Neural Networks, Vol. 2, pp. 29–51, 1989
Printed in the USA. All rights reserved.

0893-6080/89 $3.00 + .00

ORIGINAL CONTRIBUTION

Stereo Boundary Fusion by Cortical Complex Cells: A System of Maps, Filters, and Feedback Networks for Multiplexing Distributed Data

STEPHEN GROSSBERG† AND JONATHAN A. MARSHALL‡

Center for Adaptive Systems, Boston University

(*Received* 3 *June* 1988; *revised and accepted* 2 *September* 1988)

Abstract—*A neural network model of multiple-scale binocular fusion and rivalry in visual cortex is described and simulated on the computer. The model consists of three parts: a distributed spatial representation of binocular input patterns among simple cells that are organized into ocular dominance columns; an adaptive filter from simple cells to complex cells; and a nonlinear on-center off-surround shunting feedback network that joins together the complex cells. This data structure generates complex cell receptive fields which multiplex input position, orientation, spatial frequency, positional disparity, and orientational disparity, and which are insensitive to direction-of-contrast in the image. Multiple copies of this circuit are replicated in the model using receptive fields of different sizes. Within each such circuit, the simple cell and complex cell receptive field sizes covary. Together these circuits define a self-similar multiple-scale network. The self-similarity property across spatial scales enables the network to exhibit a size-disparity correlation, whereby simultaneous binocular fusion and rivalry can occur among the spatial scales corresponding to a given retinal region. It is shown that a laminar organization of the model interactions among the complex cells gives rise to conceptually simple growth rules for intercellular connections. The output patterns of the model complex cells are designed to feed into the model hypercomplex cells at the first competitive stage of a Boundary Contour System network, where they trigger a process of multiple-scale emergent binocular boundary segmentation. The modeling results are compared with psychophysical data about binocular fusion and rivalry, as well as with the cepstrum stereo model of Yeshurun and Schwartz. The results indicate that analogous self-similar multiple-scale neural networks may be used to carry out data fusion of many other types of spatially organized data structures.*

Keywords—Neural network, Binocular vision, Visual cortex, Data fusion, Multiplexing, Binocular rivalry, Computational map, Complex cells, Ocular dominance columns, Self-similarity, Size-disparity correlation, nonlinear feedback network.

1. STEREO IMAGE MULTIPLEXING BY THE COMPLEX CELL CORTICAL MAP

Neural networks are often designed as nonlinear filters that multiplex several different types of information into a spatial map. They can thereby generate compressed codes capable of reacting selectively to combinations of external environmental conditions. The complex cells in the striate visual cortex form such a map. Individual complex cells in this map selectively fire in response to a prescribed range of values of an image contrast's position, orientation, spatial frequency, stereo positional disparity, and stereo orientational disparity, yet are chromatically broad-band and insensitive to direction-of-contrast (DeValois, Albrecht, & Thorell, 1982; Poggio, Motter, Squatrito, & Trotter, 1985; Shinkman & Bruce, 1977; Thorell, DeValois, & Albrecht, 1984; von der Heydt, Hänny, & Dürsteler, 1981).

These filter properties take on functional meaning as part of a neural network model of cortical visual processing, called the Boundary Contour System (Grossberg & Mingolla, 1985a, 1985b), wherein cortical complex cells play the role of broad-band

†Supported in part by the Air Force Office of Scientific Research (AFOSR F49620-86-C-0037 and AFOSR F49620-87-C-0018) and the Army Research Office (ARO DAAG-29-85-K-0095).

‡Supported in part by the Air Force Office of Scientific Research (AFOSR 85-0149, AFOSR F49620-86-C-0037 and AFOSR F49620-87-C-0018), the Army Research Office (ARO DAAG-29-85-K-0095), and the National Science Foundaton (NSF IRI-84-17756).

Acknowledgements: We wish to thank Cynthia Suchta and Carol Yanakakis for their valuable assistance in the preparation of the manuscript and illustrations.

Requests for reprints should be sent to Stephen Grossberg, Center for Adaptive Systems, Boston University, 111 Cummington Street, Boston, MA 02215.

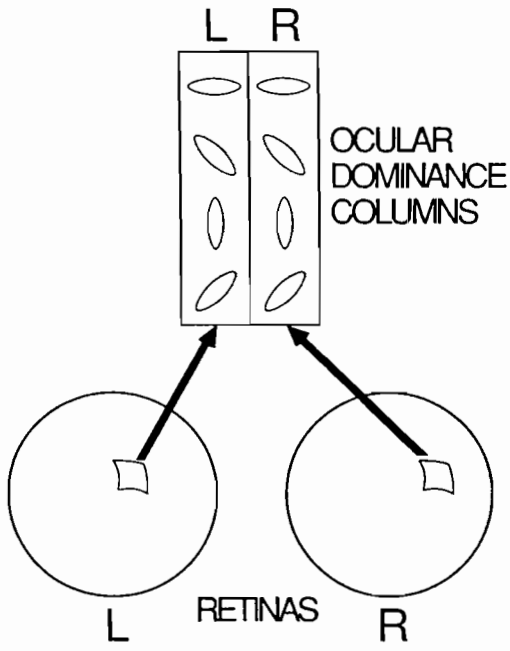

FIGURE 1. Small patches of left (*L*) and right (*R*) eye retinas project to contiguous regions of visual cortex wherein a set of oriented receptive fields, or local contrast detectors, selectively react to oriented retinal contrasts. Such a complete set of oriented *L* and *R* detectors corresponding to a pair of retinal patches is called a *hypercolumn* (Hubel & Wiesel, 1977).

boundary detectors. The present work describes computer simulations, announced in Grossberg and Marshall (1987), which clarify how the complex cell filter helps the Boundary Contour System to generate an emergent boundary segmentation that selects, regularizes, and completes binocularly consistent image data, while suppressing binocularly inconsistent image data (Grossberg, 1987a, 1987b, 1988a), including noise (Grossberg & Mingolla, 1987a). In technological applications, these properties of the complex cell filter are of interest in problems of multidimensional data fusion.

2. THE CORTICAL HYPERCOLUMN MAP

The model suggests how spatially distributed combinations of featural data at a processing level F_1 of cortical simple cells interact with a level F_2 of cortical complex cells via an adaptive filter in such a way that the individual cells at F_2 multiplex the types of information for which cortical complex cells are selective. In order to accomplish this data compression, level F_1 is modeled as a spatial map, in particular the spatial map whereby cortical simple cells are arranged into cortical modules called hypercolumns (Hubel & Wiesel, 1977). Each hypercolumn (Figure 1) brings together image data from two small corresponding regions of the left and right retinas. The retinal image data are filtered by the time they reach the hypercolumn stage into a topographically organized array of oriented filters, or receptive fields. Each hypercolumn contains a complete set of oriented filters from each eye. The hypercolumns are fit together in a spatial map in which image data from each eye form columns, called ocular dominance columns (Figure 2). Traversal of an ocular dominance column leads to a gradual change of retinal positional coordinates and of the preferred orientation of receptive field responsiveness.

Figure 3 schematizes how a scenic edge is converted into a pattern of activation across the simple cells of the model's hypercolumn map. Distinct combinations of edge positions and orientations on the two retinas are converted into distinct activation patterns across the hypercolumn map. Thus level F_1 of the model encodes important scenic parameters as distributed activation patterns across a spatially organized data structure.

3. A SELF-SIMILAR MULTIPLE-SCALE COMPETITIVE LEARNING SYSTEM

The network that converts distributed activation patterns across level F_1 into multiplexed responses of individual cells at level F_2 is a variant of a competitive learning model. Competitive learning models were developed through an interaction between Grossberg (1972, 1976) and von der Malsburg (1973), leading to several versions of the model, as well as mathematical results and computer simulations that formed the foundation for subsequent contributions (Amari & Takeuchi, 1978; Bienenstock, Cooper & Munro, 1982; Carpenter & Grossberg, 1987a, 1987b, 1988; Cohen & Grossberg, 1986, 1987; Grossberg & Kuperstein, 1986; Kohonen, 1982, 1984; Linsker, 1986a,

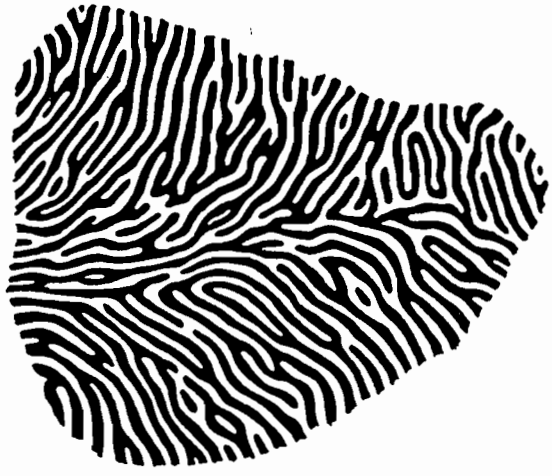

FIGURE 2. Hypercolumns in visual cortex are joined together sequentially to form ocular dominance columns. Here the black (white) bands represent sequential arrays of hypercolumns from different eyes. (Reprinted with permission from Hubel and Wiesel, 1979.)

1986b, 1986c; Rumelhart & Zipser, 1985; Singer, 1983; Willshaw & von der Malsburg, 1976). An historical discussion of the development of competitive learning models is provided in Grossberg (1987c, 1988a).

In a competitive learning model, normalized input patterns across level F_1 pass through an adaptive filter to a level F_2, whose design includes competitive interactions among its cell populations. Level F_2 contrast-enhances, or compresses, the signal pattern that it receives through the filter. Thus the activation pattern that is instated across F_2 is spatially more focused than its generative signal pattern. Only those F_2 cells

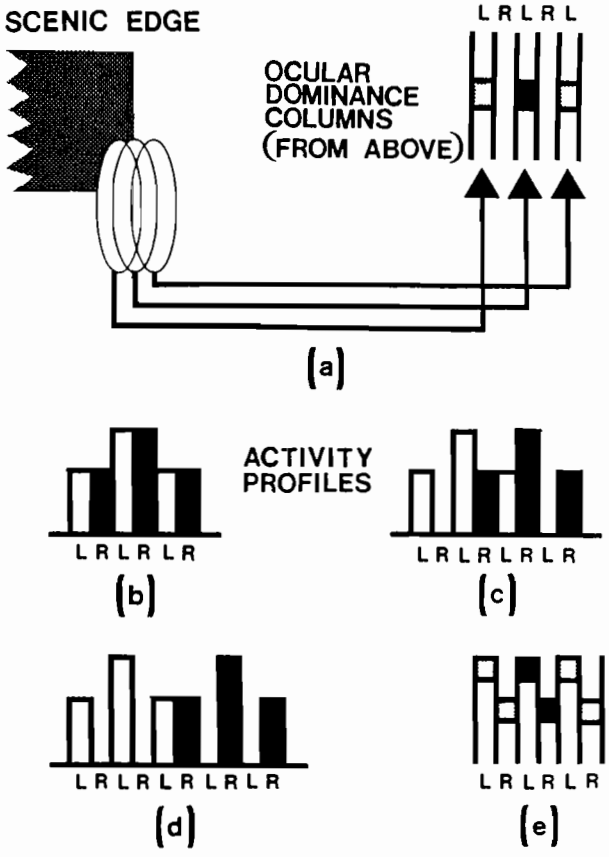

FIGURE 3. Translation of scenic contour information into spatial patterns of activity: (a) Overlapping like-oriented receptive fields generate a spatial pattern of activity at left-monocular representations in response to a left-monocularly viewed scenic edge. The figure portrays a view from above of cortical ocular dominance columns for the left (*L*) and right (*R*) eyes, and portrays increased cell activation with darker areas. (b)–(d) Binocular inputs due to a scenic edge viewed by the two eyes at increasing positional disparities create distinct, expanding activity patterns across the ocular dominance columns. Here bar heights code activities. (e) Binocular viewing can cause an orientational disparity that is coded by a positional shift in the activity pattern caused by the left eye relative to that caused by the right eye. This shift is perpendicular to the shift caused by positional disparity, which separates activity patterns caused by the two eyes in a horizontal rather than a vertical direction. (Reprinted with permission from Grossberg, 1987b.)

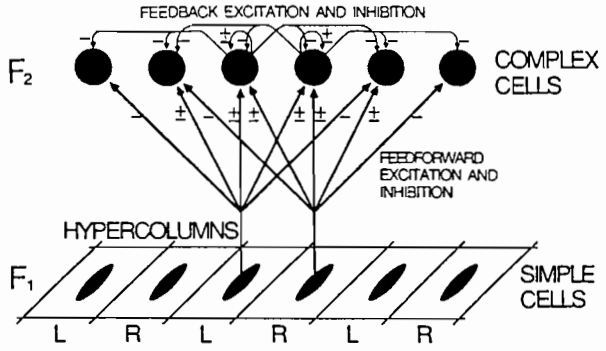

FIGURE 4. Visual signals broadly distributed across simple cells within the hypercolumn layer F_1 input to a layer F_2 of complex cells via feedforward on-center off-surround shunting interactions. Feedback on-center off-surround shunting interactions transform broad activations at F_1 into sharp, multiplexed activations at F_2. The Gaussian band-widths of the feedforward excitation and inhibition and the feedback inhibition covary with the spatial scale of the oriented receptive fields of the simple cells. See text for additional details.

which survive the competition with sufficiently positive activities can trigger associative pattern learning within the vectors of adaptive weights, or long-term memory traces, that sent them signals through the adaptive filter. Our model of the complex cell filter is a competitive learning system whose input level F_1 is modeled after the cortical hypercolumn map (Figure 4).

In order to support effective binocular processing by the Boundary Contour System, this competitive learning system needs to be designed with a multiple scale filter whose receptive fields come in several sizes that are related to each other by properties of self-similarity. A variant of such a self-similar multiple-scale competitive learning system is the masking field model, which was developed to explain properties of visual object recognition and speech recognition (Cohen & Grossberg, 1986, 1987; Cohen, Grossberg, & Stork, 1987; Grossberg, 1982). The remainder of this article focuses upon key self-similar multiple-scale aspects of the problem, rather than the learning aspects. Sections 4–8 motivate the type of self-similar multiple-scale processing that we suggest for stereopsis and depth perception processing. Section 9 begins a self-contained description of the model and computer simulations.

4. SIMPLE, COMPLEX, AND HYPERCOMPLEX CELL FILTERS OF THE BOUNDARY CONTOUR SYSTEM

The Boundary Contour System (BCS) was originally used to explain a large body of monocular visual data (Cohen & Grossberg, 1984; Grossberg & Mingolla, 1985a, 1985b; Grossberg & Todorović, 1988). Since its introduction, a variety of subsequent perceptual and neural data have provided additional support for the theory (see Grossberg, 1987a for a summary).

Several laboratories have tested the theory with positive results (e.g., Eskew, 1987; Graham, 1988; Kennedy, in press; Meyer & Dougherty, 1987; Todd & Akerstrom, 1987). It was also shown that the monocular theory may be embedded in a self-consistent fashion into a theory of binocular vision (Grossberg, 1987b).

Figure 5 schematizes the first four stages of filtering postulated by the monocular theory. The binocular theory preserves the qualitative properties of this filter, but substantially refines its description. In the first stage, pairs of simple cells sensitive to similar positions, orientations, and spatial frequencies, but opposite directions-of-contrast, rectify their output signals before adding them to form the input to a complex cell. Such a complex cell is thus sensitive to position, orientation, and spatial frequency, but insensitive to direction-of-contrast. This model of how rectified outputs from simple cells summate to activate complex cells has subsequently been partially supported by cortical data of Spitzer and Hochstein (1985).

The complex cells, in turn, are postulated to activate a subsequent filter whose interactions carry out two different types of spatially short-range competitive interaction, which give rise to two types of hypercomplex cells at successive processing stages of the BCS. Graham (1988) has recently provided additional evidence that a pair of hierarchically organized oriented filters, separated by a rectification stage, are engaged by textured images. At the first stage of the second filter, each complex cell activates an on-center off-surround interaction among like-oriented cells (Figure 5). In other words, each complex cell excites the hypercomplex cells of like-orientation corresponding to the same position at the next processing stage, but inhibits the hypercomplex cells of like-orientation that occur at nearby positions. These lateral inhibitory, or competitive, interactions carry out an operation of end stopping (Hubel & Wiesel, 1965; Orban, Kato, & Bishop, 1979) that gives rise to the hypercomplex receptive field properties at the first competitive stage of the model (Figure 5).

The second competitive stage of the BCS model executes a push–pull competition within each position between cells of different orientational preference, with the greatest inhibition occurring between perpendicular orientations. This competitive interaction gives rise to a type of higher-order hypercomplex cell receptive field (Grossberg, 1988b; Hubel & Wiesel, 1965).

Using these properties, we review those data and issues described in Grossberg (1987b) which most vividly motivate how to extend the monocular circuit of Figure 5 to a multiple-scale self-similar binocular circuit.

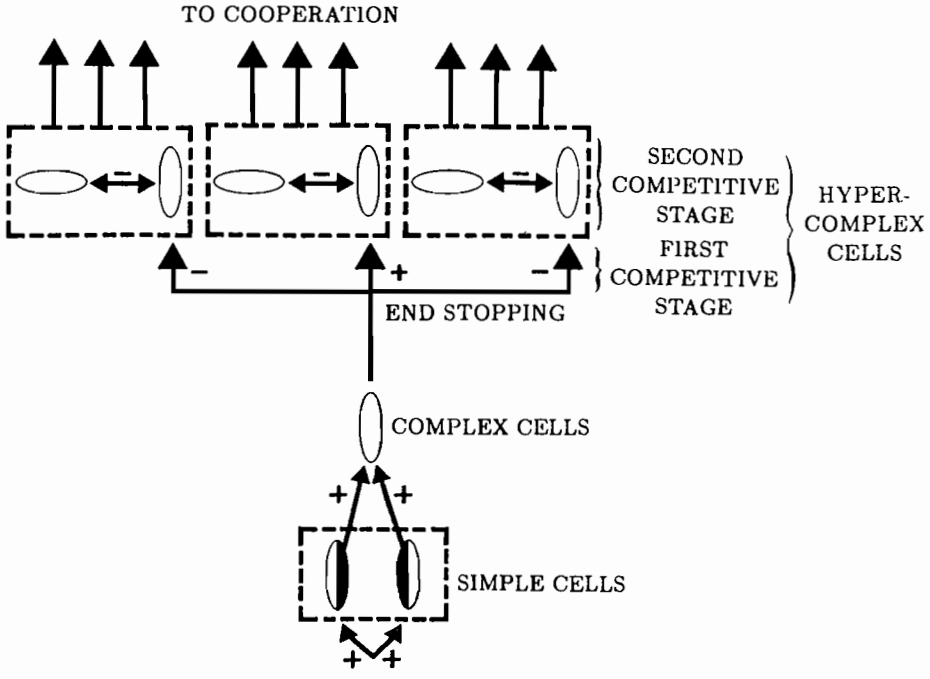

FIGURE 5. Early stages of BCS processing: At each position there exist simple cells with elongated receptive fields of various sizes which are sensitive to orientation, amount-of-contrast, and direction-of-contrast. Pairs of such simple cells sensitive to like orientation but opposite directions-of-contrast (lower dashed box) input to complex cells that are sensitive to orientation and amount-of-contrast but not to direction-of-contrast (white ellipses). The complex cells, in turn, excite like-oriented hypercomplex cells corresponding to the same position and inhibit like-oriented cells corresponding to nearby positions using the first competitive stage. At the second competitive stage (upper dashed boxes), higher-order hypercomplex cells corresponding to the same position but different orientations inhibit each other via a push-pull competitive interaction. (Reprinted with permission from Grossberg, 1987a.)

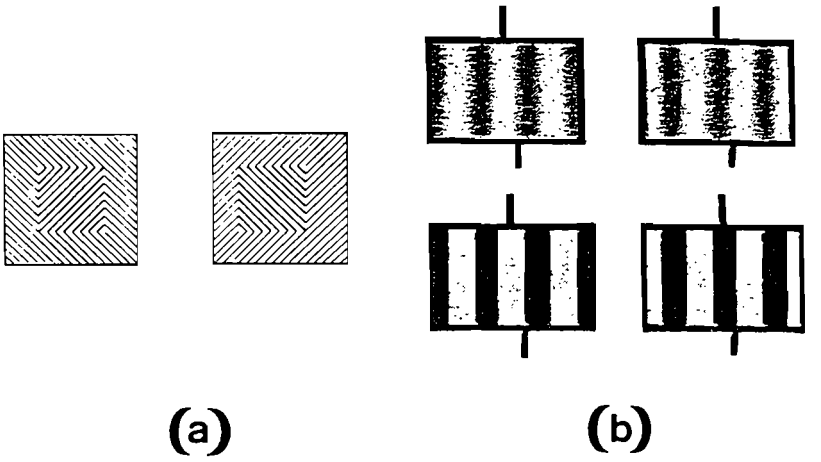

FIGURE 6. (a) The Kaufman stereogram induces a percept of a square in depth even while the perpendicular line patterns are rivalrous. (b) The Kulikowski stereograms illustrate the action of multiple spatial-frequency-sensitive scales during simultaneous fusion and rivalry: Sinusoidal gratings in antiphase can be fused to yield a percept of a grating in depth. Bar gratings in antiphase yield a percept of depth even though their edges are rivalrous.

5. BINOCULAR FUSION AND RIVALRY: THE KAUFMAN STEREOGRAM

The phenomena of binocular fusion and rivalry often occur simultaneously during the viewing of stereograms. The Kaufman (1974) stereogram (Figure 6a) illustrates several key points. When the two images are viewed through a stereoscope, a single fused square is seen in depth relative to the background. Superimposed upon the continuous percept of a square in depth is a rivalrous percept involving the oblique, mutually perpendicular lines. At any given moment, an observer can perceive either a 45° line or a 135° line at a given position, but not both. The percept switches intermittently between these mutually perpendicular lines.

Grossberg (1987b) traced this type of competition to the competition between perpendicular orientations that takes place at each position of the second competitive stage (Figure 5). In order for image data from both eyes to compete at the second competitive stage, the second competitive stage must be binocular. Thus data from both eyes are joined together no later than the second competitive stage, using the spatial map afforded by the ocular dominance columns in striate visual cortex to enable such eye-dependent, position-dependent, and orientation-dependent competition to occur.

6. BINOCULAR FUSION AND RIVALRY: THE KULIKOWSKI STEREOGRAM AND SPATIAL FREQUENCY DEPENDENCE OF STEREOPSIS

A second type of binocular rivalry is exemplified by the percepts seen in response to the Kulikowski (1978) stereograms (Figure 6b). During binocular viewing, subjects can fuse the two spatially blurred pictures. They perceive the fused image continuously in depth relative to the fused images of the two frames. Subjects experience a more complex percept when they view the stereogram composed from out-of-phase sharp bars. As in the case of the viewing of blurred bars, a fused form-in-depth is again perceived continuously. However, superimposed upon the fused percept of form-in-depth is a percept of binocular rivalry. The spatially out-of-phase edges of the bars in the left and right images are rivalrous and appear to pop alternately into and out of conscious perception.

This percept of simultaneous fusion and rivalry illustrates the operation of multiple spatial scales: The lower spatial frequency components of the two images can be binocularly fused into a continuously seen percept in depth at the same time that the high spatial frequency edges are seen to be rivalrous. Thus, the high spatial frequency scales that process the edges cannot fuse these high spatial frequency image properties at the same disparity at which the low spatial frequency scales can fuse the low spatial frequency contents of the two images. Such spatial frequency dependence of the effective disparity range for stereopsis, also called the size-disparity correlation, has been further studied by Richards and Kaye (1974), Schor and Wood (1983), and Schor, Wood, and Ogawa (1984). In contrast, the frames surrounding both images can be fused by all spatial scales because the frames can be spatially aligned with respect to both eyes. Thus, the Kulikowski stereogram provides visible evidence of the classical fact (Julesz, 1971) that a relative depth percept can be generated if different parts of two images are fused by different combinations of binocular spatial scales.

The relative depth percept that is generated by

the Kaufman stereogram (Figure 6a) can also be analyzed from this perspective. As described by Grossberg and Mingolla (1985a, 1985b), the formation of an emergent boundary segmentation around the illusory square region is initiated in the BCS by endcuts induced at the second competitive stage by the image corners at which the oblique line-ends of the square join its surround. Due to the horizontal disparity of the left and right images with respect to the image frame, not all spatial scales can form fused images from which to generate these emergent endcut boundaries. In contrast, the scenic edges that frame the pair of images can have zero disparity with respect to each other and can thus form fused boundary responses within all spatial scales. The selective activation of a subset of scales by the disparate image figures provides the basis for a relative depth difference of the figure with respect to the ground.

The Kaufman and Kulikowski stereograms illustrate that there exists an interaction between the scenic properties that can be fused within a given spatial scale and the emergent boundary segmentations that can be generated by that scale. This scale-specific interaction between binocular fusion and emergent boundary segmentation may be used to explain many of the striking data reported by Mayhew and Frisby (1976) and Frisby and Mayhew (1978) on rivalrous texture stereograms.

7. ANALYSIS OF SPATIAL-FREQUENCY DEPENDENT FUSION AND RIVALRY: RECONCILING BINOCULAR DEFORMABILITY WITH BINOCULAR ACUITY

In the Kaufman stereogram, *perpendicular* scenic edges that excite the two eyes at the *same* positions are rivalrous, thereby implicating the second competitive stage of the BCS (Figure 5). In the Kulikowski stereogram, *parallel* scenic edges that excite the two eyes at *disparate* positions are rivalrous. This is the type of competition that occurs at the first competitive stage of the BCS (Figure 5). The rivalry between the spatially disparate vertical edges of the Kulikowski stereogram is initiated when the first competitive stage causes spatially disparate, parallel orientations to compete at that copy of the second competitive stage which receives inputs from simple cells with small receptive fields.

Why does the first competitive stage not cause rivalry to occur between spatially disparate pairs of monocular images in *all* spatial scales? How can some spatial scales binocularly fuse images at the same disparity that at other spatial scales leads to binocular rivalry? Grossberg (1987b) suggested that spatially disparate pairs of monocular images that can be fused within a given scale both input to the on-center of the scale's first competitive stage (Figure 7a). In contrast, spatially disparate pairs of monocular images that are rivalrous within a given spatial scale compete via the off-surrounds of that scale's first competitive stage (Figure 7b).

This hypothesis suggests several conclusions. The input cells to the first competitive stage must be binocular, so that spatially disparate pairs of monocular images can both input through them to the on-center of the first competitive stage, as in Figure 7a. Grossberg (1987a) identified these cells with complex cells in area 17 of the striate cortex. Complex cells are, in fact, well known to be binocular (Hubel & Wiesel, 1977; Poggio, Motter, Squatrito, & Trotter, 1985). Thus, the hypercomplex cells at the first and second competitive stages are binocular because their inputs from the complex cells are already binocular. Figure 6 suggests that there exist disparities at which low spatial frequency scales can fuse pairs of monocular image elements at their complex cell level but high spatial frequency scales cannot. As a result, there exist disparities at which complex cells tuned to low spatial frequencies can excite the on-center

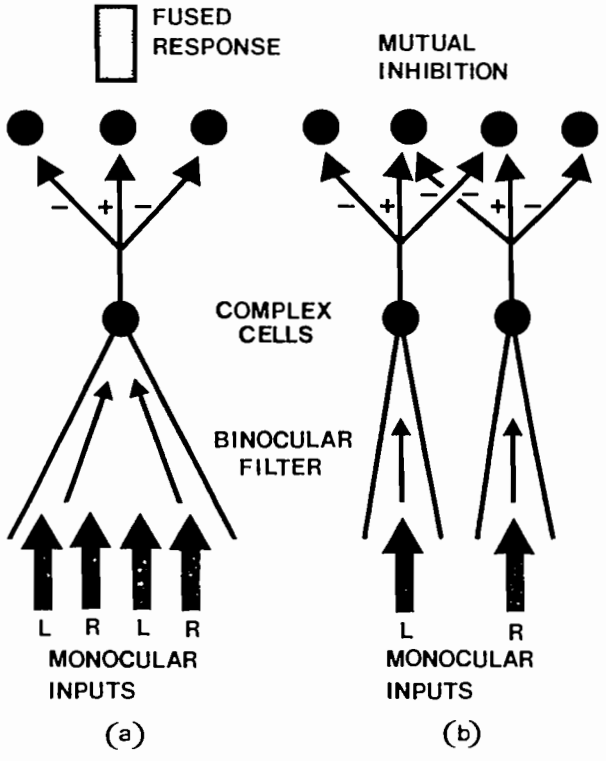

FIGURE 7. Initial processing stages leading to fusion or rivalry: (a) Spatially disparate monocular images from left (*L*) and right (*R*) eyes activate a shared population of complex cells, which in turn activate a "fused" locus of cells within an on-center of the second competitive stage. (b) At a smaller spatial scale, the same monocular images may activate spatially disjoint sets of complex cells, which input to off-surrounds at the second competitive stage, thereby initiating rivalry between the images. (Reprinted with permission from Grossberg, 1987b.)

of the corresponding second competitive stage (Figure 7a) but complex cells tuned to high spatial frequencies generate mutually inhibitory signals to the second competitive stage (Figure 7b). Figure 7 provides a pictorial way to understand how some spatial scales can fuse images that other spatial scales cannot, and how subsequent BCS mechanisms can inhibit binocularly discordant boundary signals within those spatial scales that are incapable of fusion.

Figure 7 also indicates how another important property of binocular space may be realized. In order to fuse pairs of monocular images that are spatially disparate, the binocular space must be *deformable*: two images must be deformed into one image, much as in the phenomenon of *displacement*, or *allelotropia* (von Tschermak-Seysenegg, 1952; Werner, 1937). In this phenomenon, when a pattern AB C is viewed through one eye and a pattern A BC is viewed through the other eye, the letter B can be seen in depth at a position halfway between A and C. Deformability implies that there exists a certain degree of positional uncertainty at the early stages of binocular processing, in particular at the simple cell stage where binocular scenic information is spatially distributed across the cortical hypercolumn map (Figure 3). The final binocular percept is, however, positionally sharp. A serious challenge to theories of binocular form perception is to reconcile the property of positional uncertainty, whereby deformability may permit binocular fusion, with the high degree of acuity that characterizes the final binocular percept.

One factor in our theory that contributes to an explanation is the manner in which *individual* complex cells in level F_2 multiplex *distributed activation patterns* across the level F_1 of simple cells. Because these F_1 activation patterns occur across interleaved ocular dominance columns, the multiplexed F_2 responses fuse together left eye and right eye cortical representations that are positionally disparate (Figure 7a). A second factor is the manner, described in Grossberg (1987b), in which the CC Loop of the Boundary Contour System selects a globally consistent binocular boundary segmentation from the total pattern of complex cell signals within each spatial scale, while sharpening and completing the representation via its context-sensitive cooperative–competitive feedback interactions as it rapidly approaches an equilibrium configuration.

Having motivated the functional role played by the size-disparity correlation in setting up the computation of simultaneous fusion and rivalry across multiple spatial scales, we now describe how this correlation may be achieved by imposing the property of self-similarity among the several spatial scales.

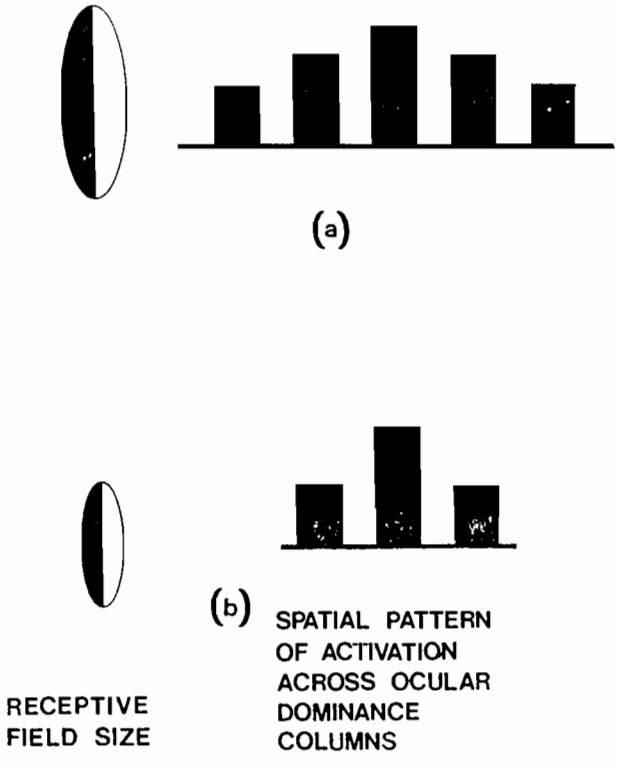

FIGURE 8. Early stages of spatial-frequency sensitivity are represented by covariation of receptive-field size with the spatial pattern of activity generated by cells of this size across F_1: Large receptive fields generate more spatially distributed patterns (a) than small receptive fields (b). Bar height represents activity at each cell position, as in Figure 3b–d, in response to an image edge. (Reprinted with permission from Grossberg, 1987b.)

8. SELF-SIMILAR SPATIAL SCALES AND THE CORTICAL MAGNIFICATION FACTOR

Let us assume, as in Figure 3, that the spatial dispersion of activation across the cortical hypercolumn map in response to localized image features covaries with the size of the simple cell receptive fields. Then larger simple cell receptive fields generate more broadly distributed activation patterns than smaller receptive fields, as in Figure 8. Assume that several different receptive field sizes exist. For simplicity of exposition, let each size be segregated into different, but topographically registered, layers within level F_1, such that each layer is organized into a hypercolumn map. Cells with larger receptive fields receive inputs from a larger retinal area than cells with smaller receptive fields. A layered organization of simple cells according to receptive field size provides one way to design a multiple-scale network of receptive fields, because cells in different layers can receive different amounts of receptive field scatter, or input fan-in. It is well-known that the amount of receptive field scatter increases with retinal eccentricity, according to the cortical magnification factor (Hubel & Wiesel, 1977). The same type of variable receptive field scat-

ter may provide a mechanism for achieving spatial frequency sensitivity across the several scale-specific layers that are posited to exist in the cortex corresponding to each retinal position.

Given a spatial organization of simple cells into hypercolumn maps with multiple layers of differently sized receptive fields, the *total activation pattern* across F_1 represents such image properties as monocular position, orientation, binocular positional disparity, binocular orientational disparity, *and* spatial frequency. Our task is to understand how these distributed activation patterns across level F_1 are converted into selective tuning curves for all these factors within individual cells of F_2, in such a way as to generate the size-disparity correlation described in Section 6.

The key hypothesis is that the input fan-in that controls the size of a simple cell's receptive field covaries with the output fan-out whereby a simple cell activates complex cells. In other words, the amount of spatial uncertainty that gives rise to the receptive field of a simple cell induces a corresponding degree of spatial uncertainty in the outputs triggered by that receptive field. Thus the input fan-in at complex cells which receive inputs from large simple cell receptive fields is larger than the input fan-in at complex cells which receive inputs from small simple cell receptive fields.

We call this hypothesis the *self-similarity hypothesis*. Self-similarity in the present model means that a cell's size and its input and output apparatus scale up or down together. This property clarifies how a *cell type* can be generated in multiple sizes across a neural network by parametrically exciting a single growth-control variable that controls the cells' overall sizes. From an information theoretical viewpoint, the self-similarity property enables a match to be established between the degree of spatial uncertainty of a cell's input and the degree of spatial uncertainty of the cell's output. This concept has been useful in a number of other neural network applications. For example, see Cohen and Grossberg (1986, 1987) for an application of multiple-scale self-similar networks in the self-organization of pattern recognition codes, Daugman (1985) and Watson and Ahumada (1987) for an application in the modelling of simple cell receptive fields, and Grossberg (1969, 1982, p. 590) for other applications of this idea.

The self-similarity hypothesis permits the functional properties depicted in Figures 7 and 8 to be mechanistically realized: Larger simple cell receptive fields (Figure 8a) can fuse together more disparate monocular images (Figure 7a) than can smaller receptive fields (Figures 8b and 7b). Thus the size-disparity correlation is a consequence of three properties of the model: (a) organization of simple cells into scale-selective hypercolumn maps; (b) self-similarity of simple cell input and output fields; and (c) organization of complex cell outputs into scale-selective and orientation-selective on-center off-surround networks (first competitive stage) that are designed to sharply contrast-enhance their input patterns.

9. NEURAL NETWORK MODEL: TRANSFORMING DISTRIBUTED HYPERCOLUMN ACTIVATION PATTERNS INTO MULTIPLEXED ACTIVITY PEAKS

We simulated the responses of complex cells with different receptive field sizes to image edges registered with different binocular disparities. In particular, we analyzed the responses of hypercolumns organized into a 1-dimensional array of alternating left and right ocular dominance slabs. The simple cells within these ocular dominance slabs respond to an image contrast by forming two interleaved activity profiles, one for each eye. We modeled four oriented mask sizes by varying the width of the activity profiles across the hypercolumns, as in Figure 8. We modeled positional disparity by varying the relative positions of the left- and right-eye activity patterns, as in Figures 3b–d.

The network model is summarized pictorially in Figure 4. Signals were transmitted from the simple cells in level F_1 to a level F_2 of binocular complex cells via a set of feedforward excitatory on-center and inhibitory off-surround connections whose strength is a Gaussian function of intercellular distance, with a small random factor added to break the spatial symmetry of the Gaussian. Each complex cell in level F_2 received inputs from both left and right ocular dominance slabs. The sum of the bottom-up inputs activated target complex cells via multiplicative, or shunting, interactions (Grossberg, 1982).

Within level F_2, cells interacted via a distance-dependent on-center off-surround feedback network. Each cell inhibited its neighbors via synapses whose strength decreased as a Gaussian function of intercellular distance. The sum of inhibitory signals received by each complex cell triggered shunting inhibition of the target cell's activity. Each complex cell also generated an autocatalytic excitatory feedback signal. Both the excitatory and inhibitory feedback signals emitted by each cell were generated as a faster-than-linear (in particular, a polynomial) function of the cell's activity, in order to sharpen the pattern of F_2 activation caused by the input pattern received from F_1.

10. LEVEL F_1 ACTIVATION PATTERN: MULTIPLE SCALES AND DISPARITIES

Such a combination of feedforward and feedback network interactions generates the desired fusion and

rivalry behavior within a robust parameter range. Figure 9 schematizes the activation patterns across the simple cells of level F_1 that are defined within four different spatial scales. Each row depicts the activation patterns within a single scale. Successive columns depict the effect of increasing the positional disparity of the left and right eye images of the edge that was registered by the simple cells. These activation patterns were chosen to generate the output patterns from level F_1 to level F_2. They were defined as follows.

Let x_{Si} be the activity at the ith cell population v_{Si} of scale S in level F_1. Let cells with even index i correspond to the input of one eye, and cells with odd index i correspond to the input of the other eye. Let

$$x_{Si} = \begin{cases} X_S\left(\frac{i}{2}\right) & \text{if } i \text{ is even} \\ X_S\left(\frac{i-1-2D}{2}\right) & \text{if } i \text{ is odd} \end{cases} \quad (1)$$

where $-36 \leq i \leq 36$. In (1), parameter D is the disparity of the two images. In Figure 9, $D = 0, 1,$

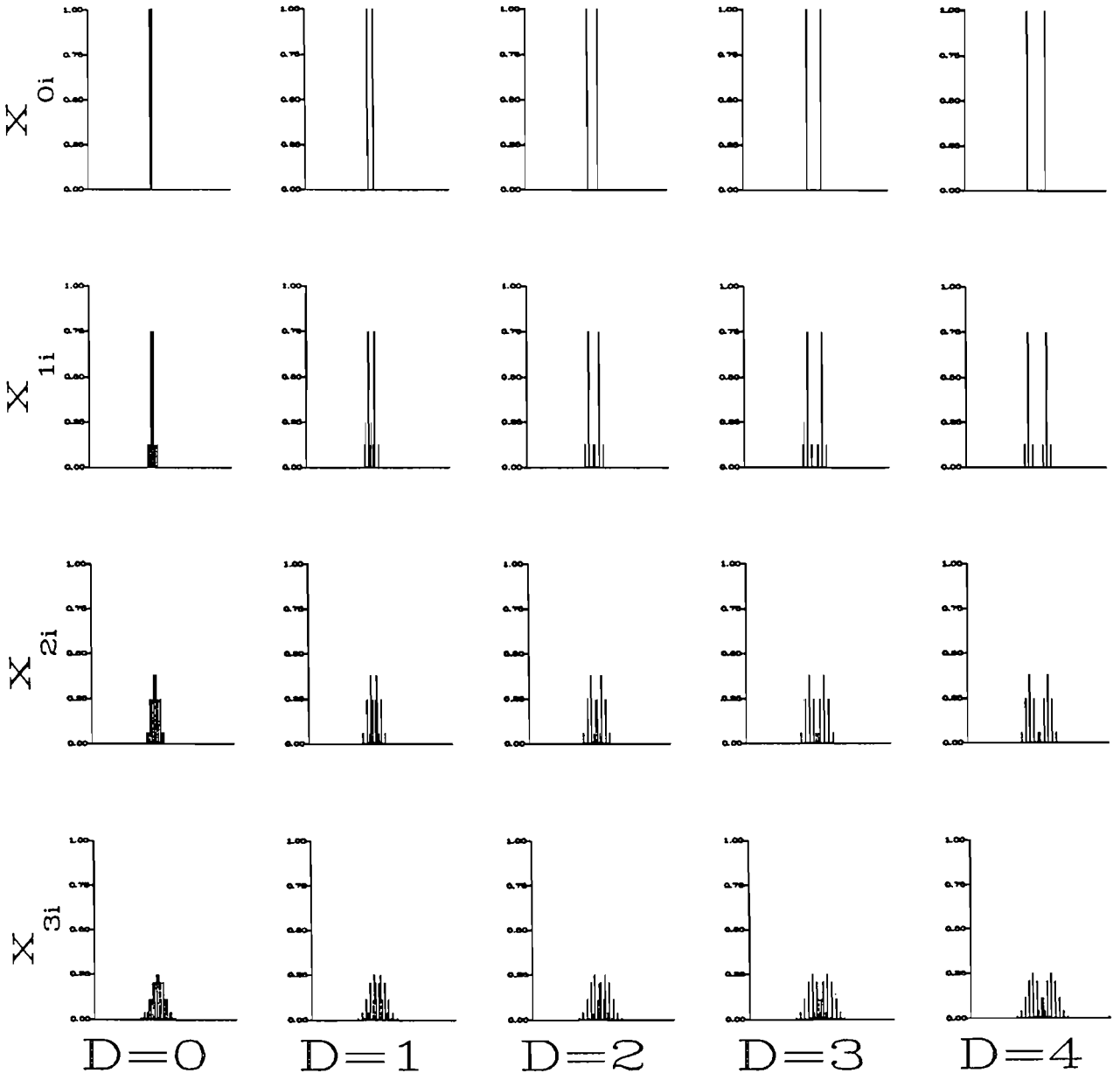

FIGURE 9. Activity patterns used as input for Figures 11–18. Bar heights code each input cell's activity level x_{Si}. Each row indicates the input patterns at a single spatial scale S, with disparity D increasing across the row. Note that the total amount of activity in all cells within each pattern is constant; this approximates the effect of a feedforward on-center off-surround shunting network.

2, 3, 4 in columns 1, 2, 3, 4, 5, respectively. Function $X_S(i)$ is a Gaussian function of distance that is parameterized as follows. Let

$$X_S(i) = \frac{1}{N_S} e^{-\nu_S i^2} \tag{2}$$

for all i such that $e^{-\nu_S i^2} \geq .02$, and $= 0$ otherwise. In (2), the spatial scale $S = 0, 1, 2, 3$; the spatial decay rate

$$\nu_S = \begin{cases} \dfrac{\ln 6}{S^2} & \text{for } S = 1, 2, 3; \\ \infty & \text{for } S = 0 \end{cases} \tag{3}$$

and the normalization factor

$$N_S = \sum_k \{e^{-\nu_S k^2} : e^{-\nu_S k^2} > .02\}. \tag{4}$$

This normalization term approximates the effect of a shunting on-center off-surround network; see subsequent discussion. The algebraic normalization in (2) was used for simplicity.

11. LEVEL F_2 NONLINEAR FEEDBACK NETWORK: BALANCED AND NORMALIZED GAUSSIAN GRADIENTS

Let y_{Sj} be the activity of the jth cell population v_{Sj} of scale S in level F_2. These activities are assumed to form an on-center off-surround network undergoing shunting feedforward and feedback interactions, as defined by the system

$$\frac{d}{dt} y_{Sj} = -\alpha y_{Sj} + (\beta - y_{Sj})(F_{Sj}^{(+)} + B_{Sj}^{(+)}) - (\gamma + y_{Sj})(F_{Sj}^{(-)} + B_{Sj}^{(-)}). \tag{5}$$

In (5), term $F_{Sj}^{(+)}$ is the total feedforward signal from level F_1 to v_{Sj}, $B_{Sj}^{(+)}$ is the total excitatory feedback signal from level F_2 to v_{Sj}, $F_{Sj}^{(-)}$ is the total inhibitory feedforward signal from level F_1 to v_{Sj}, and $B_{Sj}^{(-)}$ is the total inhibitory feedback signal from level F_2 to v_{Sj}.

Extensive computer simulations have demonstrated that both excitatory feedforward signals $F_{Sj}^{(+)}$ and inhibitory feedforward signals $F_{Sj}^{(-)}$ are needed to prevent a variety of anomalies from occurring in the responses of F_2 cells as the spatial scale S and disparity D parameters are varied. Some of these anomalies are reported below. Such a feedforward on-center off-surround anatomy is also needed to solve the fundamental *noise-saturation dilemma* (Grossberg, 1982, 1988c); that is, to enable F_2 to self-tune, and thereby self-normalize, its responses to $F_1 \rightarrow F_2$ signals as parameters such as S and D are varied. Were this not accomplished, different parameter choices could generate widely different activation levels in the responses of F_2 cells to signals from F_1. The feedback signals $B_{Sj}^{(+)}$ and $B_{Sj}^{(-)}$ are, however, chosen to be nonlinear in order to carry out basic operations such as noise suppression and contrast enhancement. These nonlinear feedback signals would respond to such highly variable activation levels in ways that do not generate the design competence being sought.

All the signals are discrete convolutions:

$$F_{Sj}^{(+)} = \sum_i g(x_{Si}) F_{Sij}^{(+)}, \tag{6}$$

$$F_{Sj}^{(-)} = \sum_i g(x_{Si}) F_{Sij}^{(-)}, \tag{7}$$

$$B_{Sj}^{(+)} = \phi \sum_k h(y_{Sk}) B_{Skj}^{(+)}, \tag{8}$$

and

$$B_{Sj}^{(-)} = \psi \sum_k h(y_{Sk}) B_{Skj}^{(-)}. \tag{9}$$

In (6) and (7), the feedforward signal function was chosen to be threshold-linear:

$$g(x_{Si}) = \max(x_{Si}, 0). \tag{10}$$

In (8) and (9), the feedback signal function was chosen to be faster-than-linear-above-threshold:

$$h(y_{Sk}) = [\max(y_{Sk} - \delta, 0)]^4. \tag{11}$$

All the spatial connections defined in (6)–(9) were chosen to vary as a Gaussian function of distance. In addition, it was assumed that *synaptic conservation* obtains; namely, that the total connection strength to each cell within each scale is the same (Cohen & Grossberg, 1986, 1987). It was also assumed that the feedforward and feedback excitatory coefficients $F_{Sij}^{(+)}$ and $B_{Skj}^{(+)}$ are influenced by small random fluctuations of their otherwise Gaussian spatial distribution. These random terms enable input patterns of different disparity D but the same scale S to activate different cells v_{Sj} even if all such input patterns are symmetrically distributed in space with respect to one another. In a multiple scale network such as a masking field (Cohen & Grossberg, 1986, 1987), the change in spatial scale of the total activation pattern across level F_1 as D increases can itself cause selection of different cells in level F_2.

These qualitative constraints are quantitatively realized as follows. Let

$$F_{Sij}^{(-)} = \begin{cases} \dfrac{\mu}{N_{Sj}} \exp[-f_S^{(-)}(i-j)^2] \\ \quad \text{if } \exp[-f_S^{(-)}(i-j)^2] > 5 \times 10^{-5} \\ 0 \quad \text{otherwise} \end{cases} \tag{12}$$

and

$$B_{Sij}^{(-)} = F_{Sij}^{(-)}, \tag{13}$$

where

$$N_{Sj} = \sum_i \{\exp[-f_S^{(-)}(i-j)^2]: \exp[-f_S^{(-)}(i-j)^2] > 5 \times 10^{-5}\}. \quad (14)$$

Term N_{Sj} is independent of j because of the isotropy of the field of integers over which it is computed.

By (13), the feedforward and feedback inhibitory signals share a common set of inhibitory interneurons. This is not assumed about the feedforward and feedback excitatory interneurons. The on-center of the feedforward excitatory interneurons increases in width with the scale S in order to generate the size-disparity correlation (Section 6). In addition, the excitatory on-center needs to include more than one cell population in order to be able to binocularly fuse disparities D that are greater than zero. On the other hand, the outcome of the feedback competition within F_2 must be able to choose a narrow focus of cell activity across all fusable combinations of S and D. Thus the spatial distribution of excitatory feedback signals $B_{Skj}^{(+)}$ was chosen more narrowly than that of the excitatory feedforward coefficient $F_{Sij}^{(+)}$. The latter coefficients were chosen as follows.

First, a Gaussian gradient of coefficients was defined and was perturbed randomly by a small amount. This it was normalized to attain the same total connection strength to the cell as did the unrandomized gradient. Thus let

$$f_{Sij}^{(+)} = \begin{cases} P_{ij} \exp[-f_S^{(+)}(i-j)^2] \\ \qquad \text{if } \exp[-f_S^{(+)}(i-j)^2] > 5 \times 10^{-5} \\ 0 \qquad \text{otherwise} \end{cases} \quad (15)$$

where

$$P_{ij} = 1 + .01(2R_{ij} - 1) \quad (16)$$

and R_{ij} is chosen pseudo-randomly within the interval [0, 1). The coefficients $f_{Sij}^{(+)}$ thus form a slightly perturbed Gaussian distribution. In order to normalize these connection strengths, we need to first define the excitatory feedback coefficients. Let

$$b_{Skj}^{(+)} = \begin{cases} P_{kj} \exp[-b_S^{(+)}(k-j)^2] \\ \qquad \text{if } \exp[-b_S^{(+)}(k-j)^2] > 5 \times 10^{-5} \\ 0 \qquad \text{otherwise} \end{cases} \quad (17)$$

where P_{kj} is defined as in (16). Now define

$$F_{Sij}^{(-)} = \frac{\nu f_{Sij}^{(+)}}{\sum_l f_{Slj}^{(+)} + \sum_m b_{Smj}^{(+)}} \quad (18)$$

and

$$B_{Skj}^{(+)} = \frac{\nu b_{Skj}^{(+)}}{\sum_l f_{Slj}^{(+)} + \sum_m b_{Smj}^{(+)}}. \quad (19)$$

In some of our computer simulations, we chose $b_S^{(+)} = \infty$ in (17) to illustrate how a sharply fused F_2 response could be elicited across all scales S. We also studied the case in which $b_S^{(+)} = f_S^{(+)}$ in (15) and (17) so as to illustrate what would happen to fusion if the feedforward and feedback excitatory signals shared a common set of interneurons.

12. LAMINAR ORGANIZATION OF NETWORK INTERACTIONS

Connections such as those defined in (6)–(19) can naturally be generated by a laminar organization of model cortical columns (Rakic, 1988). As illustrated in Figure 10, the feedforward signals from F_1 to F_2 are registered in two distinct layers of a model column (Figure 10a). One layer contains excitatory feedforward interneurons, the other inhibitory feedforward interneurons. Both sets of interneurons send pathways to the target cells v_{Sj} (Figures 10b and 10c) which are segregated in their own layer. These interneurons give rise to the terms $F_{Sij}^{(+)}$ and $F_{Sij}^{(-)}$, respectively. The target cells v_{Sj} send topographically organized positive feedback signals to the layer of inhibitory interneurons (Figure 10d). Thus these inhibitory interneurons are accessed both by feedforward and feedback pathways, as in (13). The cells v_{Sj} also project to a layer containing a separate population of excitatory interneurons (Figure 10d), which give rise to the excitatory feedback signals $B_{Skj}^{(+)}$ (Figure 10e). It is assumed that all interneurons respond rapidly and approximately linearly to their inputs, at least relative to the slower reaction rates and more pronounced shunting effects that take place within the populations v_{Sj}.

13. NETWORK RESPONSE TO FEEDFORWARD SIGNALS: COEXISTENCE OF MULTIPLE SCALES AND STM NORMALIZATION

In order to understand the network feedback dynamics, it is instructive to first shut off the feedback signals in (15)—that is, set $B_{Sj}^{(+)} \equiv B_{Sj}^{(-)} \equiv 0$—and to study the network's responses to the feedforward signals $F_{Sj}^{(+)}$ and $F_{Sj}^{(-)}$ as the scale S and disparity D are varied.

Figure 11 depicts the equilibrium activities

$$y_{Sj} = \frac{\beta F_{Sj}^{(+)} - \gamma F_{Sj}^{(-)}}{\alpha + F_{Sj}^{(+)} + F_{Sj}^{(-)}} \quad (20)$$

of F_2 cells in response to the feedforward input signals from F_1 that are derived from the activation patterns depicted in Figure 9. The solid curves represent the y_{Sj} as functions of j. Each curve corresponds to a different choice of S (across rows) and D (across columns). The dotted curves are proportional to the $F_{Sj}^{(+)}$, and the dashed curves are proportional to the

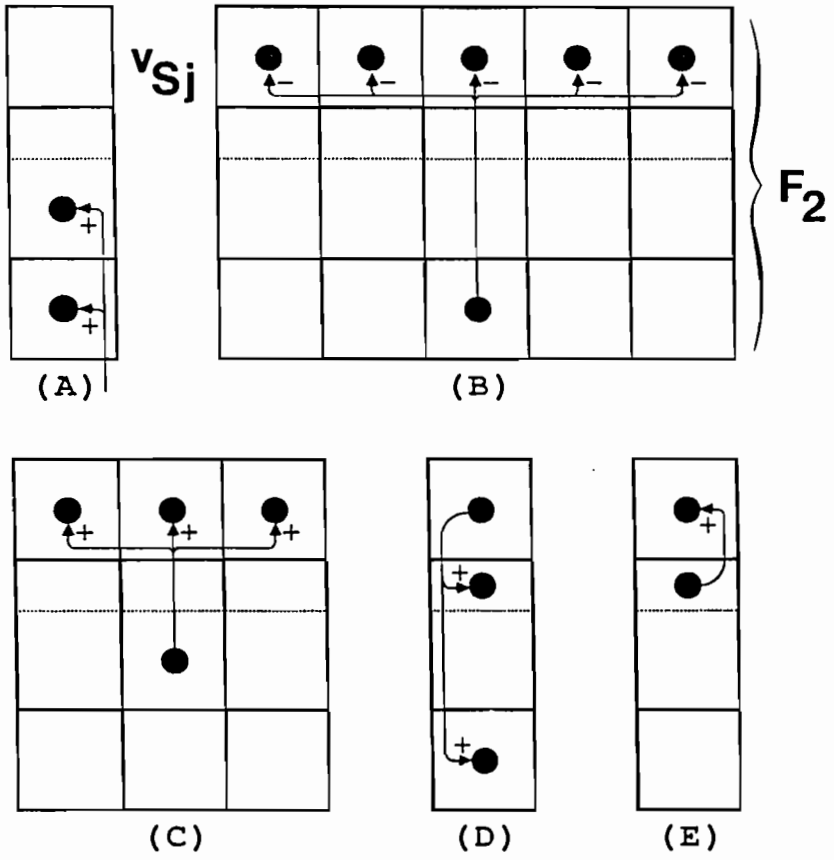

FIGURE 10. Laminar organization of level F_2. (a) Input connections map topographically from F_1 and ramify in two laminae of F_2. (b) Cells in the inhibitory lamina project feedforward connections to a broad range of nearby F_2 target cells v_{Sj}, forming a broad off-surround. (c) Cells in the excitatory lamina project feedforward connections to a range of nearby F_2 target cells v_{Sj}, forming a narrow on-center. Bandwidths of the distance-dependent connection strengths in (b) and (c) vary according to the scale parameter S. (d) Feedback connections from F_2 cells project back both to the inhibitory lamina and to a separate excitatory lamina. (e) The separate excitatory lamina projects reciprocal connections in a topographically sharp manner back to F_2. Thus, there exist two feedback loops in the network: between F_2 cells and inhibitory interneurons, and between F_2 cells and interneurons in the separate excitatory lamina. The inhibitory interneurons serve both the feedforward and feedback pathways, but the excitatory interneurons which receive input from F_1 are feedforward-only. Section 16 describes the results of applying the input to a variant of this architecture, in which the feedforward and feedback excitatory laminae are merged into a single lamina which receives both feedforward and feedback input and which projects back to F_2 as in (c).

$F_{Sj}^{(-)}$. Within each row, the scales of the $F_{Sj}^{(+)}$ and $F_{Sj}^{(-)}$ functions were chosen to fit their graphs on the y_{Sj} graphs. The parameters are provided in Table 1. Three main trends are evident in Figure 11 as S is increased in successive rows:

1. The smallest disparity $D = D_S$ at which a graph y_{Sj} becomes bimodal increases with S.
2. The graphs y_{Sj} become broader as S increases.
3. The maximal values of all graphs y_{Sj} are approximately equal.

14. NETWORK RESPONSE TO FEEDFORWARD AND FEEDBACK SIGNALS: COEXISTENCE OF SIZE-DISPARITY CORRELATION AND MAXIMAL STM COMPRESSION

Figure 12 depicts the equilibrium activities y_{Sj} in response to the same input patterns when the full system (5) of feedforward and feedback signals is operative. The solid curves represent the functions y_{Sj}. The dotted curves are proportional to the functions $F_{Sj}^{(+)} + B_{Sj}^{(+)}$, and the dashed curves are proportional to the functions $F_{Sj}^{(-)} + B_{Sj}^{(-)}$. Three main trends are evident in Figure 12:

1. The size-disparity correlation obtains. Note that fusion persists in scale $S = 0$ (row 1) until disparity $D = 1$, in scale $S = 1$ (row 2) until disparity $D = 2$, in scale $S = 2$ (row 3) until disparity $D = 3$, and in scale $S = 3$ (row 4) until disparity $D = 4$.
2. Although the curves y_{Sj} in Figure 11 become broader as S increases, the curves y_{Sj} in Figure 12 remain maximally compressed no matter how S is chosen. Thus the network as a whole resolves the design tension between increasing the spatial width of $F_1 \rightarrow F_2$ signals as S increases so as to generate the basis for the size-disparity correla-

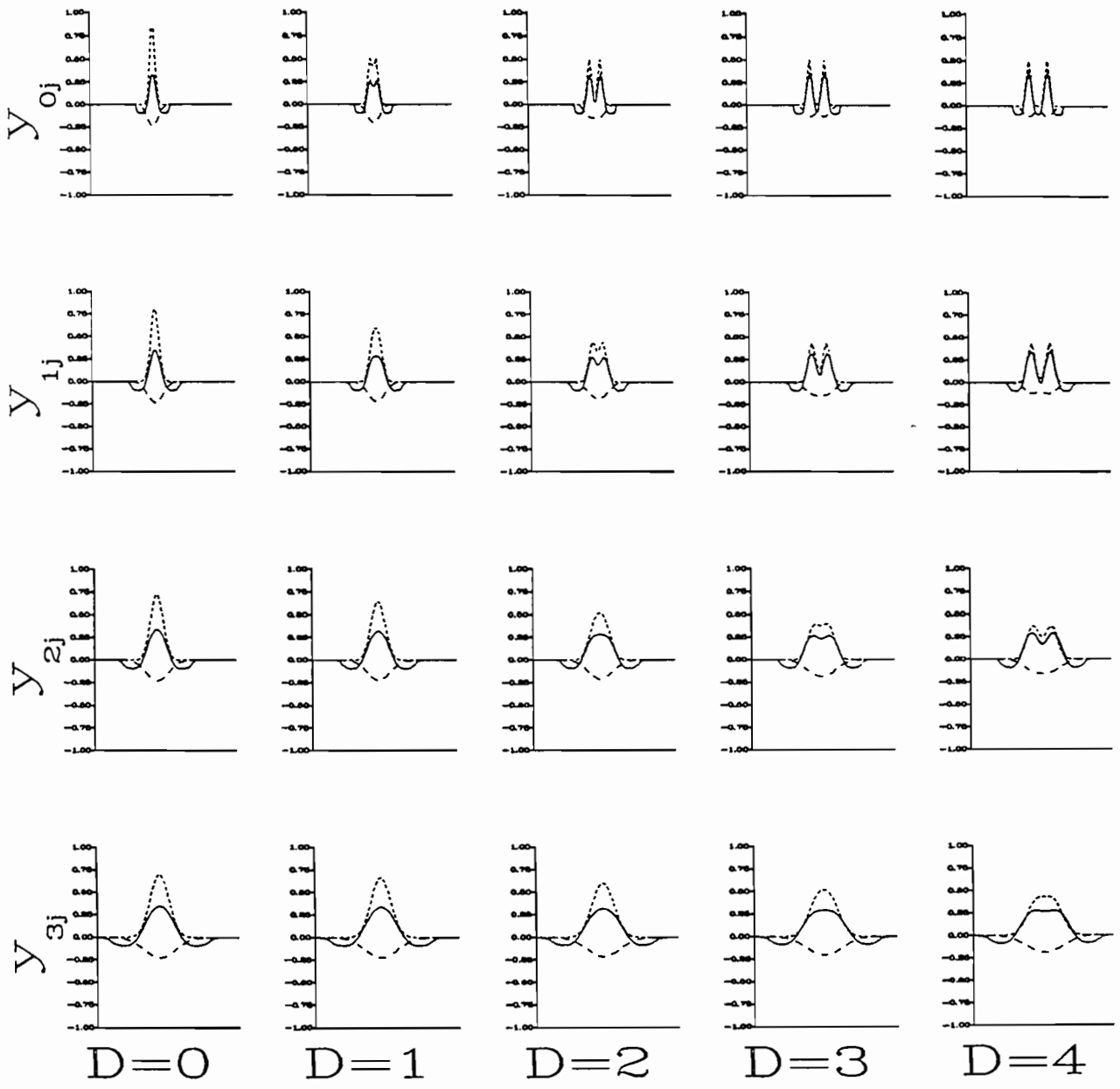

FIGURE 11. Results of applying input patterns of Figure 9 to a purely feedforward architecture (in which the feedback pathways of Figure 10d are eliminated). The abscissa of each graph represents spatial position across F_2. Solid curves, plotted as a function of cell position j, represent the activity levels y_{Sj} of all the cells across F_2. Dotted and dashed curves represent the relative shapes of the total excitatory input and the total inhibitory input to each cell across F_2, respectively. Each column of the figure represents the response of F_2 to input patterns at a single disparity $D \in \{0, 1, 2, 3, 4\}$, wherein each row represents the response of the F_2 subnetwork at a single spatial scale $S \in \{0, 1, 2, 3\}$ to input patterns of the corresponding spatial scale.

tion, yet generating highly compressed STM response patterns that subserve receptive fields that are sharply tuned to disparity at all spatial scales.

3. The maximal values of all graphs y_{Sj} are approximately equal.

Consider, in particular, the curves y_{1j} in row 2 of Figure 12. In the first column, the activity of only a single population of complex cells survives the feedback competition. Other complex cells are either at rest, such as those too far from the input activity to receive any bottom-up excitation, or are hyperpolarized due to inhibition from the single active cell population. Image features that are coded by an activity pattern distributed across many cells in F_1, and by a broadly distributed feedforward response of F_2 in Figure 11, are coded in a compact, multiplexed fashion by the activity of a single F_2 cell population.

In successive columns, disparity is increased by pulling apart the left and right input patterns. These input patterns remain coded by a single-cell peak of activity in F_2 until disparity $D = 3$ is reached. At disparity $D = 3$, the complex cells which receive the most input are so far apart that the combination of

TABLE 1
Parameters Used In Simulations

All the simulations are run to $t = 40$, by which time equilibrium has been reached. The following parameters are common to all the simulations:

$$\alpha = 0.1, \quad \beta = 1, \quad \gamma = 0.1,$$

$$\delta = 0.06, \quad \nu = 100, \quad \mu = 893.$$

The matrix below describes the parameters that varied across simulations.

Figure	ϕ	ψ	$f_S^{(+)}$	$b_S^{(+)}$	$f_S^{(-)}$
11	0	0	$2^{12-S} \times 10^{-4}$	—	$2^{13-S} \times 10^{-5}$
12	156	60	$2^{12-S} \times 10^{-4}$	∞	$2^{13-S} \times 10^{-5}$
13	0	0	$2^{11} \times 10^{-4}$	—	$2^{12} \times 10^{-5}$
14	156	60	$2^{11} \times 10^{-4}$	∞	$2^{12} \times 10^{-5}$
15	156	60	$2^{12-S} \times 10^{-4}$	$2^{12-S} \times 10^{-4}$	$2^{13-S} \times 10^{-5}$
16	780000	300000	$2^{12-S} \times 10^{-4}$	∞	$2^{13-S} \times 10^{-5}$
17[a]	156	60	$2^{12-S} \times 10^{-4}$	∞	$2^{13-S} \times 10^{-5}$
18[a]	156	60	$2^{11} \times 10^{-4}$	∞	$2^{12} \times 10^{-5}$

[a] Define $B_{Sij}^{(-)}$ as in equations (12) and (13), but then let all $F_{Sij}^{(-)} = 0$.

broad lateral inhibition and narrowly focused positive feedback causes the F_2 activity pattern to split into two sharp peaks, each of which represents the input pattern at a single eye. The outputs of the two peaks can then compete subsequently within the BCS (Figure 7b), thereby giving rise to a rivalrous percept (Grossberg, 1987b).

Each column in Figure 12 shows the F_2 response of several scales at a fixed disparity. In column 4, for example, the smaller scales, $S = 0$ and $S = 1$, show a double peak response to the image input features, whereas the larger scales, $S = 2$ and $S = 3$, respond with only a single fused peak of activity. Such a comparison illustrates how fusion and rivalry can coexist simultaneously at different scales within the visual system. When the disparity between the left and right input patterns responding to an image feature is sufficiently small for a particular scale, a narrow peak of fused activity can exclusively code the image feature at that scale. But at other scales, the same disparity may cause a double peak of activity to emerge, whence the two peaks can then compete rivalrously within the BCS.

These results illustrate how compressed multiplexed representations of abstract multiple-scale information can emerge from the nonlinear interactions of a neural network. The architectural details of these networks, being quite general, promise to clarify how such representations arise in a variety of neural systems. In particular, the properties of self-similar scaling that propagates through the network's several levels, and of balanced excitatory and inhibitory signals at each level, are robust design constraints for multiple-scale networks.

In the subsequent sections, we demonstrate that the aforementioned properties are not all obtained if any of the system's major interactions is eliminated. In particular, Section 15 illustrates a breakdown if the $F_1 \rightarrow F_2$ interactions do not preserve the multiscale character of the F_1 simple cell receptive fields. Section 16 illustrates a breakdown if the $F_2 \rightarrow F_2$ excitatory feedback signals are not more narrowly distributed than the $F_1 \rightarrow F_2$ excitatory feedforward signals. Section 17 notes a breakdown that occurs if there are imbalances between the relative strengths or distributions of the feedforward and feedback inputs to each cell. Section 18 illustrates a breakdown if the $F_1 \rightarrow F_2$ excitatory feedforward signals are not balanced by inhibitory feedforward signals. Taken together, these examples show that the design constraints embodied in system (5) all contribute to generating the size-disparity correlation within a network of sharply tuned complex cells.

15. BREAKDOWN OF SIZE-DISPARITY CORRELATION WHEN A SINGLE $F_1 \rightarrow F_2$ SPATIAL SCALE IS USED

Figures 13 and 14 illustrate that the size-disparity correlation is lost when the feedforward Gaussian connections from level F_1 to level F_2 are chosen to be independent of the scale S (see Table 1). Figure 13 depicts the equilibrium responses of level F_2 to feedforward signals, as in (20), and Figure 14 depicts the equilibrium responses of the full system (5).

Note that all scales S generate bimodal responses at the same disparity $D = 3$ in Figure 14. Figure 13 shows that this breakdown is not due entirely to a lack of differentiation among the scales S in the feedforward responses y_{Sj} in (20). In particular, the curves y_{0j} and y_{1j} in Figure 13 tend to become bimodal at disparity $D = 2$, whereas the curves y_{2j} and y_{3j} tend to become bimodal as disparity $D = 3$. The breakdown of this scale-specific tendency in Figure 14 is due to the fact that the scale of the inhibitory feed-

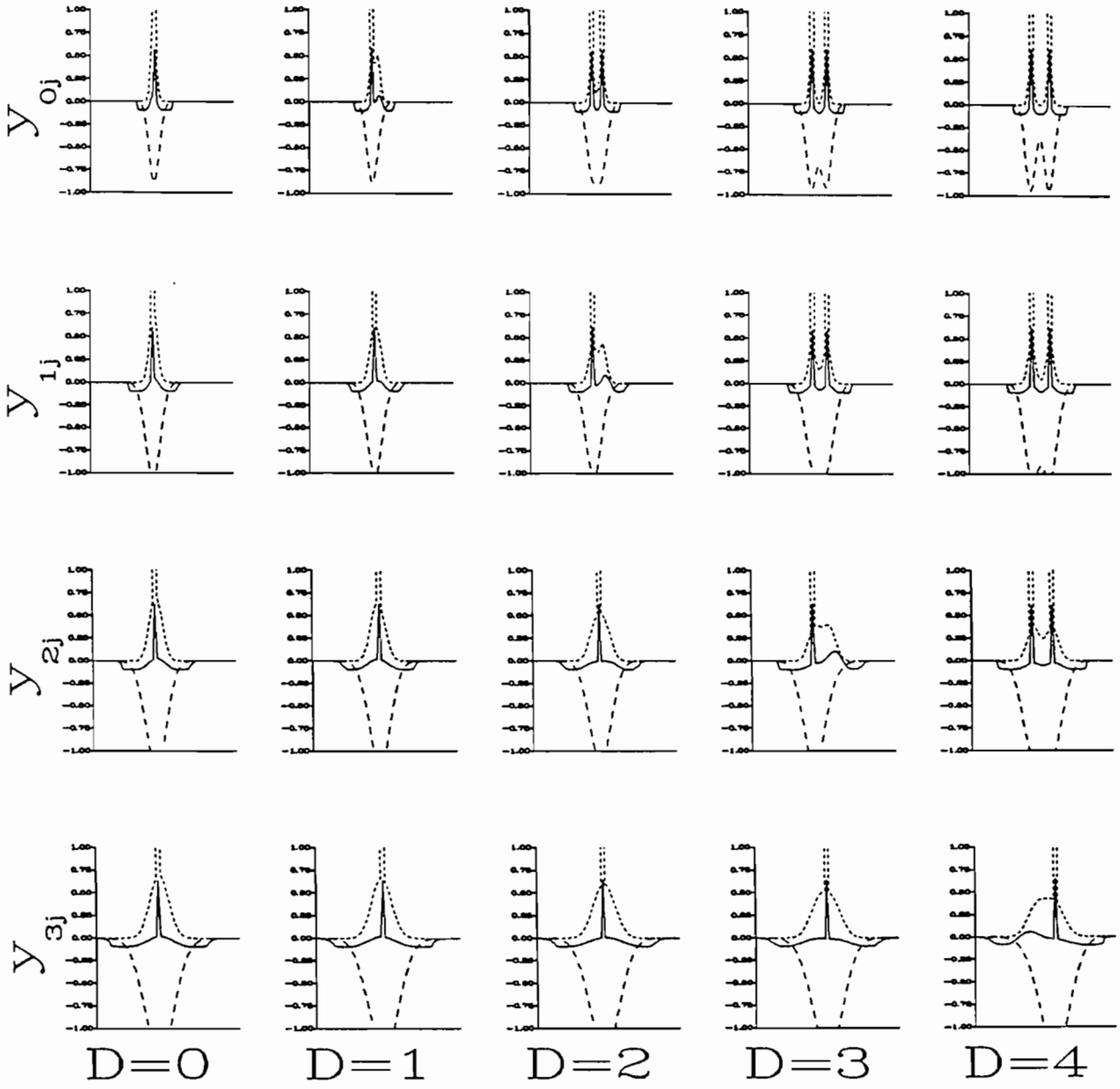

FIGURE 12. Response of the full (feedforward and feedback) network to the input patterns of Figure 9. The solid curves again represent the activity levels y_{Sj} of cells across F_2. Here the activation patterns exhibit much sharper peaks than those of Figure 11, due to the nonlinear feedback interactions illustrated schematically by Figure 10d and 10e. The organization of these peaks as a function of disparity (D) and scale (S) exhibits the size-disparity correlation.

back signals $B_{Sj}^{(-)}$ does not vary with S. The single scale of the inhibitory feedback signals prevents the feedforward activations from generating a scale-specific size-disparity correlation.

16. BREAKDOWN OF RECEPTIVE FIELDS SHARPLY TUNED TO DISPARITY WHEN A SINGLE SCALE OF EXCITATORY INTERNEURONS IS USED

Figure 15 shows that the sharply tuned responses of y_{Sj} in Figure 12 are lost when the excitatory feedback signals are not more narrowly distributed than the excitatory feedforward signals. In Figure 15, both the excitatory feedforward signals and the excitatory feedback signals use the same set of excitatory interneurons.

17. DISINHIBITORY STANDING WAVES WHEN FEEDFORWARD AND FEEDBACK SIGNALS ARE UNBALANCED

If the F_2 cells receive much more feedback input than feedforward input, then their responses become

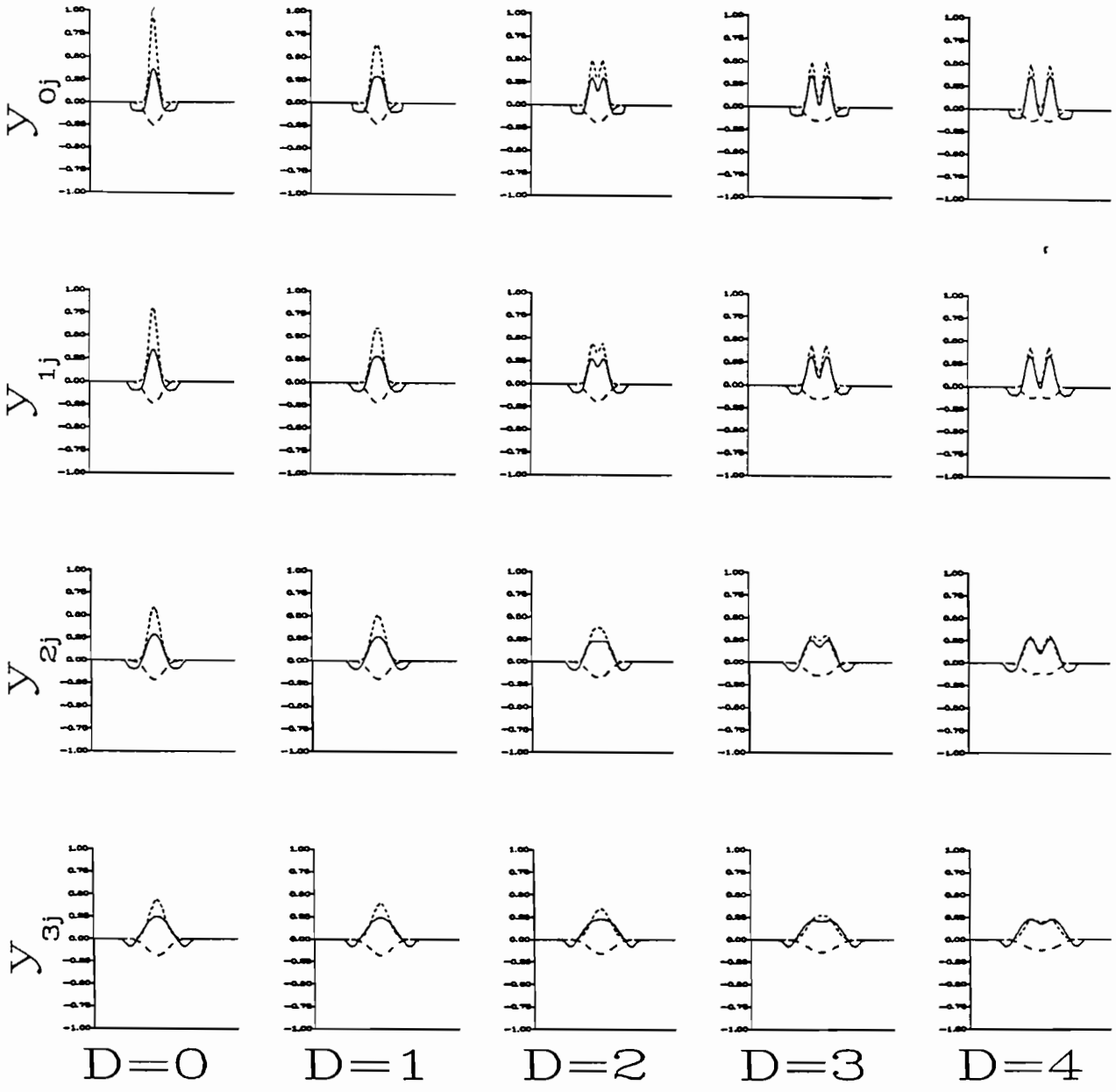

FIGURE 13. Response of a feedforward network similar to that of Figure 11, except that the spatial bandwidths of the output projections of the interneurons in Figure 10b and 10c is constant across all rows, instead of varying with the scale parameter S. Row $S = 1$ here is identical to row $S = 1$ of Figure 11.

dominated by spurious peaks of activation and are only marginally related to the F_1 input signals. This behavior is illustrated by Figure 16, in which full activation peaks appear far from the central region of $F_1 \rightarrow F_2$ excitation. The autocatalytic feedback pathway (Figures 10d–e) magnifies small perturbations in cell activity due to disinhibition or noise and overwhelms the signals elicited by visual inputs.

Pathology can also occur if the inhibitory feedback coefficients are not broadly enough distributed across the network relative to the feedforward excitatory coefficients. Then the equilibrium STM pattern can display a standing wave of multiple excitatory peaks due to propagation of disinhibition by the excitatory feedback signals. This anomaly is prevented by the constraint that the inhibitory feedback signals use the same interneurons as the inhibitory feedforward signals, as in (13), just so long as the inhibitory coefficients are sufficiently large relative to the excitatory coefficients. The same problem can occur by choosing the size of the inhibitory feedback signals too weak relative to the size of the excitatory feedback signals, even if the inhibitory feedforward and feedback signals have the same spatial distribution.

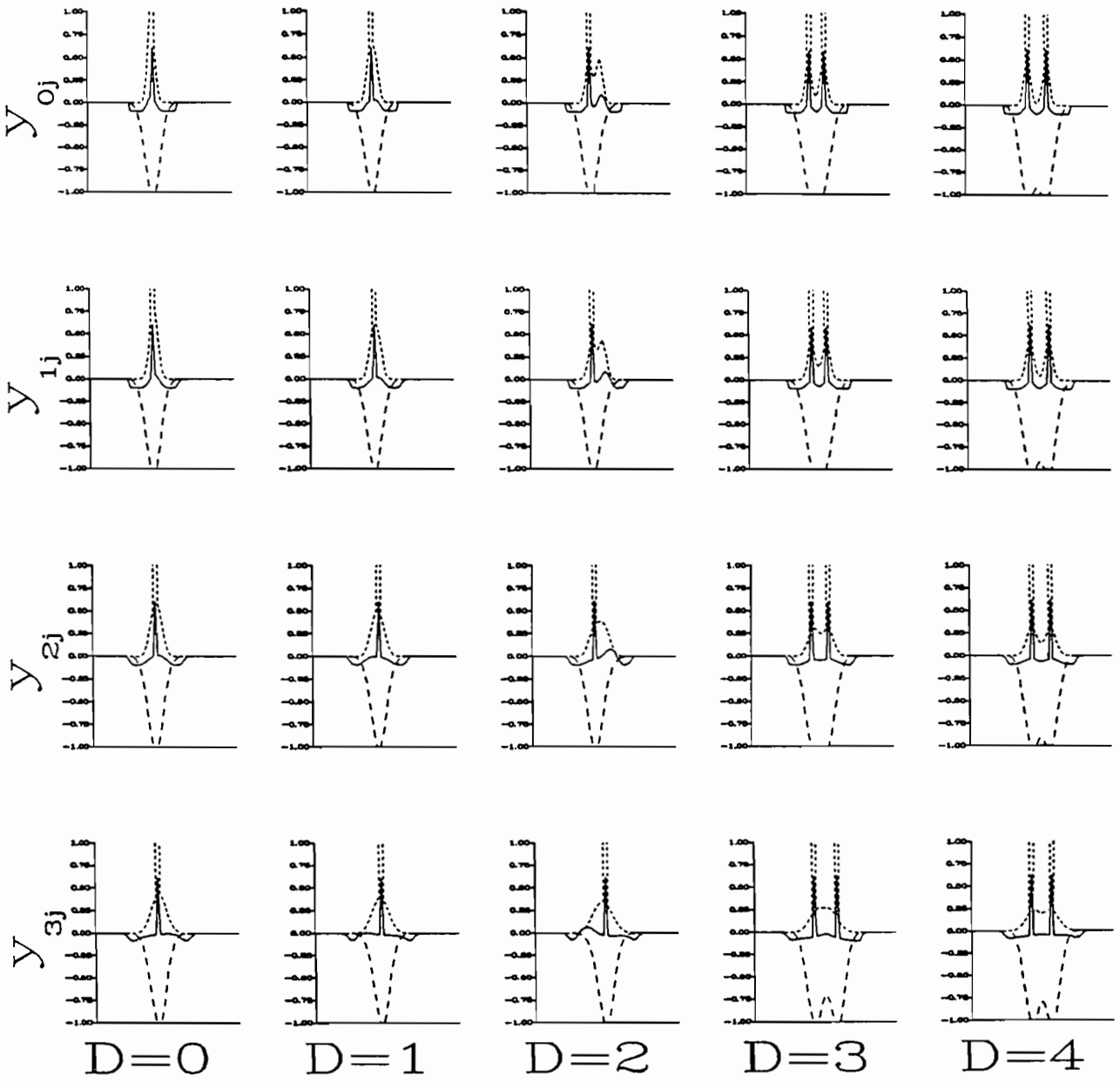

FIGURE 14. Response of network when single-scale feedback interactions are added to the single-scale feedforward interactions of Figure 13.

18. BREAKDOWN OF SIZE-DISPARITY CORRELATION WHEN $F_1 \rightarrow F_2$ EXCITATORY FEEDFORWARD SIGNALS ARE NOT BALANCED BY INHIBITORY FEEDFORWARD SIGNALS

Figure 17 shows that removal of the inhibitory feedforward signal $F_{Sj}^{(-)}$ that balances the excitatory feedforward signal $F_{Sj}^{(+)}$ in (5) can destroy the network's self-normalization property, and thereby undermine the size-disparity correlation. In Figure 18, $F_{Sj}^{(-)}$ is again equal to zero and, in addition, only a single scale is used. A more serious breakdown into an array of disinhibitory peaks is caused.

19. COMPARISON WITH THE CEPSTRUM MODEL OF DISPARITY DETECTION

Yeshurun and Schwartz (1987) have employed the cepstrum operator (Bogert, Healy, & Tukey, 1963) in an ingenious fashion to generate a neurally based model of binocular disparity detection. As in Figure 2, the inputs to their model are activation patterns distributed across the spatial map of ocular dominance columns, which parse left eye (L) and right eye (R) input patterns into thin contiguous strips. The cepstrum operator is the power spectrum of the logarithm of the power spectrum. In their application, the operator is applied to an image registered

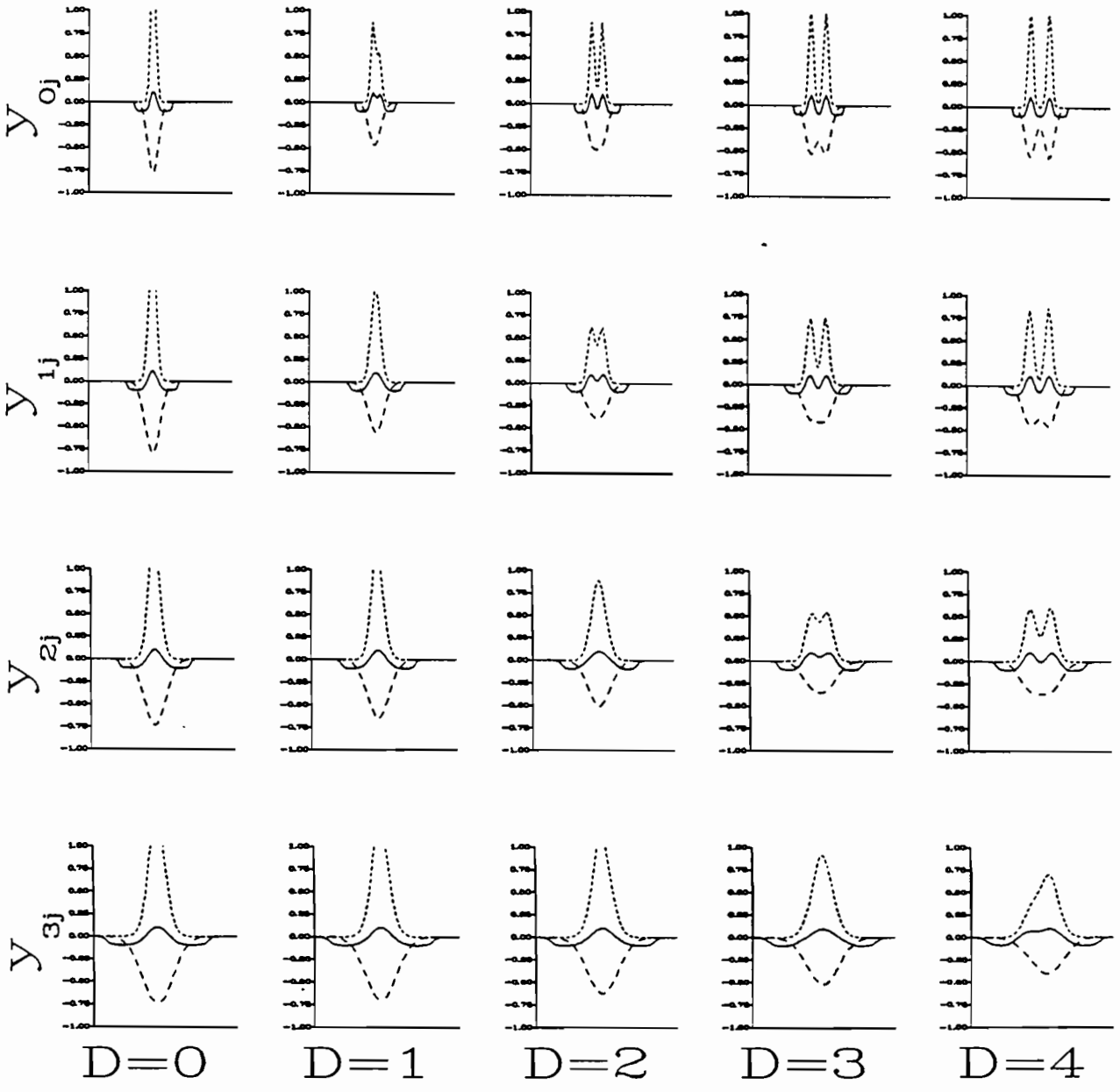

FIGURE 15. Response of network when both excitatory sublaminae are collapsed into a single excitatory lamina. See text and Figure 10 caption for more details. Even though excitatory feedback interactions are present, the degree of sharpness of the y_{sj} output curves (solid curves) is limited by the fan-out of the output projections of the excitatory interneurons (Figure 10c).

by both eyes that is distributed across an (L, R) pair of ocular dominance columns. Yeshurun and Schwartz (1987) demonstrated that the power spectrum of the operator is then sharply concentrated at a spatial location in (L, R) coordinates. This location changes as the disparity between the left eye and right eye images is varied.

The present model shares some common features with the Yeshurun and Schwartz (1987) model. Its computations are also based upon the ocular dominance column map as a data structure, and its F_2 outputs code disparity in terms of the changing locations of a maximally activated cell population. Thus the present model provides a real-time neural network realization of the type of encoding that the cepstrum operator generates through formal algebraic operations.

20. THE MULTIPLEXED FACADE REPRESENTATION

The present model also enjoys conceptual advantages over the cepstrum model. For one, the responses at the level F_2 do not represent only binocular

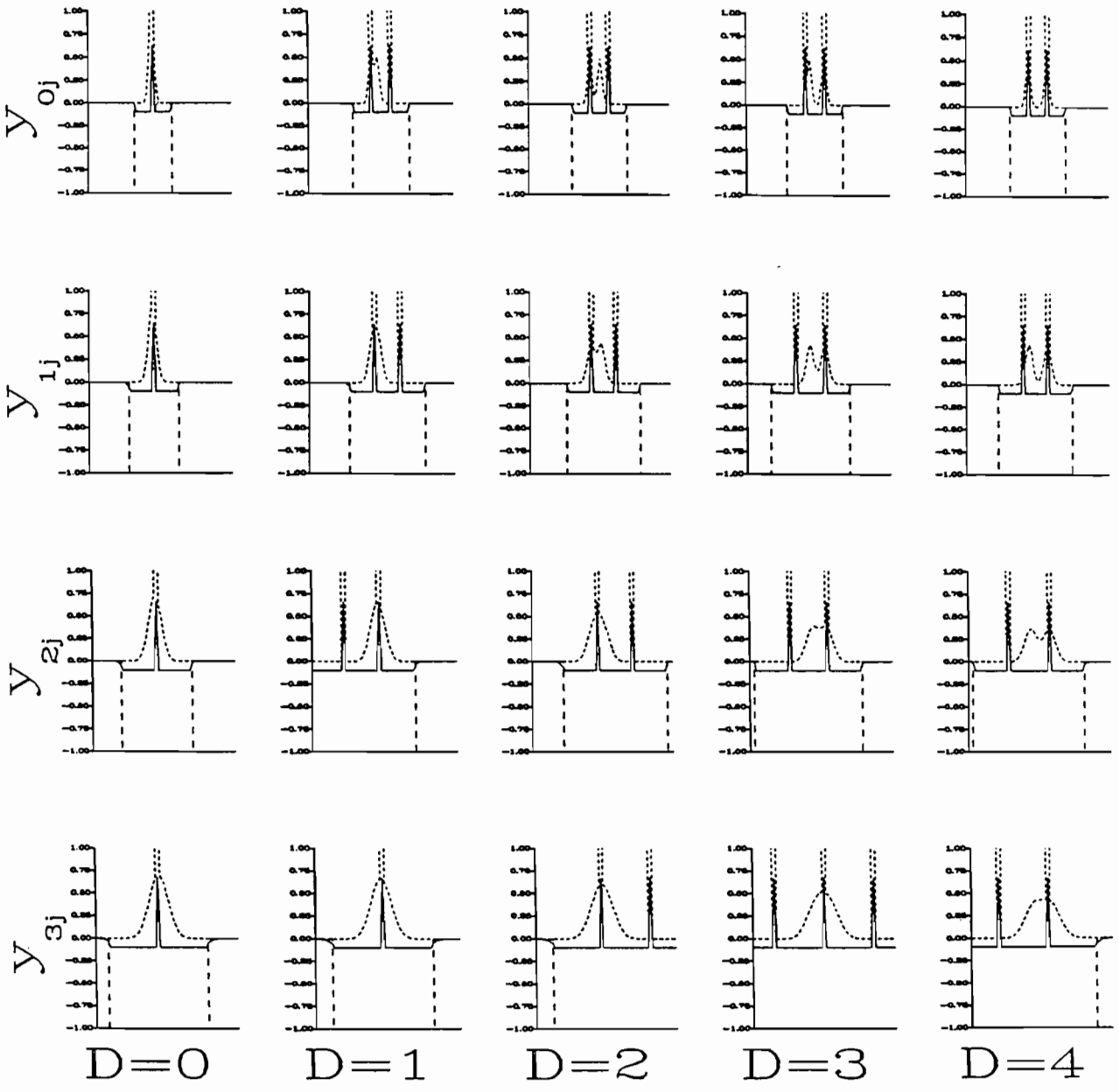

FIGURE 16. Breakdown of network response when feedback signals are too strong relative to feedforward signals. The feedback overwhelms the response of cells across F_2 near the periphery of the F_1 input, leading to the formation of additional activity peaks at increasingly distal positions.

positional disparity. The cell responses also multiplex such properties as retinal position and spatial frequency, and may readily be generalized (as indicated in Figures 2, 4, 5, 7, and 8) to encode such additional properties as orientation, orientational disparity, and insensitivity to direction-of-contrast. In addition, the cells in level F_2 are part of a larger neural network model, called the Boundary Contour System (Grossberg, 1987a, 1987b; Grossberg & Mingolla, 1985a, 1985b). The properties of the BCS have been used to expain how level F_2 responses are used to generate multiple-scale 3-D boundary segmentations capable of simultaneous binocular fusion and rivalry at a single position of perceptual space (Grossberg, 1987b). Such BCS representations, in turn, interact with a Feature Contour System (FCS) wherein they organize a multiplexed representation of surface form and color in a manner that compensates for variable illumination conditions (Grossberg, 1987b, 1988d; Grossberg & Todorović, 1988).

Thus the present results gain additional significance by forming part of a larger neural network whose BCS-FCS interactions suggest how monocular and binocular images may be transformed into a multiplexed representation of Form-And-Color-And-DEpth. Such a multiplexed FACADE representa-

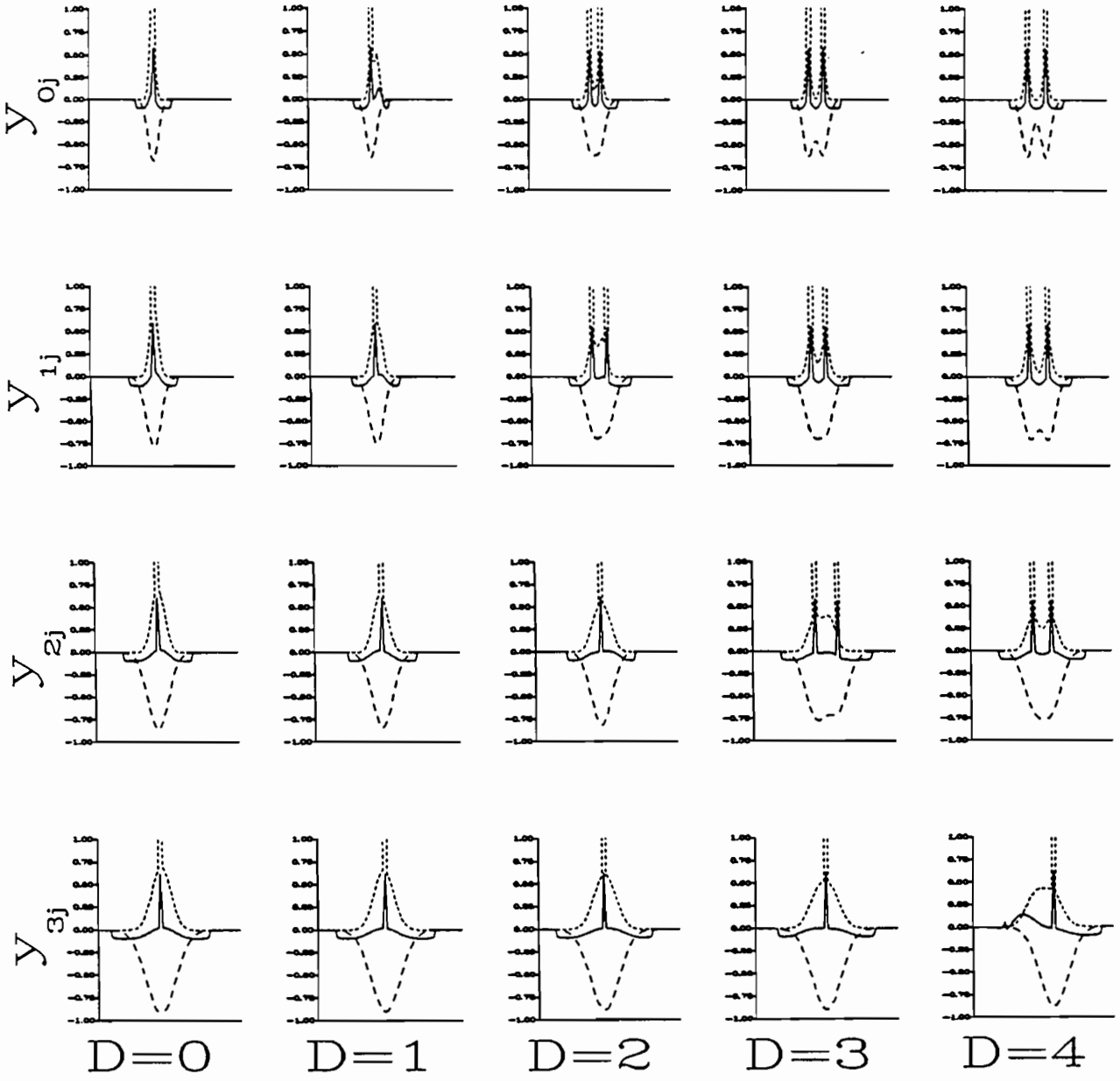

FIGURE 17. Feedforward pathway from F_1 to inhibitory interneurons (Figure 10a) is eliminated. Despite the absence of feedforward inhibition, a diagonal trend of the disparity $D = D_S$ at which the curves y_{Sj} become bimodal is present, but the size-disparity correlation is imperfect. In this case D_0, D_1, D_2, D_3 are equal to 2, 2, 3, 5, respectively. The imperfect diagonal trend indicates that the response of the network is partially sensitive to both disparity and spatial scale.

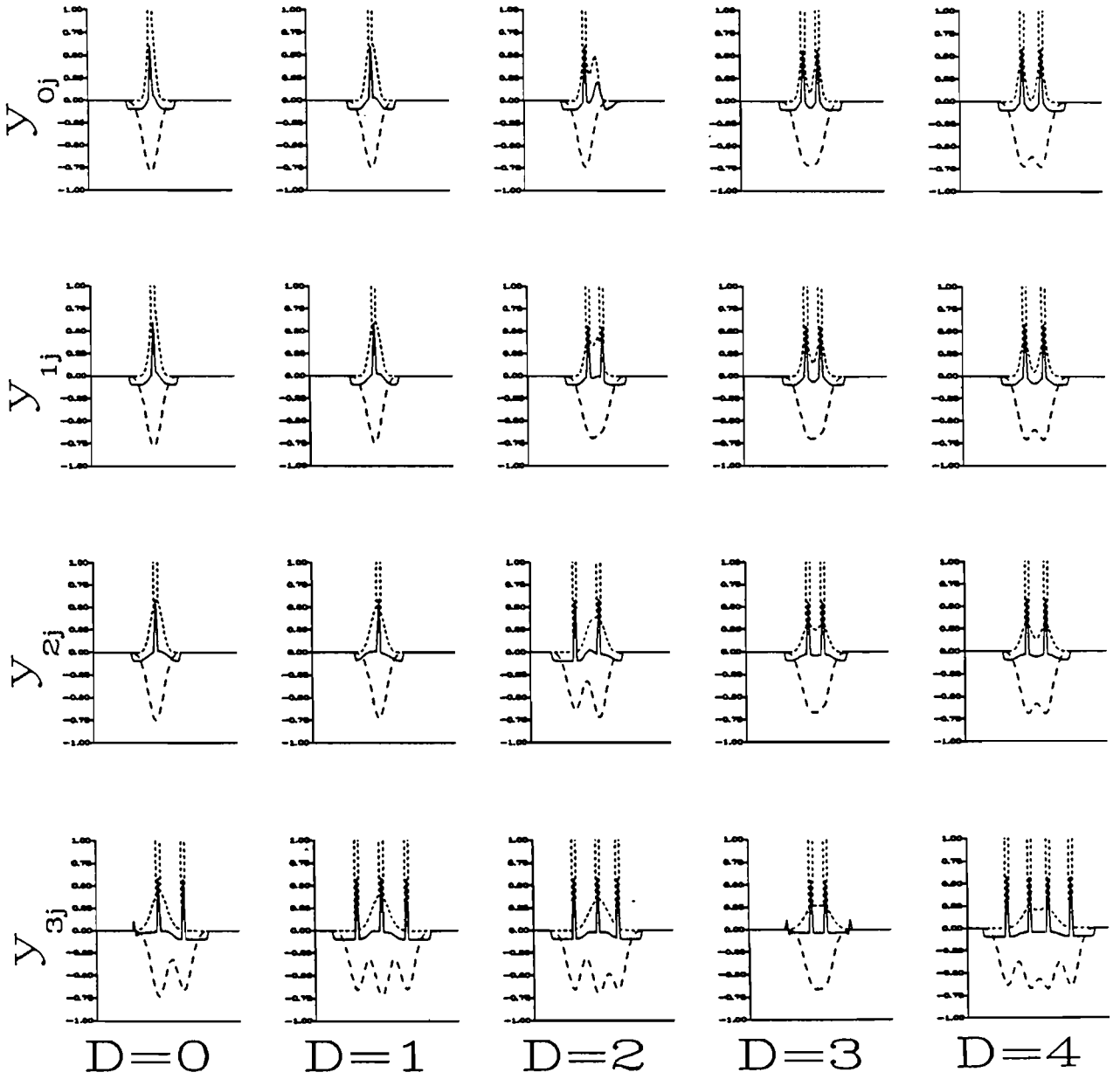

FIGURE 18. As in Figure 17, feedforward inhibition is eliminated. In addition, only the single scale $S = 1$ is used throughout. All other parameters are unchanged. A series of spurious disinhibitory peaks is caused by the resultant imbalances between single-scale feedforward and feedback signals.

tion may prove to be equally useful for the analysis of biological vision and the design of a general-purpose preattentive vision machine for technological applications.

REFERENCES

Amari, S., & Takeuchi, A. (1978). Mathematical theory on formation of category detecting nerve cells. *Biological Cybernetics*, **29**, 127–136.

Bienenstock, E. L., Cooper, L. N., & Munro, P. W. (1982). Theory for the development of neuron selectivity: Orientation specificity and binocular interaction in visual cortex. *Journal of Neuroscience*, **2**, 32–48.

Bogert, B. P., Healy, W. J. R., & Tukey, J. W. (1963). The frequency analysis of time series for echoes: Cepstrum, pseudo-autocovariance, cross-cepstrum, and saphe cracking. *Proceedings of the Symposium on Time Series Analysis* (pp. 209–243). New York: John Wiley and Sons.

Carpenter, G. A., & Grossberg, S. (1987a). A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics, and Image Processing*, **37**, 54–115.

Carpenter, G. A., & Grossberg, S. (1987b). ART 2: Self-organization of stable category recognition codes for analog input patterns. *Applied Optics*, **26**, 4919–4930.

Carpenter, G. A., & Grossberg, S. (1988). The ART of adaptive pattern recognition by a self-organizing neural network. *Computer*, **21**, 77–88.

Cohen, M. A., & Grossberg, S. (1984). Neural dynamics of brightness perception: Features, boundaries, diffusion, and resonance. *Perception and Psychophysics*, **36**, 428–456.

Cohen, M. A., & Grossberg, S. (1986). Neural dynamics of speech and language coding: Developmental programs, perceptual grouping, and competition for short term memory. *Human Neurobiology*, **5**, 1–22.

Cohen, M. A., & Grossberg, S. (1987). Masking fields: A massively parallel neural architecture for learning, recognizing, and predicting multiple groupings of patterned data. *Applied Optics*, **26**, 1866–1891.

Cohen, M. A., Grossberg, S., & Stork, D. (1987). Recent developments in a neural model of real-time speech analysis and synthesis. In M. Caudill and C. Butler (Eds.), *Proceedings of the IEEE First International Conference on Neural Networks, IV* (pp. 443–454). Piscataway, NJ: IEEE Press.

Daugman, J. G. (1985). Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America Part A*, **2**, 1160–1169.

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

Eskew, R. T., Jr. (1987). *The gap effect revisited: Slow changes in chromatic sensitivity as affected by luminance and chromatic borders*. Unpublished manuscript.

Frisby, J. P., & Mayhew, J. E. W. (1978). The relationship between apparent depth and disparity in rivalrous-texture stereograms. *Perception*, **7**, 661–678.

Graham, N. (1988). *Low-level visual processes and texture segregation*. Unpublished manuscript.

Grossberg, S. (1969). On the production and release of chemical transmitters and related topics in cellular control. *Journal of Theoretical Biology*, **22**, 325–364.

Grossberg, S. (1972). Neural expectation: Cerebellar and retinal analogs of cells fired by learnable or unlearned pattern classes. *Kybernetik*, **10**, 49–57.

Grossberg, S. (1976). Adaptive pattern classification and universal recoding, I: Parallel development and coding of neural feature detectors. *Biological Cybernetics*, **23**, 121–134.

Grossberg, S. (1982). *Studies of mind and brain: Neural principles of learning, perception, development, cognition, and motor control*. Boston: Reidel Press.

Grossberg, S. (1987a). Cortical dynamics of three-dimensional form, color, and brightness perception, I: Monocular theory. *Perception and Psychophysics*, **41**, 87–116.

Grossberg, S. (1987b). Cortical dynamics of three-dimensional form, color, and brightness perception, II: Binocular theory. *Perception and Psychophysics*, **41**, 117–158.

Grossberg, S. (1987c). Competitive learning: From interactive activation to adaptive resonance. *Cognitive Science*, **11**, 23–63.

Grossberg, S. (Ed.). (1988a). *Neural networks and natural intelligence*, Cambridge, MA: MIT Press.

Grossberg, S. (1988b). A model cortical architecture for the preattentive perception of 3-D form. In E. L. Schwartz (Ed.), *Computational neuroscience*. Cambridge, MA: MIT Press.

Grossberg, S. (1988c). Nonlinear neural networks: Principles, mechanisms, and architectures. *Neural Networks*, **1**, 17–61.

Grossberg, S. (1988d). Neural networks for visual perception in variable illumination. *Optics News*, **14**, 5–10.

Grossberg, S., & Kuperstein, M. (1986). *Neural dynamics of adaptive sensory-motor control: Ballistic eye movements*. Amsterdam: North-Holland.

Grossberg, S., & Marshall, J. A. (1987). A computational model of how cortical complex cells multiplex information about position, contrast, orientation, spatial frequency, and disparity. In M. Caudill & C. Butler (Eds.), *Proceedings of the IEEE First International Conference on Neural Networks, IV* (pp. 203–214). Piscataway, NJ: IEEE Press.

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 and Psychophysics*, **38**, 141–171.

Grossberg, S., & Mingolla, E. (1987a). A neural network architecture for preattentive vision: Multiple scale segmentation and regularization. In M. Caudill & C. Butler (Eds.), *Proceedings of the IEEE First International Conference on Neural Networks, IV* (pp. 177–184). Piscataway, NJ: IEEE Press.

Grossberg, S., & Mingolla, E. (1987b). Neural dynamics of surface perception: Boundary webs, illuminants, and shape-from-shading. *Computer Vision, Graphics, and Image Processing*, **37**, 116–165.

Grossberg, S., & Todorović, D. (1988). Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena. *Perception and Psychophysics*, **43**, 241–277.

Hubel, D. H., & Wiesel, T. N. (1965). Receptive fields and functional architectures in two nonstriate visual areas (18 and 19) of the cat. *Journal of Neurophysiology*, **28**, 229–289.

Hubel, D. H., & Wiesel, T. N. (1977). Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London (B)*, **198**, 1–59.

Hubel, D. H., & Wiesel, T. N. (1979). Brain mechanisms of vision. *Scientific American*, **241**, 150–163.

Julesz, B. (1971). *Foundations of cyclopean perception*. Chicago: University of Chicago Press.

Kaufman, L. (1974). *Sight and mind: An introduction to visual perception*. New York: Oxford University Press.

Kennedy, J. M. (in press). Line endings and subjective contours. *Spatial Vision*.

Kohonen, T. (1982). A simple paradigm for the self-organized formation of structured feature maps. In S. Amari & M. A. Arbib (Eds.). *Competition and cooperation in neural networks*. New York: Springer-Verlag.

Kohonen, T. (1984). *Self-organization and associative memory*, New York: Springer-Verlag.

Kulikowski, J. J. (1978). Limit of single vision in stereopsis depends on contour sharpness. *Nature*, **275**, 126–127.

Linsker, R. (1986a). From basic network principles to neural architecture: Emergence of spatial-opponent cells. *Proceedings of the National Academy of Sciences*, **83**, 7508–7512.

Linsker, R. (1986b). From basis network principles to neural architecture: Emergence of orientation-selective cells. *Proceedings of the National Academy of Sciences*, **83**, 8390–8394.

Linsker, R. (1986c). From basic network principles to neural architecture: Emergence of orientation columns. *Proceedings of the National Academy of Sciences*, **83**, 8779–8783.

Mayhew, J. E. W., & Frisby, J. P. (1976). Rivalrous texture stereograms. *Nature*, **264**, 53–56.

Meyer, G. E., & Dougherty, T. (1987). Effects of flicker-induced depth on chromatic subjective contours. *Journal of Experimental Psychology: Human Perception and Performance*, **13**, 353–360.

Orban, G. A., Kato, H., & Bishop, P. O. (1979). Dimensions and properties of end-zone inhibitory areas in receptive fields of hypercomplex cells in cat striate cortex. *Journal of Neurophysiology*, **42**, 833–849.

Poggio, G. F., Motter, B. C., Squatrito, S., & Trotter, Y. (1985). Responses of neurons in visual cortex (V1 and V2) of the alert macaque to dynamic random-dot stereograms. *Vision Research*, **25**, 397–406.

Rakic, P. (1988). Specification of cerebral cortical areas. *Science*, **241**, 170–176.

Richards, W., & Kaye, M. G. (1974). Local versus global stereopsis: Two mechanisms. Vision Research, **14**, 1345–1347.

Rumelhart, D. E., & Zipser, D. (1985). Feature discovery by competitive learning. *Cognitive Science*, **9**, 75–112.

Schor, C., & Wood, I. (1983). Disparity range for local stereopsis as a function of luminance spatial frequency. *Vision Research*, **23**, 1649–1654.

Schor, C., Wood, I., & Ogawa, J. (1984). Binocular sensory fusion is limited by spatial resolution. *Vision Research*, **24**, 661–665.

Shinkman, P. G., & Bruce, C. J. (1977). Binocular differences in cortical receptive fields of kittens after rotational disparate binocular experience. *Science*, **197**, 285–287.

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

Spitzer, H., & Hochstein, S. (1985). A complex-cell receptive field model. *Journal of Neurophysiology*, **53**, 1266–1286.

Thorell, L. G., DeValois, R. L., & Albrecht, D. G. (1984). Spatial mapping of monkey V1 cells with pure color and luminance stimuli. *Vision Research*, **24**, 751–769.

Todd, J. T., & Akerstrom, R. (1987). Perception of three-dimensional form from patterns of optical texture. *Journal of Experimental Psychology: Human Perception and Performance*, **13**, 242–255.

von der Heydt, R., Hänny, P., & Dürsteler, M. R. (1981). The role of orientation disparity in stereoscopic perception and the development of binocular correspondence. In E. Grastyán and P. Molnár (Eds.), *Advances in physiological science, Vol. 16: Sensory functions*. Elmsford, NY: Pergamon Press.

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

von Tschermak-Seysenegg, A. (1952). *Introduction to physiological optics*, P. Boeder (Trans.). Springfield, IL: C. C. Thomas.

Watson, A. B., & Ahumada, A. J. (1987). An orthogonal oriented quadrature hexagonal image pyramid. NASA Memo #100054, December 1987.

Werner, H. (1937). Dynamics in binocular depth perception. *Psychological Monograph* (Whole No. 218).

Willshaw, D. J., & von der Malsburg, C. (1976). How patterned neural connections can be set up by self-organization. *Proceedings of The Royal Society of London (B)*, **194**, 431–445.

Yeshurun, Y., & Schwartz, E. L. (1987). An ocular dominance column map as a data structure for stereo segmentation. In M. Caudill & C. Butler (Eds.), *Proceedings of the IEEE First International Conference on Neural Networks, IV* (pp. 371–377). Piscataway, NJ: IEEE Press.