1995-07

Temporal Dynamics of Binocular Display Processing with Corticogeniculate Interactions

Grossberg, Stephen

Boston University Center for Adaptive Systems and Department of Cognitive and Neural Systems

http://hdl.handle.net/2144/2198

Boston University
TEMPORAL DYNAMICS OF BINOCULAR DISPARITY PROCESSING WITH CORTICOGENICULATE INTERACTIONS

Stephen Grossberg and Alexander Grunewald

July 1995

Technical Report CAS/CNS-95-021

Permission to copy without fee all or part of this material is granted provided that: 1. the copies are not made or distributed for direct commercial advantage, 2. the report title, author, document number, and release date appear, and notice is given that copying is by permission of the BOSTON UNIVERSITY CENTER FOR ADAPTIVE SYSTEMS AND DEPARTMENT OF COGNITIVE AND NEURAL SYSTEMS. To copy otherwise, or to republish, requires a fee and/or special permission.

Copyright © 1995

Boston University Center for Adaptive Systems and Department of Cognitive and Neural Systems
111 Cummmington Street
Boston, MA 02215
Temporal dynamics of binocular disparity processing with corticogeniculate interactions

Stephen Grossberg and Alexander Grunewald

Department of Cognitive and Neural Systems
and
Center for Adaptive Systems
Boston University
111 Cummington Street
Boston, Massachusetts 02215

July, 1995

Technical Report CAS/CNS-TR-95-021
Boston, MA: Boston University

Send requests for reprints to:
Professor Stephen Grossberg
Boston University
Center for Adaptive Systems
111 Cummington Street
Boston, MA 02215
617-353-7857

Short Title: Cortical Binocular Disparity Processing

Manuscript: 31 pages
References: 7 pages
Figure Captions: 5 pages
Tables: 1 table
Figures: 21 figures
Abstract

A neural model of binocular vision is developed to simulate psychophysical and neurobiological data concerning the dynamics of binocular disparity processing. The model shows how feedforward and feedback interactions among LGN ON and OFF cells and cortical simple, complex, and hypercomplex cells can simulate binocular summation, the Pulfrich effect, and the fusion of delayed anticorrelated stereograms. Model retinal ON and OFF cells are linked by an opponent process capable of generating antagonistic rebounds from OFF cells after offset of an ON cell input. Spatially displaced ON and OFF cells excite simple cells. Opposite polarity simple cells compete before their half-wave rectified outputs excite complex cells. Complex cells binocularly match like-polarity simple cell outputs before pooling half-wave rectified signals from opposite polarities. Competitive feedback among complex cells leads to sharpening of disparity selectivity and normalizes cell activity. Slow inhibitory interneurons help to reset complex cells after input offset. The Pulfrich effect occurs because the delayed input from the one eye fuses with the present input from the other eye to create a disparity. Binocular summation occurs for stimuli of brief duration or of low contrast because competitive normalization takes time, and cannot occur for very brief or weak stimuli. At brief SOAs, anticorrelated stereograms can be fused because the rebound mechanism ensures that the present image to one eye can fuse with the afterimage from a previous image to the other eye. Corticogeniculate feedback embodies a matching process that enhances the speed and temporal accuracy of complex cell disparity tuning. Model mechanisms interact to control the stable development of sharp disparity tuning.

Key Words: binocular vision, binocular disparity, visual cortex, lateral geniculate nucleus, neural networks, Boundary Contour System, Pulfrich effect, binocular summation, anticorrelated stereograms
1 Introduction

This article develops a neural model of the temporal dynamics that occur during early stages of binocular vision. Some of the results were first briefly reported in Grunewald and Grossberg (1995b). To obtain a binocular representation of the environment, visual information available from the two eyes has to be combined. During normal viewing, the two eyes converge, and so the same part of the visual scene falls onto the centers of the two foveae. Since the two retinae are horizontally displaced in the head, not all information from the one retina is located at the corresponding site in the other retina. For example, the retinal image of an object that is slightly behind the fixation point will be shifted leftwards in the left eye and rightwards in the right eye (Figure 1). This disparity is used by the visual cortex as a powerful cue for depth (Julesz, 1971).

Figure 1

Since objects can be at different depths, a particular location on the retina is not always paired in the cortex with the same location on the other retina. One retinal location could be paired in the cortex with several other possible locations on the other retina. The two locations that are paired typically generate a binocular fused percept of a single location in space. One of the difficult tasks that the visual system faces is to decide which retinal location goes with which for any given visual input. This task is often called the correspondence problem (Julesz, 1971).

During free viewing, human observers tend to make about three eye movements per second. This means that the correspondence problem has to be solved very rapidly based on the two images that are being processed by the two eyes (see Figure 2). The present model simulates several types of data about the transient dynamics of binocular vision that illustrate how the brain accomplishes this task.

Figure 2

The model combines two previous modeling studies of visual perception and extends them into the dynamical domain. Both of these studies developed parts of the Boundary Contour System (BCS) of emergent boundary segmentation that was introduced in Cohen and Grossberg (1984), Grossberg (1984), Grossberg and Mingolla (1985a, 1985b) to model aspects of the interblob cortical processing stream from the lateral geniculate nucleus (LGN) to extrastriate area V4. In previous work, Grossberg (1994) further developed the theory, called FACADE theory, of which the BCS forms a part to explain perceptual and neural data about 3-D vision and figure-ground pop-out. Grossberg and McLoughlin (1995) and McLoughlin and Grossberg (1994, 1995) refined FACADE theory to simulate data about da Vinci stereopsis (Gillam & Borsting, 1988; Kaye, 1978; Lawson & Gulick, 1967; Nakayama & Shimojo, 1990; Wheatstone, 1838) and about dichoptic masking, contrast-sensitive binocular matching, and Panum’s limiting case (McKee, Bravo, Smallman, & Legge, 1994a; McKee, Bravo, Taylor, & Legge, 1994b). These simulations used model interactions from LGN ON and OFF cells to cortical simple cells and complex cells. In a parallel development, Gove, Grossberg, and Mingolla (1995) studied a monocular version of the model that included interactions from LGN cells through cortical simple, complex, hypercomplex, and bipolar cells, including feedback interactions from endstopped cortical complex cells back to the LGN. This work simulated data about brightness perception and the formation of illusory

Such corticogeniculate feedback plays a key role in studies of cortical disparity tuning for several reasons. In Grossberg (1976b), it was proposed that corticogeniculate feedback carries out a matching function whereby LGN cell activities that are consistent with cortical activations are preserved and synchronized, whereas inconsistent LGN activities are suppressed. In particular, monocular LGN activations were proposed to resonate synchronously with consistent cortical activations while binocularly inconsistent LGN activations are suppressed. Sillito, Jones, Gerstein, and West (1991) and Varela and Singer (1987) have reported neurophysiological evidence that is consistent with this prediction. The present work models how this feedback influences the dynamics of binocular disparity processing. Grossberg (1976b) also proposed that this corticogeniculate matching process plays a role in regulating and stabilizing the learning process whereby cortical complex cells achieve fine binocular disparity tuning during development. A companion paper (Grunewald & Grossberg, 1995a) models this learning process and simulates how corticogeniculate feedback may influence the development of sharp binocular tuning.

2 Review of experimental evidence

2.1 Psychophysics

Some important dynamic properties of 3-D vision are binocular summation, the Pulfrich effect, and fusion of anticorrelated stereograms. They are simulated here to illustrate various facets of the model’s dynamical interactions.

2.1.1 Binocular summation

When one views the world with only one eye, the world does not appear darker, even though the visual system is in fact receiving less visual input. This is known as Fechner’s paradox (Cogan, 1982; Hering, 1964; Levelt, 1965). However, there are situations in which Fechner’s paradox does not hold. In particular, when viewing very brief or very dim stimuli, the detection threshold is lower when the stimulus is seen binocularly rather than monocularly (Andersen & Movshon, 1989; Cogan, Clarke, Chan, & Rossi, 1990; Westendorf, Blake, & Fox, 1972). In other words, binocular summation affects the perception of surface properties. This result can also be extended to orientation discrimination tasks, in which subjects’ performance improves when both eyes are stimulated (Bears & Freeman, 1994; Legge, 1984a, 1984b). Thus, binocular summation can also influence properties of boundary segmentation. Taken together, these findings suggest that for brief and low contrast stimuli, a facilitation occurs when stimuli are viewed binocularly as opposed to monocularly, and that such binocular summation seems to involve both surface properties, such as brightness perception, and boundary properties, such as orientation discrimination. The present study focuses upon the latter type of process. We show how binocular summation can occur during brief stimulation of the BCS, even though it is normalized away by slower-acting competitive interactions in response to more prolonged stimuli. This stimulation probes the energetic aspects of binocular fusion through time.
2.1.2 Pulfrich effect

The Pulfrich effect can be elicited by positioning a dark glass in front of one eye (Pulfrich, 1922). If the observer then watches a pendulum that is moving within a frontoparallel plane, it appears as if the pendulum is moving along an ellipse in the horizontal plane (Julesz & White, 1969). In particular, when the pendulum moves from the side with the darkened eye towards the other side, then the pendulum appears further away from the observer, and it appears nearer when it is moving in the opposite direction. This is illustrated in Figure 3. This study probes how the model assigns disparity values to rapidly changing binocular combinations.

Figure 3

2.1.3 Fusion of anticorrelated stereograms

A stereogram is made up of two images, one for each eye. When these images are binocularly fused, a percept in depth becomes visible. In correlated stereograms, two images are constructed with corresponding features, which may be slightly shifted to create a disparity when each picture is viewed through a different eye. Random dot stereograms are stereograms that exclusively contain black and white dots arranged in a random fashion (Julesz, 1960). This is illustrated in Figure 4. To perceive depth from such a stereogram, the visual system typically matches dots with the same contrast polarities relative to their background.

Figure 4

In an anticorrelated random dot stereogram, the contrast polarities between the two images are reversed. This is illustrated in Figure 5. It is impossible to fuse them under normal conditions (Julesz, 1960). However, if two anticorrelated pictures are presented with a slight temporal asynchrony (about 60ms), then fusion is possible (Cogan, Lomakin, & Rossi, 1993). Afterimages occur following visual simulation and have reversed contrasts. This suggests that the fusion of one image occurs with the afterimage of an earlier image in the other eye. At similar asynchronies, fusion of correlated stereograms becomes impossible. This study probes how the model binocularly matches image contrasts with the same contrast polarities, but not opposite contrast polarities. The study also highlights how opponent processes are incorporated in the model and how rebounds within these opponent circuits can reverse internal polarity values at input offset.

Figure 5

2.2 Review of relevant neurophysiology and anatomy

Visual input from the retina projects via retinal ganglion cells to the lateral geniculate nucleus (LGN) and then to area V1 in the striate visual cortex. Area V1 is arranged in several layers (layers 1-6), and neurons in layer 6 project from area V1 back to the LGN. The following paragraphs review relevant data.
2.2.1 Retina

Ganglion cells in the retina have small receptive fields that typically consist of a central region and an annular surround (Schiller, 1992). ON cells have on-center off-surrround anatomies that are excited by an increment light flash in their center, and are inhibited by an increment light flash in their surround. A decrement light flash inhibits the center, but excites the surround. When stimulated with a uniform field they do not respond at all. OFF cells have off-center on-surrround anatomies that respond in the opposite way: a decrement light flash excites them in the center, and inhibits them in the surround. After the offset of a stimulus, a cell that responded to the stimulus will quickly cease to respond, while a cell of the opposite polarity, which was not activated by the stimulus, will respond briefly (Enroth-Cugell & Robson, 1966). For example, an increment light flash will excite an ON cell, and inhibit an OFF cell at the same location. At the end of the flash, the ON cell will be inhibited, but the opponent OFF cell will be transiently activated. This transient response will henceforth be called an antagonistic rebound.

2.2.2 Lateral geniculate nucleus

Like retinal ganglion cells, neurons in the LGN also have a center-surround structure (Wiesel & Hubel, 1966). The LGN comprises relay cells, which are excitatory and project to cortex, and interneurons that can inhibit relay cells (Sillito & Kemp, 1983). Both interneurons and relay cells receive direct input from the retina (Dubin & Cleland, 1977). When the corticothalamic input is abolished (e.g., by aspiration or chemical inactivation of cortex), then cells in the LGN show no orientation or length tuning (Sillito & Murphy, 1993). The effects of cortico-thalamic feedback will be discussed after a survey of cell properties in striate cortex.

2.2.3 Primary visual cortex

Area V1 in the visual cortex is arranged in six layers. Input from the LGN arrives at layer 4, where neurons have receptive fields with center/surround organization, very much like those found in the LGN. At least three different cell types have been identified in area V1: simple cells, complex cells, and hypercomplex cells. All of these cell types are tuned for orientation, and they show spatial summation. Simple cells have clearly identifiable ON and OFF regions. If a light increment falls within the ON region, it will excite the simple cell. Likewise, a light decrement within the OFF region excites the simple cell (Hubel & Wiesel, 1962). Unlike LGN cells, however, these two regions are in general parallel, and do not surround each other (Hubel & Wiesel, 1962). An edge parallel to the border between the ON and the OFF regions of the correct polarity is the optimal stimulus for a simple cell. When a bar of optimal orientation is swept through a simple cell’s receptive field, it will respond once. Complex cells do not have identifiable ON and OFF regions, and they respond to edges of either polarity. As a consequence, they respond twice when a bar is swept through their receptive field. This distinction between simple and complex cells has been used to define simple and complex cells (Skottun, De Valois, Grosor, Moshion, Albright, & Bonds, 1991). Hypercomplex cells respond to a large degree like complex cells. However, the responses of hypercomplex cells are reduced when the length of the stimulus gets sufficiently large (Hubel & Wiesel, 1965). This property is called endstopping.
Many simple cells are found in the upper parts of layer 4, and many of them are only responsive to stimulation from one eye. Complex cells are predominantly found in layers 2 and 3, and most of them are binocular. Hypercomplex cells are mostly found in layer 6. Hubel and Wiesel (1962) proposed that the cells in striate cortex are organized in a hierarchical way. According to that view, simple cells receive geniculate input, complex cells receive input from simple cells, and hypercomplex cells receive input from complex cells. Although this view has been challenged since then, it seems reasonable to assume that some simple, complex and hypercomplex cells are arranged in this way. In support of a hierarchical structure of these cell types in area V1 is the finding that fewer simple cells are binocular than complex cells, and fewer complex cells are binocular than hypercomplex cells (Gilbert, 1977; Hubel & Wiesel, 1962, 1968).

All these cell types discussed thus far are considered to be excitatory. There are also inhibitory interneurons in the primary visual cortex. The time constant of inhibitory interneurons appears to be slower than that of excitatory cells. There is evidence that suggests that pairs of excitatory and inhibitory neurons work together: the excitatory neuron excites the inhibitory neuron, and (slightly delayed) the inhibitory neurons inhibits the excitatory neuron (Krüger & Aiple, 1988). At present it is not clear how localized this inhibitory effect is.

Excitatory neurons in area V1 differ in the extent to which they respond to stimulation by either eye or by both eyes. This is called ocular dominance. Some cells fire only if one of the eyes is stimulated, while the input in the other eye is irrelevant. Other cells fire when either eye is stimulated, and they fire stronger when both eyes are stimulated (Gilbert, 1977; Hubel & Wiesel, 1962; Ohzawa & Freeman, 1986b, 1986a).

### 2.2.4 Corticogeniculate feedback

The LGN does not only receive projections from the retina, but also from cortex, from layer 6 of the striate cortex in particular (Robson, 1983). In fact, the majority of the input to LGN cells comes from there. The strength of this feedback connection suggests that it may play an important role in visual processing. It has been reported that cells in the LGN are endstopped, and that they can show orientation and length tuning (Cleland, Lee, & Vidyasagar, 1983), which is likely to be mediated through the cortico-thalamic projections (Sillito & Murphy, 1993). Varela and Singer (1987) also showed that cortical feedback can have a pronounced effect on the excitability of cells in the LGN. Those authors found geniculate activity reduced when the retinal input to the LGN does not match the cortical feedback, as is the case in binocular rivalry. These results were recently confirmed and extended to show that responses in the LGN synchronize at cells where retinal and cortical signals converge (Sillito et al., 1994). Thus corticogeniculate feedback can have a profound effect on neural activities in the LGN.

Feedback from cortex is excitatory (Montero, 1990), but it goes to both interneurons (Weber, Kalil, & Behan, 1989) and relay cells (Dubin & Cleland, 1977). Due to this complex pattern of connections, conflicting results have been reported: there are reports of excitatory influences (Kalil & Chase, 1970), inhibitory influences (Hull, 1968), and mixed effects (Marrocco & McClurkin, 1985). It seems clear from these results that feedback plays a role in spatially localized processing. Evidence to support this comes from the precise topography
of the feedback projections (Updyke, 1975). This means that a simple role as the source of arousal cannot be conjectured for the feedback projections.

3 A neural model of binocular processing

The model incorporates several processing stages that correspond to LGN and cortical cell types. A global overview of the model architecture is shown in Figures 6 and 7 and is briefly surveyed before a more detailed stage-by-stage description is given. At the lowest stage, retinal information is separated into ON and OFF channel responses. This separation has shown to be useful for modeling the processing of contrast information under conditions of variable illumination (Grossberg, Mingolla, & Williamson, 1995; Grossberg & Todorović, 1988; Grossberg & Wyse, 1991; Pessoa, Mingolla, & Neumann, 1994). In particular, cells that obey membrane or shunting equations, and that interact as a part of on-center off-surround or off-center on-surround networks, are capable of discounting the illuminant, extracting Weber-law modulated ratio contrasts, and normalizing their total activities (Grossberg, 1973, 1983). ON and OFF cells are linked together by an opponent processing network, called a gated dipole (Grossberg, 1980) wherein offset of an input to an ON cell can trigger a transient antagonistic rebound in the corresponding OFF cell. The net outputs of ON and OFF cells are passed on to the LGN stage, where they are combined with feedback from hypercomplex cells.

Figure 6

Spatially offset ON and OFF outputs from the LGN are activated by image contrast changes, and activate, in turn, oriented simple cells. Excitatory ON and OFF cell output signals add up at their target simple cells. Simple cells have a particular contrast polarity (light-dark vs. dark-light). Simple cells with like position and orientation but opposite contrast polarity compete before their rectified activities are output to the complex cell stage. Here, information about the polarity of an edge is lost by pooling signals from like-oriented simple cells with opposite contrast polarities. This pooling process also enables binocular disparity information to be extracted. The complex cell stage, in turn, sends its activity to the hypercomplex cell stage, whose outputs generate feedback signals to the LGN.

Two components of the model merit special attention: the organization of the complex cell field and of the feedback from hypercomplex cells to the LGN. The complex cell field rapidly matches the information from the two eyes. Complex cells pool activities from simple cells from both eyes and of both polarities. The main issue to be understood is how complex cells can binocularly match like contrasts from the two eyes, yet have output signals that pool opposite contrast polarities. Grossberg and McLoughlin (1995) and McLoughlin and Grossberg (1994, 1995) proposed that activities from simple cells of the same polarity facilitate each other while opposite polarities inhibit each other, before all polarity combinations of this interaction are half-wave rectified and added to generate the final complex cell response. This ensures that complex cells pool both polarities of contrast, yet only match across like polarities. As a result of summing these half-wave rectified signals, the complex cell computes a full-wave rectified, oriented filtering of the image.

Matches also occur only within a predefined distance. In other words, there is a limit to the disparities that can be fused. This limit is called Panum’s limiting area in the psy-
chophysical literature (Fender & Julesz, 1967). Many matches can be initiated within this distance by an image at each position, but typically only one succeeds in substantially activating the complex cells there. This is ensured through recurrent lateral inhibition across the complex cell field, which contrast enhances the input pattern received by the complex cells (Sillito, 1979; Sillito, Salt, & Kemp, 1985).

At the offset of an input, the complex cell field needs to be able to reset itself, in the sense that no node remains persistently active. The model circuit that connects simple cells and complex cells contains interneurons that control this reset process.

Figure 7

The second key element of the model is feedback from the hypercomplex cell stage to the LGN stage. This feedback stabilizes the processing at the complex cell stage. Feedback occurs when the activities of the complex cells converge onto a given disparity. This means that the winning complex cells have achieved a high level of activity, and all other cells have zero activity. Then the hypercomplex cell stage sends signals to the LGN stage which further activate those LGN cells that feed the active complex cells, while inhibiting LGN cells that do not. One can think of the feedback activity as a confirmation, or verification, signal. When activity at the LGN stage matches the confirmation signal, the corresponding LGN cell activities are enhanced, and therefore a stronger signal is fed to the simple cell, complex cell, and hypercomplex cell stages, whereupon the feedback to the LGN also increases. This feedback cycle is shown below to converge rapidly to a resonant equilibrium state between mutually consistent LGN and cortical cell activities.

3.1 Model processing stages

A more precise mechanistic description of model mechanisms is now given. All processing stages prior to the complex cell stage are monocular, thus requiring a double complement of neural fields, one for each eye (Figure 7). To achieve maximal conceptual clarity, each processing stage models only those neural properties that are rate-limiting in explaining the data.

Figure 8

3.1.1 Retinal stage

At the retinal stage, ON and OFF responses are obtained by convolving the retinal image with center and surround kernels. These cell activities, or potentials, are then half-wave rectified to yield ON and OFF signals. This is shown schematically in Figure 8. Adaptation to prolonged exposure to stimulus is achieved by incorporating the ON and OFF signals into a gated dipole opponent processing circuit that coordinates ON and OFF responses (Figure 9). A gated dipole responds to either ON or OFF signals with an initial transient overshoot that decays, or habituates, to a sustained lower value when the input stimulus persists. Habituation is mediated by chemical transmitters that multiply, or gate, the signals from the ON and OFF cells (see the square synapses in Figure 9), before the gated signals compete and the net signals are rectified. Due to the persistence of asymmetric transmitter habituation in the input channel after the input terminates, the opponent channel gets
transiently activated, after which the gated dipole gradually equilibrates to its resting status. See Öğmen and Gagné (1990a, 1990b) for gated dipole modeling of cell responses in the fly retina and Francis, Grossberg, and Mingolla (1994) for modeling of mammalian visual cortex.

3.1.2 LGN stage

At the LGN stage, the retinal ON and OFF activities are fed bottom-up into LGN ON and OFF cells. These cells also receive excitatory and inhibitory top-down signals from hypercomplex cells, as in Figure 10. As noted above, corticogeniculate feedback makes a prediction about the neural patterns that the hypercomplex cell “expects” to find at the LGN level. If the bottom-up and top-down signals match, then the LGN activity that is passed on to the next stage of processing is enhanced. If the signals do not match, then the LGN signal is attenuated. These properties are achieved by combining topographically organized excitatory feedback with nonspecific inhibitory feedback (Figure 10) to capture the main effects of the LGN feedback circuit, as in Gove et al. (1995). Target cells are activated when both bottom-up and top-down excitatory feedback converge. If only top-down inhibitory feedback converges on a previously active cell, then that cell’s activity is attenuated. This scheme is similar to the interaction between bottom-up and top-down signals that is described in Adaptive Resonance Theory, or ART (Carpenter & Grossberg, 1991, 1993; Grossberg, 1976b, 1995).

3.1.3 Simple cell stage

At the simple cell stage, ON and OFF signals from slightly shifted positions lead to maximal excitation. By itself, convergence of excitatory ON and OFF signals could activate simple cells even in the absence of a contrast difference. This is avoided by introducing competition between simple cells of opposite polarity, as in Figure 11. Such an interaction has been reported in experiments of Ferster (1988), Liu, Gaska, Jacobson, and Pollen (1992), and Pei, Vidyasagar, Volgushev, and Creutzfeldt (1994). It has been used to process visual images by Cruthirds, Gove, Grossberg, Mingolla, Nowak, and Williamson (1992), Gove et al. (1995), and Grossberg et al. (1995).

3.1.4 Complex cell stage

At the complex cell stage there are two types of neurons: complex cells and inhibitory interneurons. The complex cells receive feedforward excitatory signals from simple cells of like orientation and both contrast polarities. Moreover, at each location there are complex cells that are sensitive to different disparities. Such a cell will be maximally activated if simple cells of the same polarity are activated, and if the peak of activity at the simple cells is positionally shifted between the two eyes by the disparity to which the complex cell is best
tuned. Simple cell activities from opposite polarities do not lead to complex cell activation. This circuit is illustrated in Figure 12.

**Figure 12**

While the feedforward signals from simple cells maximally activate some cells within the complex cell field, this interaction leads to broad disparity tuning, because the disparate inputs to complex cells come from a whole neighborhood of perceptual space. To obtain sharply tuned responses, the activities within the complex cell field interact via inhibitory feedback signals (Sillito, 1979; Sillito et al., 1985). The model incorporates local competition across space and across disparities (Grossberg, 1994). For cells distributed across a cortical hypercolumn map *in vivo* (Hubel & Wiesel, 1977), a single set of lateral inhibitory interactions across the map can influence both different positions and different disparities. Each cell also sends excitatory feedback to itself. Such a recurrent competitive field is summarized in Figure 13.

**Figure 13**

In a properly designed network, most neurons within a region will receive more inhibition than excitation. As a consequence their activities decay towards zero. Only the cell population with the strongest input will receive more excitation than inhibition. The recurrent interactions amplify these small differences into large differences, so the favored population emerges as the only active one. Such a “winner-take-all” circuit may be realized using shunting interactions in a recurrent on-center off-surround anatomy if a suitably defined nonlinear feedback signal function is also incorporated (Grossberg, 1973).

Recurrent competitive fields have the advantages that they can choose clear winners and maintain their activation after the input has vanished for learning to react to the winning activation pattern (Carpenter & Grossberg, 1991, 1993; Grossberg, 1976a; Grunewald & Grossberg, 1995b; Kohonen, 1984; von der Malsburg, 1973). After the network has converged onto some value, it becomes insensitive to future inputs. Such short-term activity cannot be allowed to persist indefinitely, or else the network would become insensitive to future inputs. Thus the persisting activation is reset shortly after the input terminates.

This reset circuit works as follows. Slow inhibitory interneurons are paired with each complex cell (Figure 14). These interneurons are inhibited by simple cell input and excited by complex cell feedback. They in turn inhibit their partner complex cell. When simple cells are active, they excite complex cells and inhibit the corresponding interneurons. Once a complex cell winner has emerged through feedback interactions, it excites its interneuron, but the simple cell inhibition keeps the interneuron inactive. When the input shuts off, however, the simple cells cease to respond, the interneurons are no longer inhibited by them, and thus they are only excited by complex cells. As a consequence, the interneurons become active and inhibit the corresponding complex cells until both are no longer active.

**Figure 14**

### 3.1.5 Hypercomplex cell stage

There are as many hypercomplex cells as there are complex cells in the model. In a two-dimensional implementation, competition across space among like-oriented complex cell out-
put leads to endstopped hypercomplex cell responses (Grossberg & Mingolla, 1985b). Since the present model provides only a one-dimensional implementation, this kind of competition has no additional effect. Therefore, in the present model, hypercomplex cells and complex cells are lumped together.

3.2 Simulation results

This section outlines simulations that show how the model can explain psychophysical results about dynamic properties of binocular processing.

3.2.1 Pulfrich effect

Figure 15 shows the input that was used in the simulations of the Pulfrich effect that was summarized in Section 2.1.2. The two eyes receive input of different strength, but of zero disparity. The left input is a bar of light which is defined by:

\[
I_i^l = \begin{cases} 
0.306 & \text{if } 20 \leq i < 60 \\
0.3 & \text{otherwise.}
\end{cases}
\] (1)

The bar was moved to the right by 6 positions after 3 time units 3 times, then it was stationary, before moving leftwards at the same rate. The right input is \(I_i^r = I_i^l/3\). Since the simplified model retina does not produce large contrast-sensitive delays, a delay of 3.0 time units was introduced at the retinal stage. More realistic, but complex, retinal models do produce large contrast-sensitive delays using a cascade of membrane equations followed by an automatically gain controlled habituative gate coupled to a membrane equation (Carpenter & Grossberg, 1981; Gaudiano, 1991). Since the focus of the present simulation was to analyse how the disparity-sensitive cortical model responds to such delays, the retinal model was kept as simple as possible.

By virtue of the retinal delay, at any given time the left simple cell activities are processed at the complex cell stage together with earlier right simple cell activities. Since the image is in motion, the left and the right simple cells that are processed at the complex cell stage are not at the same zero disparity locations. In the case where the bar is moving to the right, the active left simple cell locations will be shifted rightwards with respect to the active right simple cell locations, and therefore complex cells of disparity +3 are activated. When the bar is moving to the left, the left image will be shifted leftwards with respect to the right simple cell activities, and therefore complex cells of disparity -3 are activated. When the bar is static (in the beginning, or during a reversal of motion direction), the left and right simple cell activities are at the same spatial locations, and therefore complex cells of disparity 0 are activated. This is shown in Figure 16.
3.2.2 Fusion of anticorrelated stereograms

When the two images of an anticorrelated stereogram are presented with a slight temporal asynchrony, fusion of the image is possible (Cogan et al., 1993). Within the context of the model, this can be achieved through the antagonistic rebound response obtained at the retinal stage after input offset. In this complementary response, all polarities are inverted, but spatial positions are maintained. Because complex cells can only fuse correlated images, the rebound response of the first image can fuse with the response to the second image. In other words, the rebound response to a dark-light response is a light-dark response, which can fuse with a light-dark response due to a later stimulus. A simulation of this is shown in Figure 17. Simulations of Grunewald and Grossberg (1995a) suggest that this rebound response plays a key role in controlling the development of fine disparity tuning at complex cells.

Figure 17

Figure 17 shows the anticorrelated input that was used in the simulations. The left input is defined as follows:

\[
I_i^l = \begin{cases} 
1.1 & \text{if } 17 \leq i < 37 \text{ and } 0 < t < 2 \\
1 & \text{otherwise}
\end{cases} 
\]

and the right input is given by:

\[
I_i^r = \begin{cases} 
1 & \text{if } 23 \leq i < 43 \text{ and } 3 < t < 5 \\
1.1 & \text{otherwise}
\end{cases} 
\]

The responses of the complex cell field are shown in Figure 18. In the simulation, the first stimulus leads to a response in the complex cell field of disparity 0. Shortly after the left input goes off, the right anticorrelated input comes on, and complex cells of disparity -3 respond, indicating that the new stimulus has fused with the rebound response of the first stimulus.

Figure 18

3.2.3 Binocular summation

The boundary segmentations of the Boundary Contour lose polarity information after the complex cell stage. A complementary surface representation network, called the Feature Contour System (FCS), fills-in surface "features" such as color or brightness, within the boundaries generated by the BCS (Arrington, 1994; Cohen & Grossberg, 1984; Grossberg & Todorović, 1988; Paradiso & Nakayama, 1991). Visual percepts are hypothesized to arise in the FCS, while visual recognition can be derived from either BCS boundaries or FCS surfaces (Grossberg, 1994).

Which system is responsible for binocular summation? Traditional data on binocular summation, which emphasizes the lowering of the detection threshold when stimuli are viewed binocularly as opposed to monocularly (Andersen & Movshon, 1989; Cogan et al., 1990; Westendorf et al., 1972) suggest that the FCS may be responsible for binocular summation.
Other studies, however, show that binocular summation also occurs in the discrimination between stimuli (Legge, 1984a, 1984b). In particular, orientation discrimination is enhanced for suprathreshold stimuli by binocular summation (Bearse & Freeman, 1994). This finding suggests that boundary signals can also be enhanced by binocular summation. While this does not rule out a role for the FCS (see Cohen and Grossberg (1984) for such an analysis), it suggests that binocular summation also occurs within the BCS. This is what is modeled in this section.

The BCS binocular filter from simple cells to complex cells (Grossberg, 1994; Grossberg & McLoughlin, 1995; McLoughlin & Grossberg, 1994, 1995) helps to explain how complex cell responses to binocular stimuli are typically normalized due to competitive interactions within the recurrent competitive field. The same model can also account for binocular summation because it takes time for the inhibitory interactions of the complex cell field to bring it to the normalized state. Before that state is reached, cells that are stimulated by inputs from two eyes, as opposed to just one eye, will receive more activity, and hence their convergence to the normalized state occurs faster. If the input goes off before that state is reached, then a complex cell stimulated by a binocular stimulus will have reached a higher level of activity than a complex cell stimulated by a monocular stimulus. This means that the processing of a binocular stimulus will reach threshold for briefer stimuli, or for stimuli of less contrast, than the response to a monocular stimulus. In other words, the model exhibits binocular summation.

Figure 18 compares the responses of the complex cell field at disparity 0 when the stimulus is presented monocularly and binocularly. The input is defined as follows:

\[ I_i^r = \begin{cases} 
1.1 & \text{if } 20 \leq i < 40 \text{ and } 0 < t < 0.05 \\
1 & \text{otherwise}
\end{cases} \quad (4) \]

In the simulation of monocular presentation \( I_i^r = 1 \) and in the binocular simulations \( I_i^r = I_i^l \), where \( I_i^l \) is as defined above.

### 4 Mathematical description of the model

This section describes the equations and parameters of the model in detail. A 1-D simulation of the model was used. Neural activities are governed by membrane, or shunting, equations with a hyperpolarization term (Hodgkin, 1964; Grossberg, 1973):

\[ \frac{dx}{dt} = -Dx + (U - x)E - (L + x)I. \quad (5) \]

The term \(-Dx\) in equation (5) is a passive decay term which ensures that without any external input neural activity exponentially decays to zero. Term \((U - x)E\) is the excitatory shunting term, where \(E\) denotes excitatory input to the neuron, and \(U\) is the upper bound of neural activity. Factor \(U - x\) ensures that neural activity cannot rise above \(U\), no matter how large the input \(E\). Term \(-(L + x)I\) is a shunting inhibition term, where \(I\) denotes the inhibitory activity to the neuron, and \(-L\) is the lower bound of neural activity. Factor \(L + x\) ensures that activity never drops below the lower bound. Half-wave rectified activities
$X = \max(x, 0)$ are passed on as output signals.

**Figure 19**

A simple example is used to illustrate the dynamics of each processing stage within the network. For the sequence of flash stimuli in Figure 19a, the responses at the various levels are shown below. The first stimulus is at a negative disparity, and therefore appears behind the fixation point. The second stimulus is at zero disparity. The stimuli last for 8.0 time units. Activities are shown in intervals of 1.0 time units. The rectified activities at one stage form the inputs to the next stage. At each stage, the relevant equations and parameters are given. Table 1 provides an overview of all parameters used in the model.

**Table 1**

### 4.1 Image

There are two images, left and right. Each image consists of 1’s or 2’s. The input is a bar (of 2’s) that is slowly moving rightward on a background of 1’s. The disparity between the two images varies. The activity of the retinal image is denoted by $I_i$. Figure 19b represents the inputs in Figure 19a. The mathematical definition of the left image for $0 \leq t < 8$ is:

$$I_i^l = \begin{cases} 1 & 0 \leq i < 17 \\ 3 & 17 \leq i < 37 \\ 1 & 37 \leq i \leq 100 \end{cases}$$

and for the right image is:

$$I_i^r = \begin{cases} 1 & 1 \leq i < 23 \\ 3 & 23 \leq i < 43 \\ 1 & 43 \leq i \leq 100 \end{cases}$$

For $8 \leq t < 16$, the mathematical definition of the left and the right image is:

$$I_i^l = I_i^r = \begin{cases} 1 & 1 \leq i < 40 \\ 3 & 40 \leq i < 60 \\ 1 & 60 \leq i \leq 100. \end{cases}$$

For $t > 16$, all input values are 1, in other words

$$I_i^l = I_i^r = 1.$$

### 4.2 Kernels

Most kernels used in this model are Gaussians except when otherwise indicated. A Gaussian kernel is defined as follows:

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

where $\sigma$ specifies the size of the kernel. All kernels used are normalized so that $\sum_y G(y) = 1$, and $k$ is chosen accordingly. For notational convenience, the subscript $o$ of a kernel indicates
the origin of the signal with which the kernel is to be convolved. That ensures that when, within a single equation, signals from multiple sources converge, it is clear which kernel goes with which incoming signal. The size of a kernel is the number of source nodes for which the kernel contains weights. The implemented size of the kernels throughout the simulations is 17 nodes (1 centered on the node receiving the input and 8 on either side). This does not mean each kernel is actually different from zero over all the 17 nodes used. The size from a functional point of view is determined by \( \sigma \).

### 4.3 Retinal stage

At each retinal stage, there are 4 fields of neurons: 2 eyes \( \times \) 2 polarities (ON or OFF). The equations for the activities \( r_i^+ \) and \( r_i^- \) at the first level of ON and OFF cell processing, respectively, are defined as follows:

\[
\frac{dr_i^+}{dt} = -Dr_i^+ + (U - r_i^+)F_i^+ - (L + r_i^+)F_i^-.
\]

(11)

and

\[
\frac{dr_i^-}{dt} = -Dr_i^- + (U - r_i^-)F_i^- - (L + r_i^-)F_i^+.
\]

(12)

The excitatory \( F_i^+ \) and inhibitory \( F_i^- \) feedforward activities (related directly to the image) are defined by:

\[
F_i^+ = \sum_k G_i^+(k-i)I_k
\]

(13)

and

\[
F_i^- = \sum_k G_i^-(k-i)I_k.
\]

(14)

where \( \sigma^+ = 0.3 \) and \( \sigma^- = 0.9 \). The kernels are shown in Figure 20a. The signal that is passed on to the next level of retinal processing is defined by \( P_i = M_p \max(r_i, 0) \). The signal of an ON cell is denoted by \( P^+ \) and of an OFF cell by \( P^- \). Here \( G_i^+ \) is a narrow center Gaussian kernel, and \( G_i^- \) is a wider surround Gaussian kernel. The kernels are flipped for the OFF cells, whose signals are denoted by \( P^- \). The simulated responses are shown in Figure 19c.

#### Figure 20

The opponent processing of ON and OFF cell signals is modulated by a chemical transmitter process that can multiply, or gate, the transmitted strength of activity towards the next level. For each location, there is a transmitter gate that obeys the equation (Grossberg, 1972, 1980):

\[
\frac{dg_i}{dt} = A(B - g_i) - C(P_i + T)g_i.
\]

(15)

In equation (15), parameter \( A \) defines the rate of transmitter accumulation, \( B \) gives the maximal level of accumulated transmitter, and \( C \) 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, or tonic, level of activity. This background level of activity can be interpreted as intrinsic noise within a field of neurons.
final opponent output of retinal ON and OFF cells is given by:

\[
R_i^+ = M_v \max((P_i^+ + T)g_i^+ - (P_i^- + T)g_i^-, 0)
\]  

(16)

and

\[
R_i^- = M_v \max((P_i^- + T)g_i^- - (P_i^+ + T)g_i^+, 0),
\]  

(17)

respectively.

A second upper index indicates which retina a cell belongs to (left or right), thus there are the following variables at this level: \(R_i^{l+}, R_i^{l-}, R_i^{r+}, R_i^{r-}\). Strictly speaking, there ought to be a neuronal field between the activities \(P_i\) and \(R_i\) at which the background level activity \(T\) is added to the ON and OFF cell signals \(P_i^+\) and \(P_i^-\). The intermediate field is then gated by the \(g_i\). It is assumed that these cells equilibrate more rapidly to the input than the transmitters, and hence are solved at equilibrium. This assumption reduces the number of differential equations and accordingly speeds up simulations.

When there is a sudden increase of light at a given spot, then the corresponding activity \(P_i^+\) increases, and hence \(g_i^+\) slowly decays. Thus the initial response of \(R_i^+\), which multiplies \(P_i^+\) and \(g_i^+\), is initially strong and then becomes weaker (Figure 19d). When the input shuts off, then \(P_i^+\) quickly returns to zero. However, \(g_i^+\) takes longer to reach its new resting level. Thus there will be a period during which all signals \(P_i\) are zero, but \(g_i^-\) is bigger than \(g_i^+\). Due to the tonic input, this imbalance leads to a transient response in the opponent channel \(R_i^-\). This response is the antagonistic rebound.

4.4 LGN

There are 4 fields of neurons at the LGN level: 2 eyes \(\times\) 2 polarities (ON or OFF). The shunting equations that define LGN ON and OFF activities \(l_i^+\) and \(l_i^-\), respectively, are as follows:

\[
\frac{dl_i^+}{dt} = -Dl_i^+ + (U - l_i^+)(R_i^+ + B_i^+) - (L + l_i^+)B_i^-
\]  

(18)

and

\[
\frac{dl_i^-}{dt} = -Dl_i^- + (U - l_i^-)(R_i^- + B_i^+) - (L + l_i^-)B_i^-
\]  

(19)

The specific excitatory \((B_i^+\)) and nonspecific inhibitory \((B_i^-\)) feedback signals from hypercomplex cells are given by:

\[
B_i^+ = M_h^+ \sum_{k,d} G_h(i - k, d)H_{kd}
\]  

(20)

\[
B_i^- = M_h^- \sum_{k,d} H_{kd}
\]  

(21)

The terms \(H_{kd}\) denote signals from the hypercomplex cell stage; see equation (37). The hypercomplex kernel \(G_h\) is shifted by 0.5 to compensate for the shift that arises in the transition from the LGN to simple cells. In other words, the grid corresponding to the hypercomplex cell activities is shifted by half a pixel with respect to the grid of the LGN.
cells:

$$G_h(y, d) = k \exp \left( -\frac{(y - 0.5 + ed)^2}{2\sigma^2} \right)$$

(22)

The index \( d \) denotes the disparity (which can be -3, 0, or 3) of the hypercomplex cell, \( e = -1, 1 \) denotes the ocularity (left or right) of the LGN to which feedback is going, and \( k \) is chosen to normalize the kernel. The kernel is shown in Figure 20b.

The LGN output signal is defined as follows:

$$L_i^+ = \max (l_i^+, 0)$$

$$L_i^- = \max (l_i^-, 0).$$

(23) (24)

A second upper index indicates which LGN a cell belongs to (left or right), thus there are 4 types of output signals from the LGN: \( L_i^{++}, L_i^{+-}, L_i^{-+}, L_i^{--} \) are shown in Figure 19e.

4.5 Simple cells

There are 4 fields of neurons at the simple cell level: 2 eyes \( \times 2 \) orientations (light-dark or dark-light). The responses of simple cells are built up from convolutions of the LGN cell responses with odd-symmetric kernels:

$$s_i^+ = \sum_k K_s(i - k)L_k^+,$$

(25)

and similarly for \( s^- \). In this expression, \( K_s \) is an odd-symmetric kernel such that:

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

(26)

where \( \sigma = 0.3 \) gives the width of the kernel, and \( k \) normalizes the kernel. See Figure 20c. In this kernel, \( y \) is shifted by 0.5 so that the simple cell is positioned between a pair of LGN cells. This affords good edge localization.

Simple cell responses \( S_i^+ \) and \( S_i^- \) are derived from \( s_i^+ \) and \( s_i^- \) as follows (see Figure 11):

$$S_i^+ = M_s \max(s_i^+ - s_i^- - \alpha |s_i^+ + s_i^-|, 0)$$

(27)

and

$$S_i^- = M_s \max(s_i^- - s_i^+ - \alpha |s_i^+ + s_i^-|, 0),$$

(28)

where the upper indices stand for dark-light (+) and light-dark (-) edges, \( M_s \) is a scaling constant, and \( \alpha \) reduces spurious responses. The activities \( s_i^+ \) and \( -s_i^- \) give the contributions of the ON and OFF cells to the dark-light simple cell. Both need to be sufficiently active to fire the simple cell. This type of model is supported by neurophysiological studies of simple cells in the cat and monkey (Perster, 1988; Liu et al., 1992; Pei et al., 1994). The absolute value term \(-\alpha |s_i^+ + s_i^-|\) ensures that the simple cell does not fire when both \( s_i^+ \) and \(-s_i^-\)
are activated, or if only one of the two is active. This rule combines two similar simple cell models previously described (Gove et al., 1995; Grossberg & McLoughlin, 1995).

Another upper index is added to (27)-(28) to denote the eye of origin (l or r): $S^l_{i+}$, $S^l_{i-}$, $S^r_{i+}$, $S^r_{i-}$. The simple cell signals are shown in Figure 19f. Note how the rebound responses generated at the retinal stage lead to rebound responses at the simple cell stage.

### 4.6 Complex cells

At the complex cell stage there are 3 fields of complex cells: one each for zero, crossed and uncrossed disparities. The disparities that were used are 0, -3, 3. A disparity of -3 means that the left image has been shifted by -3 (3 to the left), and the right image by 3 (3 to the right). So the actual distance between corresponding points is six units. Associated with each complex cell is also an inhibitory interneuron, as in Figure 14. The equations for the excitatory complex cells $c^+_id$ and the inhibitory interneurons $c^-id$ are as follows:

$$\frac{dc^+_id}{dt} = -Dc^+_id + (U - c^+_id)(F^+_id + B^+_id) - (L + c^-id)(F^-id + B^-id + \beta c^-id)$$  \hspace{1cm} (29)

and

$$\frac{dc^-id}{\epsilon \ dt} = -Dc^-id + (U - c^-id)f(c^+_id) - (L + c^-id)F^-id.$$  \hspace{1cm} (30)

The parameter $\beta$ in (29) 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 complex cell positive feedback $B^+_id$ after inputs shut off. The small parameter $\epsilon$ in (30) ensures that the inhibitory interneuron reacts more slowly than the complex cell.

The feedforward activities define the binocular disparity filter between simple and complex cells (Grossberg & McLoughlin, 1995):

$$F^+_id = M^f_c \left| \sum_k G^+_s(k - i, d)S^l_k - \sum_k G^+_s(k - i, d)S^r_k \right|$$

$$- \sum_k G^+_s(k - i, d)S^l_k + \sum_k G^+_s(k - i, d)S^r_k$$  \hspace{1cm} (31)

and

$$F^-id = M^f_c \left| \sum_k G^-s(k - i, d)S^l_k - \sum_k G^-s(k - i, d)S^r_k \right|$$

$$- \sum_k G^-s(k - i, d)S^l_k + \sum_k G^-s(k - i, d)S^r_k,$$  \hspace{1cm} (32)

where $|$ denotes the absolute value and $M^f_c$ scales the strength of feedforward activities. The difference within each absolute value expression ensures that maximal activation occurs when simple cells of the same polarity are active in the two eyes, as in Figure 12b. If the
polarities differ, only a weak signal can be generated. The absolute value operation, on the other hand, performs a full-wave rectification which ensures that the feedforward signal does not depend on what polarity the simple cells have. Each of these full-wave rectification terms may be interpreted as arising from the sum of two half-wave rectification terms that respond to one or the other polarity match, but not both, before the results are pooled at the complex cells; see Grossberg and McLoughlin (1995). 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 output pools opposite contrast polarities, which has been viewed as a defining characteristic of complex cells (Hubel & Wiesel, 1962; Gilbert, 1977; Skottun et al., 1991). This property is illustrated in Figure 12b. For complex cells at zero disparity ($d = 0$), the feedforward weight $M_c^f$ is scaled by a factor of 1.05 times its value for non-zero disparity cells. This factor ensures that, during monocular presentation, cells at the zero disparity level respond maximally.

The feedforward inhibitory Gaussian kernels $G^-$ in (31) - (32) are not disparity tuned. They are characterized by parameter $\sigma^-$, as in equation (10). The feedforward excitatory kernels $G^+$ are disparity tuned. The left and right kernels are given by:

$$G^+_s(y, c) = G(y - c)$$
$$G^+_r(y, c) = G(y + c)$$

with $\sigma^+ = 0.3$, where $d$ gives the disparity shift of a given complex cell. As noted above, the disparities that are used in these simulations are -3, 0 and 3. Since $G(y)$ is a Gaussian that is centered on 0 as in equation (10), $G^+_s$ peaks at $y = d$ and $G^+_r$ peaks at $y = -d$. For $d = 3$, this means that the left kernel is shifted rightwards, and the right kernel is shifted leftwards; i.e., $d = 3$ corresponds to crossed disparities. Similarly, $d = -3$ corresponds to uncrossed disparities, and $d = 0$ corresponds to zero disparity. The convolution with these kernels in equation (31) implies that the input to a given complex cell will be maximal when the simple cell activities from the two eyes not only have the same polarity, but are also offset by the correct amount. When the disparity of a kernel does not match the disparity of simple cell activities (and therefore of contrast edges in the image), that particular complex cell will not survive the feedback competition defined by (29).

The feedback signals that realize the competition in equation (29) are given by:

$$B_{id}^+ = M_c^b \sum_{j,c} G^+_c(j - i)f(c_{jc}^+)$$

and

$$B_{id}^- = M_c^b \sum_{j,c} G^-_c(j - i)f(c_{jc}^+)$$

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

The feedback signal function in (30), (35) and (36) is a faster-than-linear nonlinearity $f(x) = x^4$ in order to achieve winner-take-all dynamics (Grossberg, 1973). The feedback kernel parameters $\sigma^+$ and $\sigma^-$ are given in Table 1. The profiles of the feedforward kernels
are shown in Figure 20d, and those of the feedback kernels in Figure 20e. The complex cell outputs are shown in Figure 19g.

When the complex cells receive an input, the excitatory and inhibitory feedback interactions in (35) and (36) use the faster-than-linear signal function to contrast-enhance the input pattern. As indicated by equation (36), competition occurs across space and also across disparities. Since the competitive interactions are limited in their spatial extent, several winners can emerge, although within each local spatial region, there is only one winner. Once a winner has emerged, excitatory feedback $B_{id}^f$ in (29) can keep the winner active in short term memory. This raises the problem of how to reset the reverberating activity when its input shuts off, so that the next input may be correctly processed. The inhibitory interneurons in (30), as in Figure 14, are posited herein to reset active complex cells after their input shuts off. A functionally similar alternative to this reset mechanism is one which uses a habituative transmitter gate, (15), as in the retinal opponent process (16) - (17), within the recurrent competitive field (Grossberg, 1976a). Such a gate would occur in each complex cell feedback pathway. As a complex cell becomes active, its gate variable habituates, much as in equation (15). The net feedback signal gets weaker but persists at a lower level. When a new input arrives, it is then able to reset the previous winner, since the latter's influence on the network is decreased. When a new winner emerges, the transmitter gate of the previous winner can recover, so that it can become a winner at a later stage.

Both the inhibitory interneuron in (30) and a habituative transmitter gate have a slower rate of integration than the complex cell. This property allows the complex cell to remain active for a while after its input shuts off. This activity persists while rebound responses in the model retina and LGN occur. These rebound responses cause simple cells of opposite polarity to re activate the persistent complex cell activity, since simple cells of opposite polarity converge upon a single complex cell, by (31). In this way, a single input produces a single complex cell response, rather than two separate responses. This property is shown elsewhere to enable oppositely polarized simple cells to become associated with the same complex cell during the developmental process whereby fine disparity tuning is learned (Grunewald & Grossberg, 1995a).

### 4.7 Hypercomplex cell

In the present one-dimensional model, hypercomplex cells are a replica of the complex cell signals. Thus:

$$H_{id} = f(c_{id}^+)$$

(37)

in the LGN equations (20) and (21).

### 4.8 Closed Loop Model Dynamics

So far, the organization of each individual level has been discussed. Since there is long-loop feedback from hypercomplex cells to the LGN, the dynamics of the model as a whole exhibit a greater level of complexity. Feedback from hypercomplex cells occurs only when a winner has emerged at the complex cell level. The inhibitory ($G_{X}^{-}$) and excitatory ($G_{X}^{+}$) feedback kernels from hypercomplex cells to the LGN in (18) and (19) are chosen to selectively activate those
regions of the LGN that contributed to the winner. In other words, the feedback signals tend to match the feedforward excitation. The cells that receive feedback activation, in turn, pass their increased activation on to the complex cell stage. As a result of this positive feedback loop, the complex cell winner gets even stronger. Figure 21 shows the complex cell activities in the example used previously when there is no feedback to LGN. The relative advantage of winners over the other close matches is decreased. It takes longer for cells in the complex cell field to reach maximal activity. Moreover, the lack of corticogeniculate inhibition means that low rebound responses are not suppressed at the LGN, and therefore complex cell activity persists longer at disparity -3, and complex cells at disparity 3 get incorrectly activated for a brief time.

Figure 21

4.9 Implementation details

Each field of neurons has the same size. In the present simulations, a field of 100 units was used. 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.

5 Discussion

This article describes a model of binocular disparity processing which includes corticogeniculate feedback interactions. The model is able to dynamically process information obtained from the two eyes, and to compute a spatial representation of the disparities of the images from the two eyes at model cortical complex cells. Although these cells self-normalize their activities through time, they also exhibit binocular summation in response to brief and low contrast stimuli. The model can simulate the Pulfrich effect, thereby illustrating how signals from the two eyes can get desynchronized at early stages of processing, and lead to erroneous percepts. The model is capable of fusion of slightly delayed anticorrelated stereograms by using the rebounds at input offset in opponent processing channels.

At each stage within the model local interactions between neurons are used to generate global behavior. The model is part of a larger theory of 3-D vision (Grossberg, 1994), within which similar mechanisms based on local neuronal processing are used to generate a 3-D percept of the visual environment. The applicability of this theory to 3-D vision manifests itself in its ability to explain such diverse effects as binocular rivalry, da Vinci stereopsis, and neon color spreading, among others. The present work shows how this theory can be extended to include various dynamical properties of binocular vision.
Acknowledgments

Stephen Grossberg was partially supported by the Air Force Office of Scientific Research (F49620-92-J-0499), the Advanced Research Projects Agency (ONR N00014-92-J-4015), and the Office of Naval Research (ONR N00014-95-1-0409 and N00014-95-1-0657). Alexander Grunewald was partially supported by the Air Force Office of Scientific Research (F49620-92-J-0334 and F49620-92-J-0225), the Advanced Research Projects Agency (ONR N00014-92-J-4015), and the Office of Naval Research (N00014-91-J-4100). Present address: Division of Biology, California Institute of Technology, Mail Code 216-76, Pasadena, CA 91125. The authors wish to thank Robin Locke for her valuable assistance in the preparation of the manuscript.
References


Table 1. The parameters used in the binocular model.

Figure 1. Since the two eyes are horizontally displaced, the two retinal images are not identical. The two images of an object in the line of sight are at the same retinal locations in both eyes, in this example a person. The two images of objects that are slightly in front or behind the fixation point differ by a slight horizontal displacement, or disparity. In this example, the flower is shifted rightwards in the left eye, and leftwards in the right eye. The table, on the other hand, is shifted leftwards in the left eye, and rightwards in the right eye. The visual system uses these disparities to calculate the depth of an object.

Figure 2. When an observer looks around in a visual scene, and fixates different objects, the disparities of the images of these objects on the two retinae change. Thus when looking at the person, the images corresponding to the person have zero disparity, but when looking at the table, the two images differ, thus creating a disparity. Similarly, the disparities of the flower differ, depending on whether the fixation point is the person, the table, or the flower.

Figure 3. The Pulfrich effect is elicited by viewing a pendulum move first to the right and then to the left while one eye (here the right eye) is covered with a dark glass (top part of Figure). The pendulum appears to be moving in depth, coming closer when the pendulum moves rightward, and receding in depth when the pendulum is moving leftward (bottom part of Figure).

Figure 4. In a correlated stereogram, the left and the right images have the same contrast polarities, but the images may differ due to disparity differences between the left and the right images. In this example, all dots have zero disparity, except the two middle dots, which are slightly shifted to the left in the left image, and to the right in the right image.

Figure 5. In an anticorrelated stereogram the left and the right images have opposite contrast polarities, and the images may differ due to disparity differences between the left and the right images. In this example all dots have zero disparity, except the two middle dots, which are slightly shifted to the left in the left image, and to the right in the right image.

Figure 6. Model processing stages. The retina receives visual information, extracts contrast information and passes it on to the lateral geniculate nucleus (LGN), where it is combined with corticogeniculate feedback signals. Filtering of orientation-selective simple cell activities leads to a disparity-selective the complex cell stage. Processing is hierarchically organized, but feedback signals from the hypercomplex cell stage to the LGN play an important part in confirming complex cell activities.

Figure 7. Model architecture. The left and right images impinge on the left and right retinae respectively. The retina processes contrast information using ON and OFF cells whose opponent signals are separately transmitted to the lateral geniculate nucleus (LGN), where they are combined with corticogeniculate feedback signals. The simple cell stage combines LGN responses to yield orientation selectivity which retains information about the polarity of image contrast. The complex cell stage pools rectified simple cell activities across polarities, across the two eyes, and across space. Complex cells can binocularly match like
polarities, yet also pool opposite polarities, to become selective to disparities between left and right images. The hypercomplex cell stage generates feedback signals to the LGN.

**Figure 8.** ON and OFF retinal cells: The image is convolved with center and surround kernels, which are subtracted from each other to yield ON and OFF cell responses.

**Figure 9.** Opponent processing of retinal ON cells (left column) and OFF cells (right column) by a gated dipole: Transmitter inactivation leads to habituation (in square synapses) of the response to a persistent input. Habituation combined with inhibition between ON and OFF channels, half-wave rectified output signals, and tonic background activity result in an antagonistic rebound response at the offset of a stimulus. The ON channel (left hand column) responds to a phasic step input on a constant tonic background (lowest left graph) by habituating its transmitter (next lowest graph). The phasic-plus-tonic ON signal is multiplied by the transmitter to generate overshoot and undershoot responses (next lowest graph). The OFF channel (right hand column) responds to only the constant tonic background (lowest right graph) which creates a constant baseline level of habituation and output (next two graphs). The habituated OFF channel output is subtracted from that of the ON channel by an opponent interaction. The result is half-wave rectified to generate a habituative, but sustained, ON response (upper left graph). When the same is done in the OFF channel, a transient antagonistic rebound occurs at the offset of the input (upper right graph).

**Figure 10.** Combined bottom-up and top-down processing at the LGN stage. Feedforward signals from the retina excite LGN cells. Topographic feedback signals amplify LGN activities if feedforward activities match feedback activities. If there is no match, the nonspecific feedback inhibition decreases LGN activities.

**Figure 11.** Simple cells are excited by LGN ON cells and spatially displaced LGN OFF cells (left figure). Opposite polarity simple cells are due to different spatial distributions of inputs from ON and OFF cells. Opposite polarity simple cells compete before generating rectified output signals. In response to a light-dark vertical contrast (middle figure), only one simple cell of the pair gets activated, so an output is generated. In response to an input of spatially uniform luminance (right figure), both simple cells are equally activated and mutually inhibit one another.

**Figure 12.** (a). Feedforward processing at the complex cell stage. Complex cells pool across simple cell polarities, across space and across eyes. A complex cell is maximally excited if the disparity between simple cells matches its preferred disparity, and if the simple cells are of the same polarity. Input from the simple cell stage is arranged in a center-surround fashion, whereby simple cells excite nearby complex cells, but inhibit more distant complex cells. (b). Disparity tuning of complex cells]Disparity tuning at complex cells is obtained by pooling activities from simple cells of the same polarities from the two eyes. After adding activities for each polarity, the activities corresponding to the two polarities are subtracted and then the absolute value is taken. This ensures that only simple cells of the same polarities can be fused, while rendering the complex cell insensitive to the direction of contrast.

**Figure 13.** (a). Feedback interactions at the complex cell stage across space. A complex cell excites itself, but inhibits complex cells that are further away in the network. (b). Feedback interactions at the complex cell stage across disparities. A complex cell excites itself, but inhibits complex cells that code different disparities.

28
Figure 14. Complex cells reset circuit interneurons. See text for details.

Figure 15. The stimulus used in the simulations of the Pulfrich effect is a moving bar. The right image is weakened by a factor of 3 in comparison to the left image. Initially the bar is at rest, then it moves rightwards. After a brief pause it moves leftwards. The bar is at zero disparity.

Figure 16. Simulation of the Pulfrich effect. The direction of motion determines the perceived depth of the object. When the bar is not moving, it is perceived at zero disparity. When moving towards the covered eye, it is seen with disparity 3 (crossed), i.e. in front. When it is moving in the other direction, it is seen at disparity -3 (uncrossed), i.e. behind. Top: activities of complex cells of disparity -3, middle: activities of complex cells of disparity 0, bottom: activities of complex cells of disparity 3.

Figure 17. (a). The stimulus used in the simulations of fusion of delayed anticorrelated stereograms. First a bar is presented to the left eye, and then a bar of opposite polarity is presented to the other eye. The two bars are at a disparity of -3 with respect to each other. (b). Simulation of fusion of delayed anticorrelated stereograms. The first stimulus has no disparity (since the stimulus is monocular), and therefore the cells at zero disparity are activated. When the delayed anticorrelated stimulus is presented, the new stimulus fuses with the rebound response of the first stimulus. The stimulus is seen at disparity -3. Top: activities of complex cells of disparity -3, middle: activities of complex cells of disparity 0, bottom: activities of complex cells of disparity 3.

Figure 18. The activities of the complex cell field at disparity 0 when the same brief low contrast stimulus is presented monocularly (left) and binocularly (right). Note that the complex cell responses are half as strong in the monocular case than in the binocular case.

Figure 19. (a). Two flashes that are used as an example to illustrate processing within the model. The first stimulus comes on at \( t = 0 \) and it has a disparity of -3 (top). It is seen behind the fixation point. The second stimulus comes on at \( t = 8 \), and it has zero disparity (bottom). It is seen at the same depth as the fixation point. The two stimuli are also spatially offset. (b). The raw image falling on the two retinæ. The left image depicts the input of the left retina, the right image that of the right retina. (c). The activities at the first stage of retinal processing. Top row: ON responses, bottom row: OFF responses. Left column: left retina, right column: right retina. (d). Opponent output signals from the model retina. Top row: ON responses are habituative but sustained. Bottom row: OFF responses are transient rebounds. Left column: left retina, right column: right retina. (e). Outputs from the LGN. The top row shows ON responses, and the bottom row shows OFF responses. Left column: left LGN, right column: right LGN. (f). Neuronal responses of the simple cells. Top row shows the responses of dark-light cells, and the bottom row responses of light-dark cells. Left column: left simple cells, right column: right simple cells. (g). Complex cell outputs at disparity -3 are correctly activated. Cells at disparity 0 also get activated before the offset of complex cells at disparity -3, since they are responding to the rebound response at the simple cell stage. Top: activities of complex cells of disparity -3, middle: activities of complex cells of disparity 0, bottom: activities of complex cells of disparity 3.

Figure 20. (a). Center and surround kernels at the retinal stage. (b). Feedback kernel used for LGN processing. (c). Feedforward kernels of simple cells. (d). Feedforward kernels
at the complex cell stage of a cell of disparity -3. The left graph depicts the kernels with which left simple cells are convolved, the right those with which the right simple cells are convolved. The shift of the kernels leads to disparity sensitivity. For cells of disparity +3, the left and right kernels are interchanged, and for zero disparity cells both kernels are centered at 0. Note that the inhibitory kernels are not disparity tuned. Figure 21. Feedback kernels at the complex cell stage. Excitatory kernels are narrowly tuned, while inhibitory kernels are broadly tuned. Feedback kernels are not disparity tuned.

Figure 21. Complex cell activities in the absence of feedback from hypercomplex to LGN cells. As before, the complex cells at disparity -3 respond to the first stimulus. The response lasts longer than before, and cells at disparity 3 respond (incorrectly) after the cells at disparity -3 shut off. Top: activities of complex cells of disparity -3, middle: activities of complex cells of disparity 0, bottom: activities of complex cells of disparity 3.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D$</td>
<td>1</td>
<td>passive decay constant of cell activity</td>
</tr>
<tr>
<td>$U$</td>
<td>1</td>
<td>upper limit of cell activity</td>
</tr>
<tr>
<td>$L$</td>
<td>1</td>
<td>lower limit of cell activity</td>
</tr>
<tr>
<td>$A$</td>
<td>0.2</td>
<td>rate of gate recovery</td>
</tr>
<tr>
<td>$B$</td>
<td>1</td>
<td>baseline gate activity</td>
</tr>
<tr>
<td>$C$</td>
<td>2</td>
<td>active gate decay</td>
</tr>
<tr>
<td>$T$</td>
<td>0.3</td>
<td>background activity</td>
</tr>
<tr>
<td>$\sigma^+_r, \sigma^-_r$</td>
<td>0.3, 0.9</td>
<td>width of retinal kernels</td>
</tr>
<tr>
<td>$\sigma^+_g, \sigma^-_g$</td>
<td>0.6, 2</td>
<td>width of corticogeniculate feedback kernels</td>
</tr>
<tr>
<td>$\sigma_s$</td>
<td>0.3</td>
<td>width of simple cell kernels</td>
</tr>
<tr>
<td>$\sigma^+_c, \sigma^-_c$</td>
<td>0.3, 5</td>
<td>width of feedforward complex cell kernels</td>
</tr>
<tr>
<td>$\sigma^{b+}_c, \sigma^{b-}_c$</td>
<td>0.3, 4</td>
<td>width of feedback complex cell kernels</td>
</tr>
<tr>
<td>$M_p$</td>
<td>10</td>
<td>coefficient of gate signals</td>
</tr>
<tr>
<td>$M_r$</td>
<td>200</td>
<td>coefficient of retinal signals</td>
</tr>
<tr>
<td>$M_s$</td>
<td>2</td>
<td>coefficient of simple cell signals</td>
</tr>
<tr>
<td>$M^+_d$</td>
<td>2</td>
<td>coefficient of feedforward complex cell signals (non-zero disparity)</td>
</tr>
<tr>
<td>$M^-_d$</td>
<td>2.1</td>
<td>coefficient of feedforward complex cell signals (zero disparity)</td>
</tr>
<tr>
<td>$M^+_b$</td>
<td>300</td>
<td>coefficient of feedback complex cell signals</td>
</tr>
<tr>
<td>$M^-<em>b, M^-</em>{\lambda}$</td>
<td>10, 1</td>
<td>coefficient of hypercomplex cell signals</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>1.3</td>
<td>simple cell threshold</td>
</tr>
<tr>
<td>$\beta$</td>
<td>20</td>
<td>weight of complex cell inhibitory interneuron</td>
</tr>
</tbody>
</table>

Table 1:
Visual Scene

Left Eye

Right Eye

Figure 1:
Figure 2:
Figure 3:
Figure 4:
Figure 6:
Figure 7:
Figure 8:
Figure 9:
Figure 10:
Figure 11:
Figure 12
Complex

Simple

Left Right Left Right

Figure 12b
Figure 13: a
Complex

Figure 13b
Figure 14:
Figure 15:
Figure 16:
Figure 17 a
Figure 17b
Figure 18:
Figure 19 a
Figure 19e
Figure 19d
Figure 19e
Figure 19f
Figure 19g
Figure 20a
Figure 20b
Figure 20c
Left feedforward kernels disparity: -3

Right feedforward kernels disparity: -3

Figure 20d
Feedback kernels

Figure 20e
Figure 21: