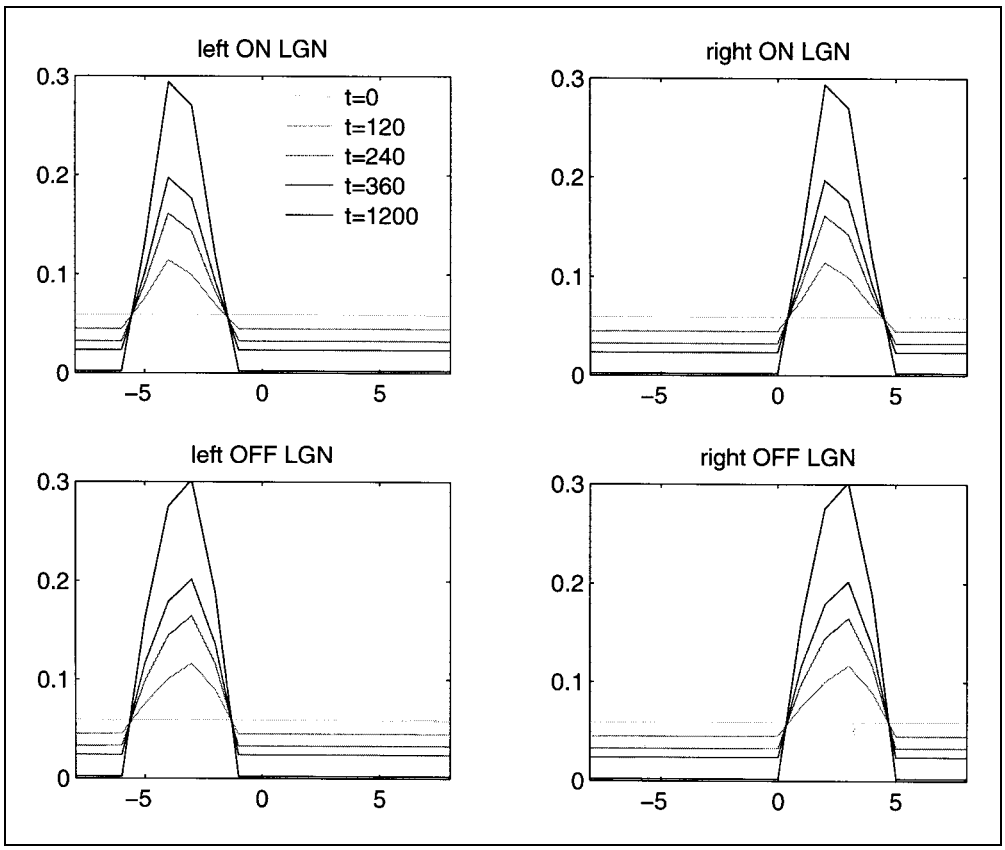

Figure 4. The development of individual top-down kernels from a complex cell of far, or uncrossed, disparity preference. It is the same complex cell as in Figure 3. The top panels show the kernels between the complex cell and ON LGN cells. On the left is shown the kernel between the left ON LGN cells and the complex cell, and on the right the kernel between the right LGN cells and the complex cell.

how the width of the top-down kernels changes over time.

Impaired Development of Complex Cells

This section shows that two key design components of the model are needed to obtain stable learning: the existence of rebound responses and corticogeniculate feedback.

Learning with No Rebound Responses

Four bottom-up kernels b are associated with each complex cell: one for each eye and one for each contrast polarity. Binocular activation of a complex cell occurs only when both eyes received the same type of edge. The learning law specified above means that learning takes place between the complex cell and the active simple cells, as occurs in long-term potentiation (LTP), whereby postsynaptic activity gates an increase of synaptic strength (Frégnac, Burke, Smith, & Friedlander, 1994; Kirkwood & Bear, 1994a). However, learning also occurs between the complex cell and the inactive simple cells. This property realizes a type of long-term depression (LTD), in which postsynaptic activity gates the decrease of synaptic strength when no presynaptic signal is present (Artola & Singer, 1993; Frégnac et al., 1994; Grossberg, 1976a, 1980; Kirkwood & Bear, 1994b; Singer, 1983). Without a rebound response, learning occurs only for simple cells of one polarity, whereas selectivity for the opposite polarity is gradually lost. The rebound response ensures that immediately after one polarity is learned, the other polarity starts to also be learned. Thus, whereas one polarity is initially learned via LTP, and the other forgotten via LTD, during the rebound response the latter is learned and the former forgotten. Learning proceeds at a faster rate than forgetting, ensuring that the net effect to a single stimulus presentation is a net increase in selectivity for both polarities. Figure 7 shows that, in the absence of a rebound response, the complex cells develop the kernel corresponding to only one subpopulation of simple cells, whereas the other kernel decays away, thus rendering the complex cell sensitive to the polarity of contrast in the image. This would disrupt processing at later stages of visual processing, where pooling of opposite contrast polarities is needed to build long-range boundary groupings of object boundaries on textured backgrounds.

Learning with No Corticogeniculate Feedback

Through corticogeniculate feedback, the cortex sends an expectation signal back to the LGN stage, and it can thereby shut off a pattern that was incorrectly classified by the complex cell stage. On the other hand, if the LGN pattern matches the feedback pattern, an enhancement

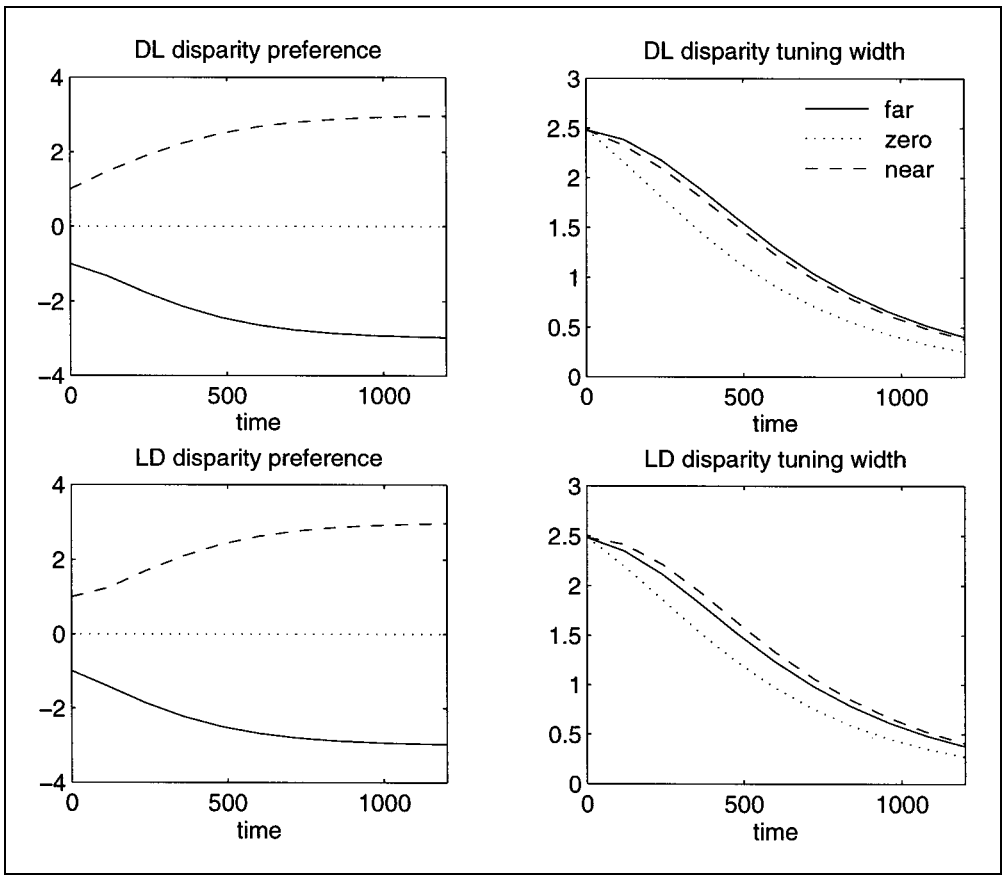

Figure 5. The development of bottom-up complex cell kernels over time. In this figure kernels from all different positions have been superimposed. The top panels show the kernel statistics of the kernels between left DL simple cells and complex cells. On the left the disparity preference is shown and on the right the disparity tuning width is shown (for precise definitions see text). The bottom panels show the same statistics for kernels between left LD simple cells and complex cells.

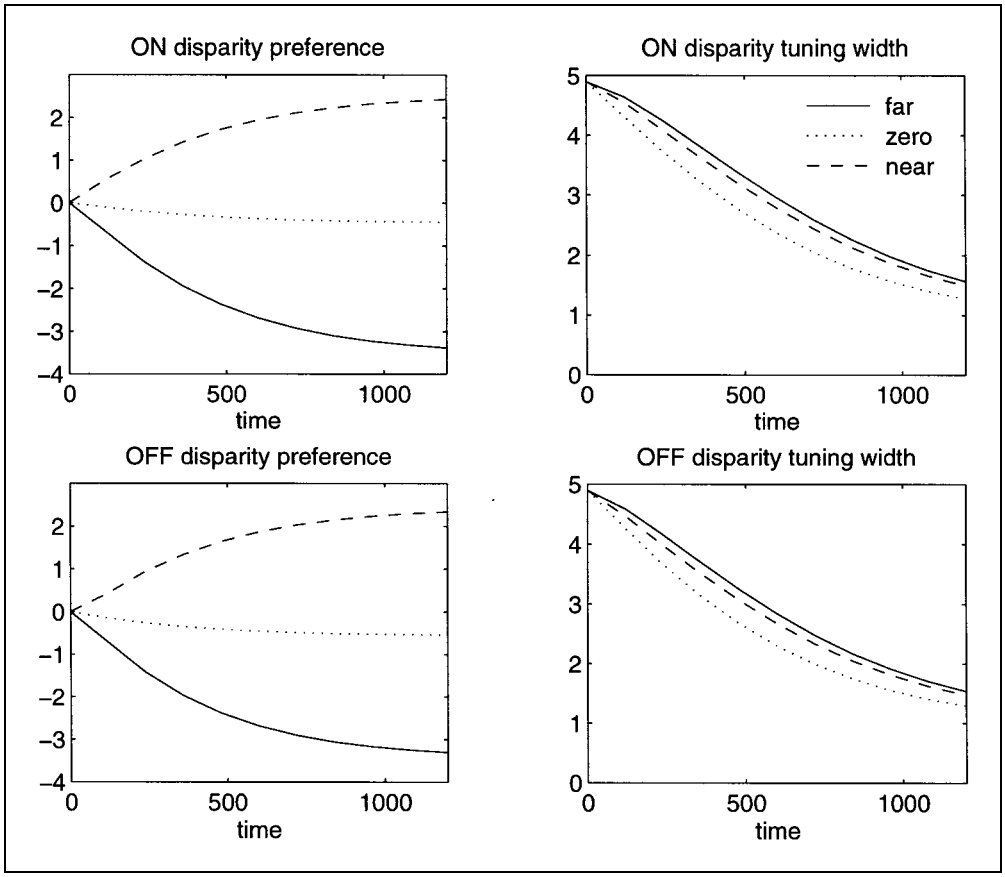

Figure 6. The development of top-down complex cell kernels over time. The top panels show the kernel statistics of the kernels between complex cells and left ON LGN cells. The bottom panels show the same statistics for kernels between complex cells and left OFF LGN cells. Note that all kernels are shifted by 0.5 in comparison to the bottom-up kernels.

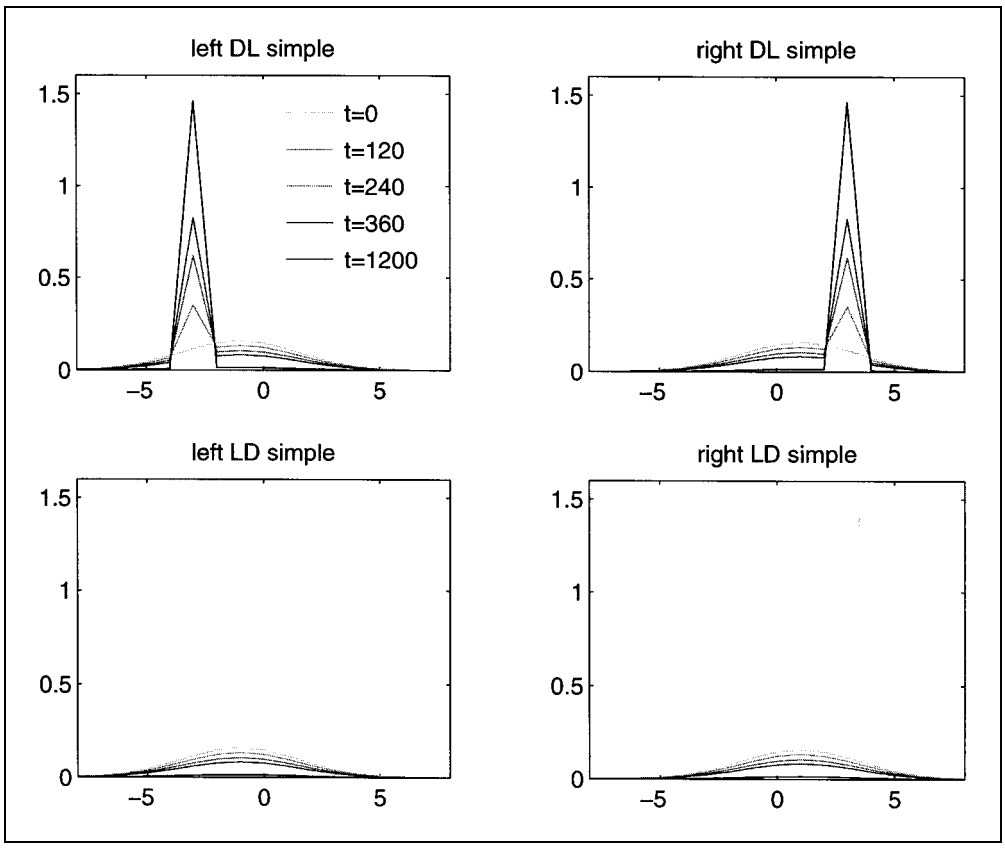

Figure 7. In the absence of rebound responses, complex cells learn to receive input from simple cells of only one polarity (the first one to which they happen to respond), whereas the input from simple cells of the other polarity decays away. As a consequence, complex cells become highly sensitive to the polarity of contrast. The kernels shown in this figure belong to the same complex cell as the kernels shown in Figures 3 and 4. The kernels for one polarity are decaying away, and the oppositely polarized kernels are learned as before.

of LGN activities can take place. In the absence of feedback, this strengthening effect does not take place, and hence the second time a pattern is presented, a different winner could emerge, which might not be corrected in the absence of confirmative feedback. When this happens, a recoding of the kernels could occur. This is shown in Figure 8.

DISCUSSION

This study models how disparity selectivity can arise in a neural network model of binocular development. The model shows how complex cells can develop that binocularly fuse edges of the same contrast polarity, that pool fused signals from opposite contrast polarities, and that are highly disparity selective. The disparity preference that emerges matches the disparities present in the model's environment. The ability of complex cells to pool signals from opposite contrast polarities has proven to be important in models of long-range boundary grouping and texture segregation (Grossberg, 1994; Grossberg & Mingolla, 1985a, 1985b; Grossberg & Pessoa, 1998). This property clarifies how the brain builds boundaries around objects whose relative contrast with respect to their backgrounds can reverse along their perimeters. Because the output signals of the oppositely polarized simple cells that converge on complex cells are half-wave rectified (e.g., Equations (17) and (18)), complex cells, in effect, perform an oriented full-wave rectification of the input image. This property has become commonplace in recent models of texture segregation (e.g., Chubb & Sperling, 1989; Grossberg & Mingolla, 1985b; Sutter, Beck, & Graham, 1989). Thus the present study clarifies how this key texture segregation property is learned along with the equally crucial property for depth perception of fusing only like-polarity contrasts.

In order for development to take place, the model needs to have some innate but coarse disparity tuning. This works as a symmetry-breaking mechanism because it ensures that the entire network does not respond equally to any given stimulus. During the course of development these biases are enhanced dramatically by intracortical recurrent excitation and inhibition in the model to yield high disparity selectivity.

Complex cells in the model can learn to pool opposite contrast polarities using the antagonistic rebound that occurs at the offset of a visual stimulus. Ringach et al., (1996) have reported rebound phenomena using reverse correlation techniques to analyze orientational tuning in neurons of area V1. Rebound responses have heretofore been used to explain data about visual aftereffects (Francis & Grossberg, 1996; Grossberg, 1976b; Grunewald & Lankheet, 1996), persistence (Francis et al., 1994), and binocular rivalry (Grossberg, 1987). The adaptive role for rebounds that is envisaged here specializes the general

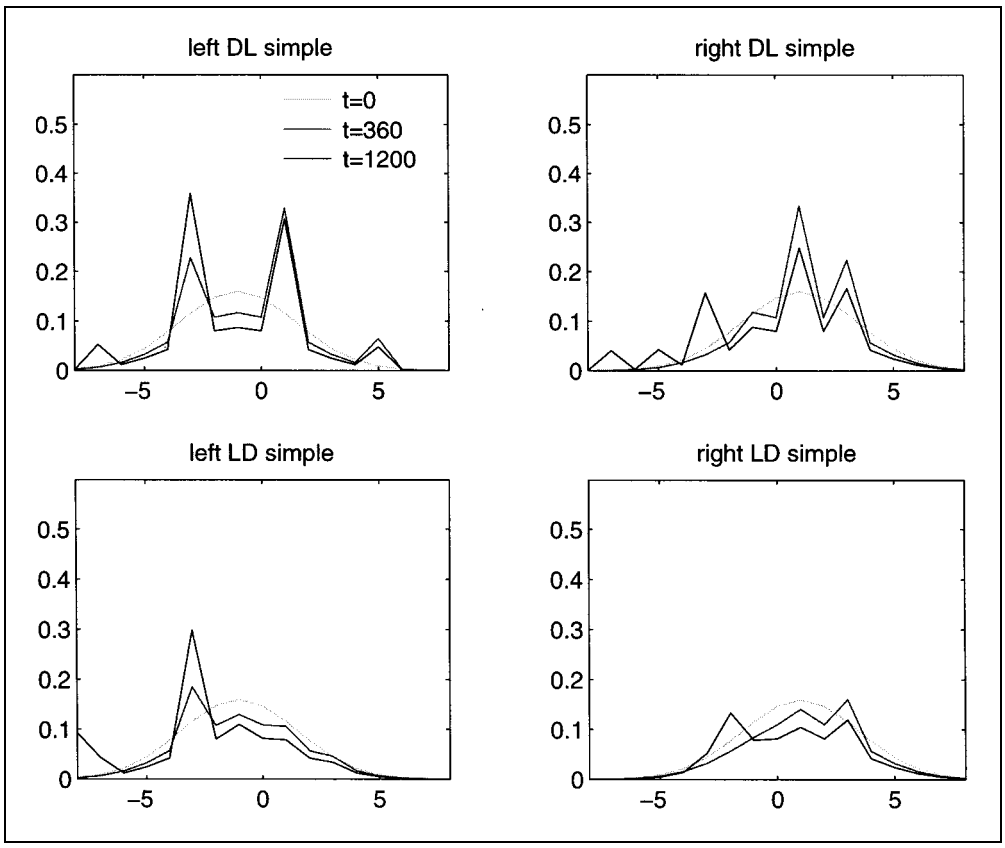

Figure 8. In the absence of corticogeniculate feedback, no matching between feedback and feedforward signals takes place, which opens the possibility for unstable recoding of the disparities learned by the kernels. In this figure, the kernels learn the leftward shift, but due to recoding, they also learn the (incorrect) rightward shift. In other words, this complex cell loses its disparity selectivity.

concept from ART that rebounds between ON and OFF cells help to stabilize the learning process (Grossberg, 1976b, 1980).

A second important feature of the model, which also specializes ART concepts, is that a top-down attentive matching process stabilizes learning. In the present example, this general concept is realized as a corticogeniculate matching process that stabilizes complex cell disparity learning. Similar top-down attentive matching properties are found in data on spatial attention in areas V2 and V4 (Reynolds et al., 1994, 1995) and on feature-selective attention in areas V4 and IT (Carpenter & Grossberg, 1993; Desimone, 1992; Grossberg, 1995; Motter, 1994a, 1994b).

The present model helps to explain how human infants can rapidly develop disparity selectivity from an initial stage at which they have only limited stereoscopic capabilities (Birch et al., 1983; Held et al., 1980; Shimojo et al., 1986). Learning within the model also occurs rapidly, with 10 stimulus presentations within the model being enough to achieve learning.

The model builds on the assumption that opponent pairs of simple cells exist that are tuned to similar orientations but opposite contrast polarities, as has been found in several neurophysiological experiments (Ferster, 1988; Liu et al., 1992; Ohzawa et al., 1990). Olson and Grossberg (1996) have modeled how pairs of simple-cell receptive fields can develop with similar orientational tuning but opposite contrast polarity preference. This developmental process goes on while the model develops cortical maps of orientation and ocular dominance that exhibit the characteristic singularities, linear zones, and fractures that are observed *in vivo* (Blasdel, 1992a, 1992b). This model uses mechanisms like those used in the present work. Taken together, these studies suggest how the stream of LGN-to-simple-to-complex cell receptive fields develops in a coordinated fashion to achieve sharp disparity tuning in adults.

This total set of connections has earlier been used to explain psychophysical data about adult 3-D vision and figure-ground separation (Grossberg, 1994, 1997), as well as recent data of McKee and her colleagues (McKee, Bravo, Taylor, & Legge, 1994; McKee, Bravo, Smallman, & Legge, 1995; Smallman & McKee, 1995) about contrast-sensitive binocular matching, Panum's limiting case, and dichoptic masking (McLoughlin & Grossberg, 1998). Related modeling work (Grossberg & Grunewald, 1995) has shown how key dynamic properties of human binocular vision may be explained by these interactions. In particular, the inability of humans to fuse simultaneous anticorrelated stereograms together with their ability to fuse delayed anticorrelated stereograms (Cogan, Lomakin, & Rossi, 1993; Julesz, 1960) illustrates several model properties; namely, complex cells fuse only like-contrast polarities, pool fused signals from opposite contrast polarities, and are activated by opponent simple

cells that can experience an antagonistic rebound at stimulus offset that activates the opposite contrast polarity cells. Simultaneous anticorrelated stereograms cannot be fused because only like-contrasts fuse. Delayed anticorrelated stereograms can be fused because the antagonistic rebound reverses contrast polarity, so the delayed response can be fused with the later response. Similarly, a simulation of the Pulfrich effect (Julesz & White, 1969; Pulfrich, 1922) depends upon the model's sharp disparity tuning, whereas a simulation of binocular summation (Andersen & Movshon, 1989; Cogan, Clarke, Chan, & Rossi, 1990; Westendorf, Blake, & Fox, 1972) illustrates the importance of using self-normalizing kernels with balanced excitation and inhibition. Taken together, these studies show how adult psychophysical and neurobiological data can be explained as consequences of the developmental mechanisms that are modeled herein.

Acknowledgments

This research was supported in part by the Air Force Office of Scientific Research (F49620-92-J-0334 and F49620-92-J-0225), the Defense Advanced Research Projects Agency (ONR N00014-92-J-4015), and the Office of Naval Research (N00014-91-J-4100) (A. Grunewald) and by the Office of Naval Research (N00014-95-1-0657 and N00014-95-1-0409) (S. Grossberg). The authors wish to thank Diana Meyers for her valuable assistance in the preparation of the manuscript.

Requests for reprints should be sent to Stephen Grossberg, Department of Cognitive and Neural Systems, Boston University, 677 Beacon Street, Room 201, Boston, MA 02215.

REFERENCES

Abbott, L. F., Varela, J. A., Sen, K., & Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science, 275,* 220–224.

Andersen, P. A., & Movshon, J. A. (1989). Binocular combination of contrast signals. *Vision Research, 29,* 1115–1132.

Anderson, B. L., & Nakayama, K. (1994). Toward a general theory of stereopsis: Binocular matching, occluding contours, and fusion. *Psychological Review, 101,* 414–445.

Artola, A., & Singer, W. (1993). Long-term depression of excitatory synpatic transmission and its relationship to long-term potentiation. *Trends in Neurosciences, 16,* 480–487.

Birch, E. E., Gwiazda, J., & Held, R. (1983). The development of vergence does not account for the onset of stereopsis. *Perception, 12,* 331–336.

Blakemore, C., Hawken, M. F., & Mark, R. F. (1982). Brief monocular deprivation leaves subthreshold synaptic input on neurones of the cat's visual cortex. *Journal of Physiology (London), 327,* 489–505.

Blakemore, C., & van Sluyters, R. C. (1974). Reversal of the physiological effects of monocular deprivation in kittens: Further evidence for a sensitive period. *Journal of Physiology (London), 265,* 195–216.

Blasdel, G. G. (1992a). Differential imaging of ocular dominance and orientation selectivity in monkey striate cortex. *Journal of Neuroscience, 12,* 3115–3138.

Blasdel, G. G. (1992b). Orientation selectivity, preference, and continuity in monkey striate cortex. *Journal of Neuroscience, 12,* 3139–3161.

Blasdel, G. C., & Fitzpatrick, D. (1984). Physiological organization of layer 4 in macaque striate cortex. *Journal of Neuroscience, 4,* 880–895.

Carpenter, G. A., & Grossberg, S. (1981). Adaptation and transmitter gating in vertebrate photoreceptors. *Journal of Theoretical Neurobiology, 1,* 1–42.

Carpenter, G. A., & Grossberg, S. (Eds.) (1991). *Pattern recognition and self-organizing neural networks.* MIT Press, Cambridge, MA.

Carpenter, G. A., & Grossberg, S. (1993). Normal and amnesic learning, recognition, and memory by a neural model of cortico-hippocampal interactions. *Trends in Neurosciences, 16,* 131–137.

Chubb, C., & Sperling, G. (1989). Two motion perception mechanisms revealed through distance-driven reversal of apparent motion. *Proceedings of the National Academy of Science, 86,* 2985–2989.

Cogan, A. I., Clarke, M., Chan, H., & Rossi, A. (1990). Two-pulse monocular and binocular interactions at the differential luminance threshold. *Vision Research, 30,* 1617–1630.

Cogan, A. I., Lomakin, A. J., & Rossi, A. F. (1993). Depth in anticorrelated stereograms: Effects of spatial density and interocular delay. *Vision Research, 33,* 1959–1975.

Daw, N. W. (1994). Mechanisms of plasticity in the visual cortex. *Investigative Ophthalmology & Visual Science, 35,* 4168–4179.

Daw, N. W., & Wyatt, H. J. (1976). Kittens reared in a unidirectional environment: Evidence for a critical period. *Journal of Physiology (London), 257,* 155–170.

Desimone, R. (1992). Neural circuits for visual attention in the primate brain. In G. A. Carpenter & S. Grossberg (Eds.), *Neural networks for vision and image processing* (pp. 343–364). MIT Press, Cambridge, MA.

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex, 1,* 1–47.

Ferster, D. (1988). Spatially opponent excitation and inhibition in simple cells of the cat visual cortex. *Journal of Neuroscience, 8,* 1172–1180.

Francis, G., & Grossberg, S. (1996). Cortical dynamics of boundary segmentation and reset: Persistence, afterimages, and residual traces. *Perception, 25,* 543–567.

Francis, G., Grossberg, S., & Mingolla, E. (1994). Cortical dynamics of feature binding and reset: Control of visual persistence. *Vision Research, 34,* 1089–1104.

Freeman, R. D., & Ohzawa, I. (1992). Development of binocular vision in the kitten's striate cortex. *Journal of Neuroscience, 12,* 4721–4736.

Frégnac, Y., Burke, J. P., Smith, D., & Friedlander, M. J. (1994). Temporal covariance of pre- and postsynaptic activity regulates functional connectivity in the visual cortex. *Journal of Neurophysiology, 71,* 1403–1421.

Gaudiano, P. (1994). Simulations of X and Y retinal ganglion cell behavior with a nonlinear push-pull model of spatiotemporal retinal processing. *Vision Research, 34,* 1767–1784.

Gilbert, C. D. (1977). Laminar differences in receptive field properties of cells in cat primary visual cortex. *Journal of Physiology (London), 268,* 391–421.

Gove, A., Grossberg, S., & Mingolla, E. (1995). Brightness perception, illusory contours, and corticogeniculate feedback. *Visual Neuroscience, 12,* 1027–1052.

Grossberg, S. (1968). Some nonlinear networks capable of learning a spatial pattern of arbitrary complexity. *Proceedings of the National Academy of Sciences, 59,* 368–372.

Grossberg, S. (1972). A neural theory of punishment and avoidance: II. Quantitative theory. *Mathematical Biosciences, 15,* 253–285.

Grossberg, S. (1973). Contour enhancement, short term memory, and constancies in reverberating neural networks. *Studies in Applied Mathematics, LII,* 213–257.
Grossberg, S. (1974). Classical and instrumental learning by neural networks. In R. Rosen & F. Snell (Eds.), *Progress in theoretical biology, vol. 3* (pp. 51–141). New York: Academic Press.
Grossberg, S. (1976a). Adaptive pattern classification and universal recoding, I: Parallel development and coding of neural feature detectors. *Biological Cybernetics, 23,* 121–134.
Grossberg, S. (1976b). Adaptive pattern classification and universal recoding, II: Feedback, expectation, olfaction, illusions. *Biological Cybernetics, 23,* 187–202.
Grossberg, S. (1980). How does a brain build a cognitive code? *Psychological Review, 87,* 1–51.
Grossberg, S. (1983). The quantized geometry of visual space: The coherent computation of depth, form, and lightness. *Behavioral and Brain Sciences, 6,* 625–657.
Grossberg, S. (1987). Cortical dynamics of three-dimensional form, color and brightness perception, II: Binocular theory. *Perception and Psychophysics, 41,* 117–158.
Grossberg, S. (1994). 3-D vision and figure-ground separation by visual cortex. *Perception & Psychophysics, 55,* 48–120.
Grossberg, S. (1995). The attentive brain. *American Scientist, 83,* 438–449.
Grossberg, S. (1997). Cortical dynamics of 3-D figure-ground perception of 2-D pictures. *Psychological Review, 104,* 618–658.
Grossberg, S., & Grunewald, A. (1995). Temporal dynamics of binocular disparity processing with corticogeniculate interactions. Tech. rep. CAS/CNS-TR-95-021, Boston University, Boston, MA.
Grossberg, S., & McLoughlin, N. (1997). Cortical dynamics of 3-D surface perception: Binocular and half-occluded scenic images. *Neural Networks, 10,* 1583–1605.
Grossberg, S., & Merrill, J. W. L. (1996). The hippocampus and cerebellum in adaptively timed learning, recognition, and movement. *Journal of Cognitive Neuroscience, 8,* 257–277.
Grossberg, S., & Mingolla, E. (1985a). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review, 92,* 173–211.
Grossberg, S., & Mingolla, E. (1985b). Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations. *Perception & Psychophysics, 38,* 141–171.
Grossberg, S., & Pessoa, L. (1998). Texture segregation, surface representation, and figure-ground separation. *Vision Research,* in press.
Grunewald, A. (1995). *Temporal dynamics of visual perception.* Ph.D. thesis, Boston University.
Grunewald, A., & Lankheet, M. J. M. (1996). Orthogonal motion-aftereffect predicted by a model of cortical motion processing. *Nature, 384,* 358–360.
Hawken, M. J., & Parker, A. J. (1984). Contrast sensitivity and orientation selectivity in lamina IV of the striate cortex of old world monkeys. *Experimental Brain Research, 54,* 367–372.
Held, R., Birch, E. E., & Gwiazda, J. (1980). Stereoacuity of human infants. *Proceedings of the National Academy of Sciences USA, 77,* 5572–5574.
Hirsch, J., & Gilbert, C. (1991). Synaptic physiology of horizontal connections in the cat's visual cortex. *Journal of Neuroscience, 11,* 1801–1809.
Hodgkin, A. L. (1964). *The conduction of the nervous impulse.* Liverpool, UK: Liverpool University.
Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology (London), 160,* 106–154.
Julesz, B. (1960). Binocular depth perception of computer-generated patterns. *Bell System Technical Journal, 38* 1001–1020.
Julesz, B., & White, B. (1969). Short term visual memory and the Pulfrich phenomenon. *Nature, 222,* 639–641.
Kirkwood, A., Bear, M. F. (1994a). Hebbian synapses in visual cortex. *Journal of Neuroscience, 14,* 1634–1645.
Kirkwood, A., & Bear, M. F. (1994b). Homosynaptic long-term depression in the visual cortex. *Journal of Neuroscience, 14,* 3404–3412.
Kohonen, T. (1984). *Self-organization and associative memory.* New York: Springer-Verlag.
Leventhal, A. G., & Hirsch, H. V. B. (1980). Receptive field properties of different classes of neurons in visual cortex or normal and dark-reared cats. *Journal of Neurophysiology, 43,* 1111–1132.
Leventhal, A. G., Thompson, K. G., Liu, D., Zhou, Y., & Ault, S. J. (1995). Concomitant sensitivity to orientation, direction, and color of cells in layers 2, 3, and 4 of monkey striate cortex. *Journal of Neuroscience, 15,* 1808–1818.
Liu, Z., Gaska, J. P., Jacobson, L. D., & Pollen, D. A. (1992). Interneuronal interaction between members of quadrature phase and anti-phase pairs in the cat's visual cortex. *Vision Research, 32,* 1193–1198.
Macchi, G., & Rinvik, E. (1976). Thalamo-telencephalic circuits: A neuroanatomical survey. In A. Remond (Ed.), *Handbook of electroencephalography and clinical neurophysiology, vol. 2.* Amsterdam: Elsevier.
McGuire, B., Gilbert, C., Rivlin, P., & Wiesel, T. (1991). Targets of horizontal connections in macaque primary visual cortex. *Journal of Comparative Neurology, 305,* 370–392.
McKee, S. P., Bravo, M. J., Smallman, H., & Legge, G. E. (1995). The "uniqueness constraint" and binocular masking. *Perception, 24,* 49–65.
McKee, S. P., Bravo, M. J., Taylor, D. G., & Legge, G. E. (1994). Stereo matching precedes dichoptic masking. *Vision Research, 34,* 1047–1060.
McLoughlin, N., & Grossberg, S. (1998). Cortical computation of stereo disparity. *Vision Research, 38,* 91–99.
Motter, B. C. (1994a). Neural correlates of attentive selection for color or luminance in extrastriate area V4. *Journal of Neuroscience, 14,* 2178–2189.
Motter, B. C. (1994b). Neural correlates of feature selective memory and pop-out in extrastriate area V4. *Journal of Neuroscience, 14,* 2190–2199.
Movshon, J. A., & Düsteler, M. R. (1977). Effects of brief periods of unilateral eye closure on the kitten's visual system. *Journal of Neurophysiology, 40,* 1255–1265.
Murphy, P. C., & Sillito, A. M. (1987). Corticofugal influences on the generation of length tuning in the visual pathway. *Nature, 329,* 727–729.
Ohzawa, I., DeAngelis, G. C., & Freeman, R. D. (1990). Stereoscopic depth discrimination by the visual cortex: Neurons ideally suited as disparity detectors. *Science, 249,* 1037–1041.
Olson, S., & Grossberg, S. (1996). A neural network model for the development of simple and complex cell receptive fields within cortical maps of orientation and ocular dominance. Tech. rep. CAS/CNS-TR-96-021, Boston University, Boston, MA. *Neural Networks,* in press.
Poggio, G. F., & Fischer, B. (1977). Binocular interaction and depth sensitivity in striate and prestriate cortex cortex of behaving rhesus monkeys. *Journal of Neurophysiology, 40,* 1392–1407.
Pulfrich, C. (1922). Die stereoskopie im dienste der iso-

chromen und heterochromen photometrie. *Die Naturwissenschaften, 10,* 553–564, 570–574, 596–601, 714–722, 735–743, 751–761.

Reynolds, J., Chelazzi, L., Luck, S., & Desimone, R. (1994). Sensory interactions of selective spatial attention in macaque area V2. *Society for Neuroscience Abstracts, 20,* 1054.

Reynolds, J., Nicholas, J., Chelazzi, L., & Desimone, R. (1995). Spatial attention protects macaque V2 and V4 cells from the influence of non-attended stimuli. *Society for Neuroscience Abstracts, 21,* 1759.

Ringach, D., Carandini, M., Sapiro, G., & Shapley, R. (1996). Cortical circuitry revealed by reverse correlation in the orientation domain. *Perception, 25,* Supplement 31.

Schiller, P. (1992). The ON and OFF channels of the visual system. *Trends in Neurosciences, 15,* 86–92.

Shimojo, S., Bauer, J., O'Connell, K. M., & Held, R. (1986). Prestereoptic binocular vision in infants. *Vision Research, 26,* 501–510.

Sillito, A. M. (1977). Inhibitory processes underlying the directional specificity of simple, complex, and hypercomplex cells in the cat's visual cortex. *Journal of Physiology, 271,* 699–720.

Sillito, A. M. (1979). Inhibitory mechanisms influencing complex cell orientation selectivity and their modification at high resting discharge levels. *Journal of Physiology, 33–53,* 289.

Sillito, A. M., Jones, H. E., Gerstein, G. L., & West, D. C. (1994). Feature-linked synchronization of thalamic relay cell firing induced by feedback from the visual cortex. *Nature, 369,* 479–482.

Singer, W. (1983). Neuronal activity as a shaping factor in the self-organization of neuron assemblies. In E. Basar, H. Flohr, H. Haken, & A. Mandell (Eds.), *Synergetics of the brain* (pp. 89–101). New York: Springer-Verlag.

Skottun, B. C., De Valois, R. L., Grosof, D. H., Movshon, J. A., Albright, D. G., & Bonds, A. B. (1991). Classifying simple and complex cells on the basis of response modulation. *Vision Research, 31,* 1079–1086.

Smallman, H., & McKee, S. P. (1995). A contrast ratio constraint on stereo matching. *Proceedings of the Royal Society of London, B,* 265–271.

Sutter, A., Beck, J., & Graham, N. (1989). Contrast and spatial variable in texture segregation: Testing a simple spatial-frequency channels model. *Perception and Psychophysics, 46,* 312–332.

Tsumoto, T., Creutzfeldt, O. D., & Legendy, C. R. (1978). Functional organization of the cortiofugal system from visual cortex to lateral geniculate nucleus in the cat. *Experimental Brain Research, 32,* 345–364.

van Essen, D. C., & Maunsell, J. H. R. (1983). Hierarchical organization and functional streams in the visual cortex. *Trends in Neurosciences, 6,* 370–375.

Varela, F. J., & Singer, W. (1987). Neuronal dynamics in the visual corticothalamic pathway revealed through binocular rivalry. *Experimental Brain Research, 66,* 10–20.

von der Malsburg, C. (1973). Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik, 14,* 85–100.

von Helmholtz, H. (1910/1925). *Treatise on physiological optics* (J. P. Southall, trans.). New York: Dover.

Weber, A. J., Kalil, R. E., & Behan, M. (1989). Synaptic connections between corticogeniculate axons and interneurons in the dorsal lateral geniculate nucleus of the cat. *Journal of Comparative Neurology, 289,* 156–164.

Westendorf, D. H., Blake, R. R., & Fox, R. (1972). Binocular summation of equal-energy flashes of unequal duration. *Perception & Psychophysics, 12,* 445–448.