

## Theory and Evaluative Reviews

# 3-D vision and figure-ground separation by visual cortex

STEPHEN GROSSBERG

*Center for Adaptive Systems and Department of Cognitive and Neural Systems  
Boston University, Boston, Massachusetts*

A neural network theory of three-dimensional (3-D) vision, called *FACADE* theory, is described. The theory proposes a solution of the classical figure-ground problem for biological vision. It does so by suggesting how boundary representations and surface representations are formed within a boundary contour system (BCS) and a feature contour system (FCS). The BCS and FCS interact reciprocally to form 3-D boundary and surface representations that are mutually consistent. Their interactions generate 3-D percepts wherein occluding and occluded object parts are separated, completed, and grouped. The theory clarifies how preattentive processes of 3-D perception and figure-ground separation interact reciprocally with attentive processes of spatial localization, object recognition, and visual search. A new theory of stereopsis is proposed that predicts how cells sensitive to multiple spatial frequencies, disparities, and orientations are combined by context-sensitive filtering, competition, and cooperation to form coherent BCS boundary segmentations. Several factors contribute to figure-ground pop-out, including: boundary contrast between spatially contiguous boundaries, whether due to scenic differences in luminance, color, spatial frequency, or disparity; partially ordered interactions from larger spatial scales and disparities to smaller scales and disparities; and surface filling-in restricted to regions surrounded by a connected boundary. Phenomena such as 3-D pop-out from a 2-D picture, Da Vinci stereopsis, 3-D neon color spreading, completion of partially occluded objects, and figure-ground reversals are analyzed. The BCS and FCS subsystems model aspects of how the two parvocellular cortical processing streams that join the lateral geniculate nucleus to prestriate cortical area V4 interact to generate a multiplexed representation of Form-And-Color-And-Depth, or *FACADE*, within area V4. Area V4 is suggested to support figure-ground separation and to interact with cortical mechanisms of spatial attention, attentive object learning, and visual search. Adaptive resonance theory (ART) mechanisms model aspects of how prestriate visual cortex interacts reciprocally with a visual object recognition system in inferotemporal (IT) cortex for purposes of attentive object learning and categorization. Object attention mechanisms of the What cortical processing stream through IT cortex are distinguished from spatial attention mechanisms of the Where cortical processing stream through parietal cortex. Parvocellular BCS and FCS signals interact with the model What stream. Parvocellular FCS and magnocellular motion BCS signals interact with the model Where stream. Reciprocal interactions between these visual, What, and Where mechanisms are used to discuss data about visual search and saccadic eye movements, including fast search of conjunctive targets, search of 3-D surfaces, selective search of like-colored targets, attentive tracking of multielement groupings, and recursive search of simultaneously presented targets.

---

Supported in part by the Air Force Office of Scientific Research (AFOSR 90-0175 and AFOSR F49620-92-J-0499), ARPA (AFOSR 90-0083 and ONR N00014-92-J-4015), and the Office of Naval Research (ONR N00014-91-J-4100). Thanks are extended to Cynthia E. Bradford and Diana J. Meyers for their valuable assistance in the preparation of the manuscript, and to Harald Ruda and Niall McLoughlin for their assistance with the figures. The author's mailing address is Center for Adaptive Systems, Boston University, 111 Cummington St., Boston, MA 02215.

—Accepted by previous editor, Charles W. Eriksen

### 1. Introduction

This article describes a neural network theory of biological vision, called *FACADE* theory, that suggests solutions to some long-standing problems concerning how we perceive a three-dimensional (3-D) world, notably the classical figure-ground problem of biological vision. *FACADE* theory provides a unified explanation of many visual phenomena that may, upon first encounter, seem to be unrelated. The theory explains them all as manifestations of the mechanisms whereby the visual cortex generates

informative 3-D representations of boundaries and surfaces with which to perceive the visual world. These phenomena include: how a 2-D image may generate a 3-D percept; how figures pop out from cluttered backgrounds; how spatially sparse disparity cues can generate continuous surface representations at different perceived depths; how binocular fusion of objects at different depths can deform perceptual space by different amounts, as during allelotropia; how representations of occluded regions can be completed and recognized without usually being seen; how occluded regions can sometimes be seen during percepts of transparency; how high-spatial-frequency parts of an image may appear closer than low-spatial-frequency parts; how sharp targets are detected better against a figure and blurred targets are detected better against a background; how low-spatial-frequency parts of an image may be fused while high-spatial-frequency parts are rivalrous; how sparse blue cones can generate vivid blue surface percepts; how depth attraction may occur between nearby targets and depth repulsion between farther away targets; how 3-D neon color spreading, visual phantoms, and tissue contrast percepts are generated; and how conjunctions of color and depth or other 3-D object properties may rapidly pop out as single attributes during visual search. These explanations are motivated by an ecological analysis of how monocularly viewed parts of an image inherit the appropriate depth from contiguous binocularly viewed parts, as during Da Vinci stereopsis, the equidistance tendency, and the viewing of texture stereograms. The theory is thus supported by its proposed explanations of many challenging and paradoxical psychophysical and neurobiological data that have heretofore eluded explanation. It also makes many experimental predictions whereby its mechanisms can be further tested.

FACADE theory describes a set of computational rules that are applied in a prescribed order. Consistent use of these rules provides an explanation of the types of perceptual data that are listed above. These rules may be naturally instantiated by neural processes. The ordering of the rules places strong constraints upon the processing stages at which they are most likely to occur in the visual cortex. The theory hereby suggests how key processing stages in the two parvocellular processing streams from the lateral geniculate nucleus (LGN) through prestriate cortical area V4 are organized, and how they interact (Figure 1). Although the individual model processes are few and conceptually simple, their interactions in multiple processing stages lead to subtle perceptual properties. The processing stages that are needed in the theory are linked to stages in the LGN Parvo → Interblob → Interstripe → V4 processing stream and the LGN Parvo → Blob → Thin Stripe → V4 processing stream.

Corresponding to these two parvocellular processing streams, FACADE theory contains two subsystems called the boundary contour system (BCS) and the feature contour system (FCS) for representation of 3-D boundaries and surfaces, respectively. In particular, the BCS generates emergent 3-D boundary segmentations that combine edge, texture, shading, and stereo information. The FCS

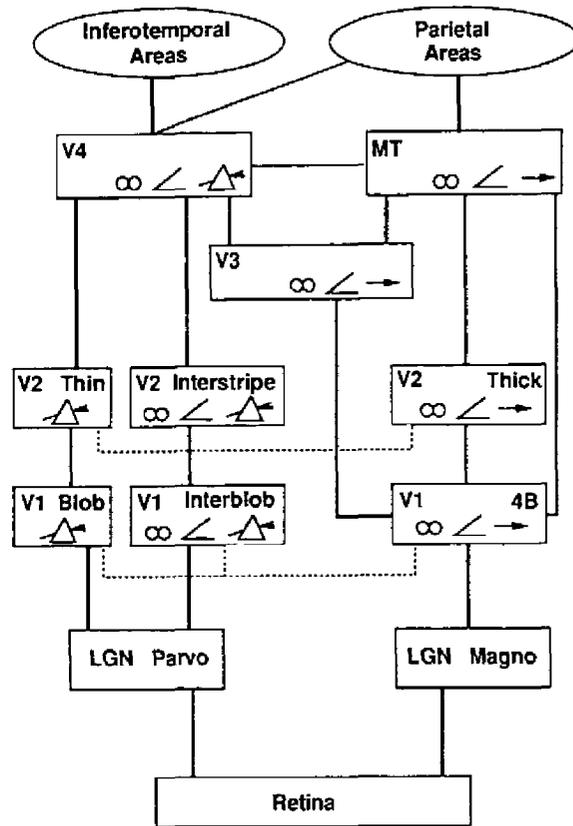


Figure 1. Schematic diagram of anatomical connections and neuronal selectivities of early visual areas in the macaque monkey. LGN = lateral geniculate nucleus (parvocellular and magnocellular divisions). Divisions of V1 and V2: blob = cytochrome oxidase blob regions, interblob = cytochrome oxidase-poor regions surrounding the blobs, 4B = lamina 4B, thin = thin (narrow) cytochrome oxidase strips, interstripe = cytochrome oxidase-poor regions between the thin and thick stripes, thick = thick (wide) cytochrome oxidase strips, V3 = visual area 3, V4 = visual area(s) 4, and MT = middle temporal area. Areas V2, V3, V4, and MT have connections to other areas not explicitly represented here. Area V3 may also receive projections from V2 interstripes or thin stripes. Heavy lines indicate robust primary connections, and thin lines indicate weaker, more variable connections. Dotted lines represent observed connections that require additional verification. Icons: rainbow = tuned and/or opponent wavelength selectivity (incidence at least 40%), angle symbol = orientation selectivity (incidence at least 20%), spectacles = binocular disparity selectivity and/or strong binocular interactions (V2) (incidence at least 20%), and pointing hand = direction of motion selectivity (incidence at least 20%). From "Concurrent processing streams in monkey visual cortex" by E. A. De Yoe and D. C. Van Essen, 1988, *Trends in Neurosciences*, 11, 219-226. Copyright 1988 by Elsevier Trends Journals. Adapted by permission.

discounts the illuminant and fills-in surface properties of brightness, color, and depth within FCS domains that are defined by BCS - FCS interactions. These interactions give rise to a multiplexed representation of Form-And-Color-And-DEpth at the final processing stage of the FCS, which is called a binocular Filling-In-Domain (FIDO).

Such a FACADE representation is hypothesized to occur in area V4 of prestriate cortex. Area V4 is hereby predicted to support figure-ground separation and to play a key role in attentive object recognition and visual search, rather than being merely a region for color perception, as originally proposed (Zeki, 1983a, 1983b). This characterization does not, however, imply that area V4 does not also participate in color perception. It is in the nature of multiplexed processing that multiple properties and tasks may be supported by each brain region. However, the binocular FIDO, and by interpretation area V4, is the *first* stage in the theory that is fully capable of 3-D figure-ground separation.

The BCS and FCS also interact reciprocally via adaptive pathways with a visual object recognition system, interpreted to occur in inferotemporal (IT) cortex. This interaction binds preattentive boundary and surface representations together into attentive object representations. This *object-attention* process (Duncan, 1984) generates categorical and prototypical representations for purposes of visual object recognition (Carpenter & Grossberg, 1993). Its circuits are interpreted in terms of interactions of IT with prestriate visual cortex, notably V4. The proposed object-attention circuits are conceptually distinguished from *spatial-attention* circuits (Posner, 1980) that transform visual information into spatial maps, interpreted to occur in parietal cortex (Figure 1). The theory suggests how parallel outputs from the FCS and a motion system called the motion BCS (Grossberg and Mingolla, 1992; Grossberg & Rudd, 1989, 1992) converge upon this spatial-attention system. The FCS, for example, enables visual search to be restricted to items of a particular color (Egeth, Virzi, & Garbart, 1984) or color-depth combination (Nakayama & Silverman, 1986). The motion BCS is suggested to operate via the magnocellular cortical processing stream (Figure 1), using mechanisms that are suggested to be closely related to apparent motion mechanisms (Grossberg, 1991, in press) and that may contribute to predictive eye tracking (Fischer, 1986; Mountcastle, 1978; Wurtz, Goldberg, & Robinson, 1982). These object-attention and spatial-attention processes thus model aspects of the What cortical stream through IT cortex, which subserves visual object recognition, and the Where cortical stream through parietal cortex, which subserves spatial localization & orientation (Goodale & Milner, 1992; Mishkin, Ungerleider, & Macko, 1983; Ungerleider & Mishkin, 1982). BCS and FCS interactions with these attentive object recognition and spatial localization systems are suggested to control visual search for targets among distractors, a task that often depends upon a competence for figure-ground separation (Schiller & Lee, 1991).

The article does not describe mathematical equations or computer simulations. In order to be maximally accessible, it relies on consistent use of the same small set of qualitatively described concepts and mechanisms to explain a large interdisciplinary data base. Each such use is supported by quantitative articles where such processes have been mathematically characterized and computationally simulated in order to explain related perceptual and

neural data. These references are described in the text. Thus, the article is self-contained, but builds upon a decade of quantitative modeling which ensures that these qualitatively described processes actually work the way they are claimed to work.

These concepts and mechanisms will be explained in stages in the text. Some of them are briefly summarized here, without explanation, for the reader who desires a compressed glimpse of what lies ahead. The main new insights concern how and why the brain processes visual information about near and far objects in an asymmetric way. This asymmetry prevents erroneous boundary segmentations and surface filling-in events from occurring among occluding and occluded objects. As a result, many interactions between and within the BCS and FCS are *partially ordered* from larger scales and disparities to smaller scales and disparities. Moreover, a *boundary contrast* process occurs, such that stimuli that strengthen one boundary more than they do a contiguous boundary, whether due to contrast, spatial frequency, or disparity differences, help to initiate figure-ground pop-out.

On a finer level of detail: Model complex cells with large receptive fields can binocularly fuse a broader range of disparities than can cells with small receptive fields. Complex cells of the same receptive field size compete across disparities at each position. A pool of near-zero disparity complex cells cooperates with multiple pools of nonzero disparity complex cells to form multiple boundary segmentations. The cortical magnification factor helps to convert disparity computations at different retinal eccentricities that correspond to a prescribed range of relative depths from the observer into a single boundary segmentation. Cortical ocular dominance columns help to organize the amount of allelotropia and the size-disparity correlation that occur during binocular viewing into a decision as to whether a fused binocular boundary segmentation or binocular rivalry will occur. Multiple self-similar networks of simple cells, complex cells, hypercomplex cells, higher order hypercomplex cells, and bipole cells form the multiple boundary segmentations that correspond to different relative depths from the observer. Ordered competitive and cooperative interactions across position, spatial frequency, orientation, and disparity give rise to the coherent 3-D boundary segmentations that separate and complete the boundaries of occluding and occluded objects. Moreover, each boundary segmentation cooperates across multiple receptive-field sizes at each position to generate the positionally and orientationally most accurate segmentations that are possible. Each 3-D boundary segmentation interacts with monocular featural data within the FCS to select the monocular featural data that are consistent with the segmentation. The first such FCS filling-in event occurs within a circuit called a monocular FIDO. Only binocularly consistent featural data are allowed to fill-in FIDO surface regions. Only regions surrounded by a *connected* boundary can fill-in. The regions that manage to fill-in surface regions at a given depth excite BCS boundary and FCS feature signals at the same position and depth, but inhibit them at the same position

but farther depths. Boundary-surface consistency is hereby ensured. All surviving boundary segmentations that correspond to nearer depths are added to the boundary segmentations corresponding to farther depths. These last two operations illustrate the partial ordering of interactions that is used to separate figure from ground. The augmented boundary segmentations activate the surviving FCS signals to trigger surface filling-in events that generate the final FACADE representation within a circuit that is called a binocular FIDO.

The article is broken into three parts. Part I discusses some challenging perceptual data about 3-D vision and figure-ground separation in order to highlight some of the key data properties and conceptual paradoxes that need to be explained. Part II reviews just enough of the theory as it existed before the present work to make the article self-contained. Part III uses this foundation to extend the theory and to explain the data summarized in Part I as well as a large body of related data. Readers who are familiar with the theory can skip directly from Part I to Part III.

## PART I

### 2. Da Vinci Stereopsis and Filling-In

The theory may be motivated by the following example, which is experienced ubiquitously as we view 3-D layouts during our daily lives. When we view a farther surface that is partly occluded by a nearer surface, one eye typically registers more of the farther surface than the other eye does. Our conscious percept of the farther surface is often derived from the view of the eye that registers more of this surface. For example, under the viewing conditions depicted in Figure 2, observers see the right-eye view in depth, even though the image region that lies between the vertical lines B and C is registered by only the right eye. This type of ubiquitous perceptual condition has been known since the time of Leonardo da Vinci, and is often called Da Vinci stereopsis (Gillam & Borsting, 1988; Kaye, 1978; Lawson & Gulick, 1967;

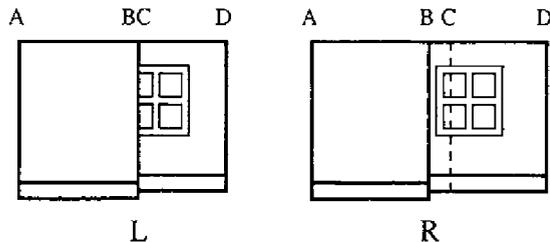


Figure 2. When a scene is viewed by both eyes, most of it may be binocularly detected, such as regions AB and CD, but part of it may be detected by only one eye, such as region BC. An appropriate depth of the monocularly viewed region is often filled-in using information from retinally contiguous, binocularly viewed regions.

Nakayama & Shimojo, 1990; Wheatstone, 1838). Some of the challenging perceptual properties that subserve this apparently innocuous percept will now be illustrated by considering them under simpler stimulus conditions.

#### A. Deformable Fusion by Allelotropia

Because each eye views the world from a different position in the head, the same material point on an object is registered at a different location on the two retinas, except for the object region that is foveally fixated by both eyes. In order to binocularly fuse such a disparate pair of monocular images, the two images must be deformed into one image. A simple case of this process is the phenomenon of *displacement*, or *allelotropia* (Kaufman, 1974; von Tschermak-Seysenegg, 1952; Werner, 1937). In this phenomenon, when a pattern EF G is viewed through one eye and a pattern E FG is viewed through the other eye, the letter F can be seen in depth at a position halfway between E and G. Thus, the process of binocular fusion deforms the two monocular appearances of F into one binocular percept of F whose spatial position differs from either monocular position of F with respect to E and G. This deformation of F's relative position is necessitated by the large disparity of the two monocular F positions when E and G are binocularly fused.

During inspection of a 3-D scene, the amount of deformation needed to achieve binocular fusion depends upon how far away each object is with respect to an observer's retinas. For example, crossed-disparity images of closer objects are more disparate than images of farther objects. Thus, different parts of the left-eye and right-eye images are deformed by different amounts to generate a single binocular percept of the world. During Da Vinci stereopsis, the vertical boundaries of regions AB and CD in the left/right-eye images of Figure 2 are deformed by different amounts in order to be binocularly fused. Given that different amounts of deformation are required to binocularly fuse the monocular boundaries of objects at different distances, why are no "holes" in binocular perceptual space created?

#### B. Distance of Zero Disparity Points

Some other basic facts about binocular vision also have profound implications for vision theories. For example, the retinal images of objects at optical infinity have zero disparity on the two retinas, and the disparities on the two retinas of corresponding object points tend to increase as an object approaches the observer. This is the familiar reason for assuming that larger disparities are an indicator of relative closeness.

On the other hand, when both eyes focus on a single point on a planar surface viewed in depth, the fixation point is a point of zero disparity. Points of the surface that are registered by the retinas farther from the fixation point generate larger binocular disparities. Why does a plane not recede toward optical infinity at the fixation point and curve toward the observer at the periphery of the visual field? Why does the plane not become distorted

in a new way every time our eyes fixate on a different point within its surface?

For present purposes, a key fact is that zero disparity also occurs under monocular viewing conditions. In particular, the region BC in the right-eye image of Figure 2 is monocularly viewed. Yet this region is perceived as a continuous extension in depth of the binocularly viewed region CD. How does the monocularly viewed region BC inherit the depth of the binocularly viewed region CD? Why, then, are unpaired monocularly viewed regions of stereograms always seen in back (Julesz, 1971; Nakayama & Shimojo, 1988)?

### C. Equidistance Tendency and Emmert's Law

These properties of Da Vinci stereopsis are closely related to the *equidistance tendency* that has been studied by Gogel (1956, 1965, 1970). Gogel noted that if one object is viewed monocularly through a mirror arrangement, whereas all other objects in a scene are viewed binocularly, then the monocularly viewed object seems to lie at the same distance as the retinally most contiguous binocularly viewed object. Emmert (1881) earlier reported the analogous percept that a monocular afterimage seems to be located on any surface that a subject fixates binocularly while the afterimage is active. Collett (1985) and Buckley, Frisby, and Mayhew (1989) have studied this phenomenon psychophysically by using textured stereograms in which a pair of textured regions could be matched binocularly, but an intervening region was defined by monocular information to only one eye.

How does the region BC in Figure 2 inherit the depth of the region CD, and analogously, how do Gogel's equidistance tendency and Emmert's law obtain? These effects may be explained by a filling-in process that selectively completes a BC surface representation at a depth corresponding to that of region CD. A variety of recent experiments have demonstrated that a filling-in process does, indeed, complete various depthful surface properties (Nakayama, Shimojo, & Ramachandran, 1990; Nakayama, Shimojo, & Silverman, 1989; Takeichi, Watanabe, & Shimojo, 1992; Watanabe & Cavanagh, 1992). To explain how this occurs, the theory utilizes the following types of processes.

### D. Binocular and Monocular Boundary Representation

The filling-in process is contained by internal representations of scenic boundaries. Some boundaries are viewed binocularly, others are viewed monocularly. In the example of Figure 2, we need to show how the boundaries A and B in the left and right images are binocularly fused, and how the boundaries within region CD are binocularly fused. As noted above, fusion of the AB boundaries and the CD boundaries causes different amounts of allelotropia. The monocularly viewed boundaries in region BC of the right-eye view are not binocularly fused; hence, they do not register a binocular disparity in their internal cortical representation. The same is true for all horizon-

tal boundaries in the image. Thus, there are at least three ways in which an image can be registered with zero, or near-zero, disparity: as an occluded region during Da Vinci stereopsis, as an entire image that is monocularly viewed, or as a horizontal boundary during either monocular or binocular viewing. The theory suggests that all such near-zero disparity boundaries are processed in a separate pool of near-zero disparity cortical cells. The following discussion indicates how the theory makes use of this property.

### E. The Near-Zero Disparity Cell Pool

We need to explain how the monocularly viewed, near-zero disparity vertical and horizontal boundaries in region BC are joined with the binocularly fused, large-disparity vertical boundaries and horizontal near-zero disparity horizontal boundaries in region CD to form the window frame in Figure 2. Disparity-sensitive cortical cells are tuned to a limited range of disparities. The theory assumes that active near-zero disparity cells, whether they are monocularly or binocularly activated, give rise to spatially organized boundary signals that are combined with the spatially organized activations of cells that code nonzero disparities to create a more complete boundary representation. The nonzero disparity cells are themselves assumed to be segregated into separate cell pools that are organized, in a manner described below, to correspond to different relative depths of an observed image feature. Thus near-zero disparity cells are assumed to add their boundary activations to multiple boundary representations, each corresponding to a differently tuned pool of nonzero disparity cells. This property suggests a new functional interpretation of psychophysical evidence (Regan, Erkelens, & Collewijn, 1986; Richards & Regan, 1973) and neurophysiological evidence (Poggio & Talbot, 1981) that near-zero disparities, crossed disparities, and uncrossed disparities are processed by separate cell pools in the visual cortex.

The theory also segregates disparity-sensitive cells according to their receptive field sizes, or spatial scales, and suggests how, and for what functional purpose, different receptive field sizes binocularly fuse a different range of binocular disparities, as in the size-disparity correlation (Kulikowski, 1978; Richards & Kaye, 1974; Schor & Tyler, 1981; Schor & Wood, 1983; Schor, Wood, & Ogawa, 1984; Tyler, 1975, 1983). Thus it is assumed that BC boundaries are added to the CD boundaries at the scales and disparities that are capable of computing binocularly fused CD boundaries. For those spatial scales and nonzero disparities at which all these boundaries exist, the composite BCD boundaries enclose *connected* regions, such as the connected window frame in the right-eye image of Figure 2.

It should not be inferred from this verbal discussion of "separate cell pools" of disparity-tuned and spatial-frequency-tuned cells that this separation is maintained in the final visible percept. Instead, it is suggested below how cells sensitive to different positions, spatial frequencies, orientations, and disparities are combined by context-

sensitive competitive and cooperative interactions into the 3-D boundary segmentations that organize visible percepts. In this regard, the composite BCD boundaries in Figure 2 will enclose a connected region only if the following problem can be solved.

**F. 3-D Emergent Boundary Completion**

Due to allelotropia, the binocularly fused boundaries within region CD may be positionally displaced relative to the monocularly viewed boundaries within region BC. As a result, gaps may occur between the cortical loca-

tions of cells that represent these boundaries. When the monocularly and binocularly viewed regions contain oblique contours, the responses of cortical cells may be both orientationally and positionally displaced. These gaps and misalignments need to be corrected by a boundary-completion process. The theory explains how each pool of cells corresponding to a different range of nonzero disparities is capable of generating an emergent boundary segmentation that is triggered by the active cells in its disparity range augmented by the active near-zero disparity cells. Such a process realigns and connects the bound-

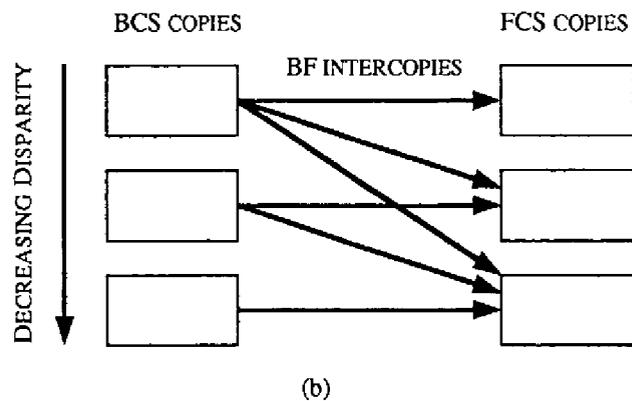
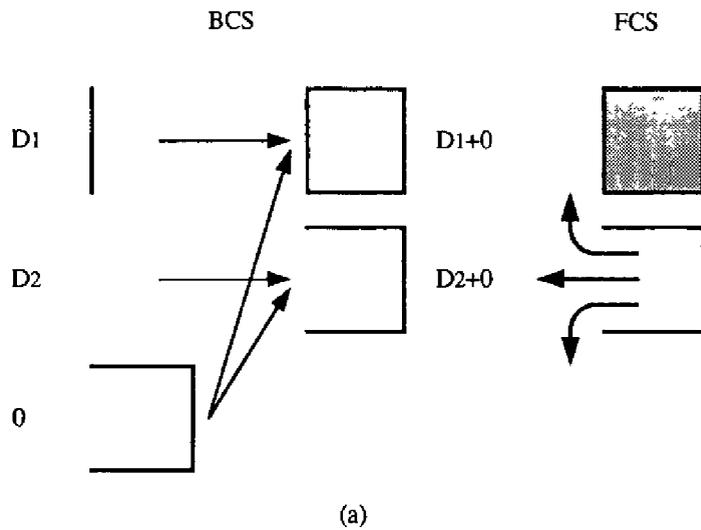


Figure 3. (a) Near-zero disparity boundaries add to the boundaries of all nonzero disparity cell pools. Only the boundaries that enclose a connected region can contain the filling-in of surface properties. (b) Boundaries corresponding to nearer objects are added to boundaries corresponding to farther objects to contain the filling-in of figure-ground-separated surface representations in the perceived visual representation. In more technical terms, each feature contour system (FCS) copy receives inhibitory boundary-gating signals, or BF intercopies, from one or more boundary contour system (BCS) copies. The BF intercopy inputs are partially ordered from larger disparity to smaller disparity BCS copies. Each FCS copy contains three pairs of opponent Filling-In Domains (FIDOs). A FIDO is defined in the text.

aries that join regions BC and CD, thereby generating boundaries that completely enclose the window frame in Figure 2.

### G. Filling-In Surface Properties of Connected Regions

The connected boundaries within region BCD form a sparse and discontinuous representation of the scene. How are the scene's continuous surface properties, including their brightnesses, colors, and surface depths, generated? The theory explains how boundaries that enclose *connected* regions in BCD, and *only* these boundaries, can trigger filling-in of surface properties of these regions that form part of the final visible 3-D percept (Figure 3a). It is assumed that multiple filling-in domains exist. Each filling-in domain corresponds to boundaries that are sensitive to a restricted range of binocular disparities. Thus, the filled-in representations combine properties of surface depth, position, orientation, brightness, and color. These multiplexed properties may be compared with analogous receptive field profiles of cells in cortical area V4 (Desimone, Schein, Moran, & Ungerleider, 1985; Zeki, 1983a, 1983b). A key insight of the theory is thus to show how the monocularly viewed region BC selectively fills-in depthful surface properties within the filling-in domain corresponding to the binocularly fused boundaries of region CD.

### H. Boundary Invariance Under Eye Fixation Changes

One can begin to see from this property how the near-zero disparity cell pool helps to compensate for changes in eye fixation. Suppose, for example, that the eyes binocularly fixate one of the vertical boundaries in region CD, and therefore register it via near-zero disparity cells rather than nonzero disparity cells. These near-zero disparity activations are added to all the nonzero disparity cell activations, in their respective cell pools. If there are enough nonzero disparity activations in a given pool to form a connected boundary, then this boundary can support the filling-in of surface properties in a filling-in domain that represents the corresponding relative depth from the observer.

Changes in eye fixation also alter all the binocular disparities that are registered by the brain. The theory suggests how the brain combines responses of cells that correspond to the same relative depth from the observer into a single boundary representation. It uses the cortical magnification factor to start the process whereby cells that code larger disparities at larger distances from the fovea may input to the same boundary representation.

### I. Near Boundaries Obstruct Filling-In of Occluded Regions

Why does not the filling-in of BC by CD also propagate across region AB? The binocular boundary B is fused at a disparity corresponding to a nearer surface than are the boundaries of region CD. Without further processing, boundary B could not form a connected boundary around

region BD. Nor could it prevent filling-in of region AB within the filling-in domain whose depth corresponds to region CD from continuing across region AB. Filling-in would also occur within the "correct" filling-in domain whose depth corresponds to boundaries A and B of region AB. If both filling-in events could occur, region AB would appear transparent; it would be represented by two different filled-in representations at two different depths from the observer. This example illustrates the general problem that if filling-in is the basis for many surface depth percepts, why do not *all* such surfaces look transparent?

The theory suggests that this does not happen because the boundaries that correspond to closer objects are added to the boundaries that correspond to further objects in the filling-in domains (Figure 3b). As a result, filling-in that is initiated in region BD does not flow behind region AB. This restriction upon filling-in of surface properties does not prevent *boundaries* from being completed behind an occluding region. Since direct interactions are assumed to exist from boundary representations to the object recognition system, some occluded objects or object parts may be recognized via their completed boundaries, even if visible surface properties are not filled-in behind the occluding object. The object recognition system may, in turn, send attentive prototypical signals back to the boundary segmentation system, thereby using prior knowledge to prime the system, select among ambiguous boundary segmentations, and attentively enhance some parts more than others. Thus, the preattentively completed boundary segmentation may be attentively modified.

These properties of Da Vinci stereopsis will be explained below in terms of the theory's computational rules. The theory will also be used to explain a variety of other basic perceptual processes that challenge traditional views about how 3-D visual perception takes place. For example, the discussion in Section 2 illustrates that a new analysis is needed of how the multiple spatial scales that are used for early visual filtering interact with later boundary segmentation processes that group, or bind, visual features into surface and object representations. The need for a fresh analysis of these interactions is also indicated by a powerful set of demonstrations of how figure-ground perception depends upon spatial frequency. I will collectively call these demonstrations the Weisstein effect.

## 3. Spatial Frequency Influences on Figure-Ground Perception

The Weisstein effect shows how paradoxical 3-D perceptual properties can occur in response to even simple images that are constructed from multiple spatial frequencies. These images show that our understanding of early filtering and how it interacts with grouping processes is incomplete. In particular, it is often stated that low spatial frequencies selectively process near objects and high spatial frequencies selectively process far objects, because the images of an object on an observer's two retinas increase in size and disparity as the distance between object and observer decreases. This relationship is often

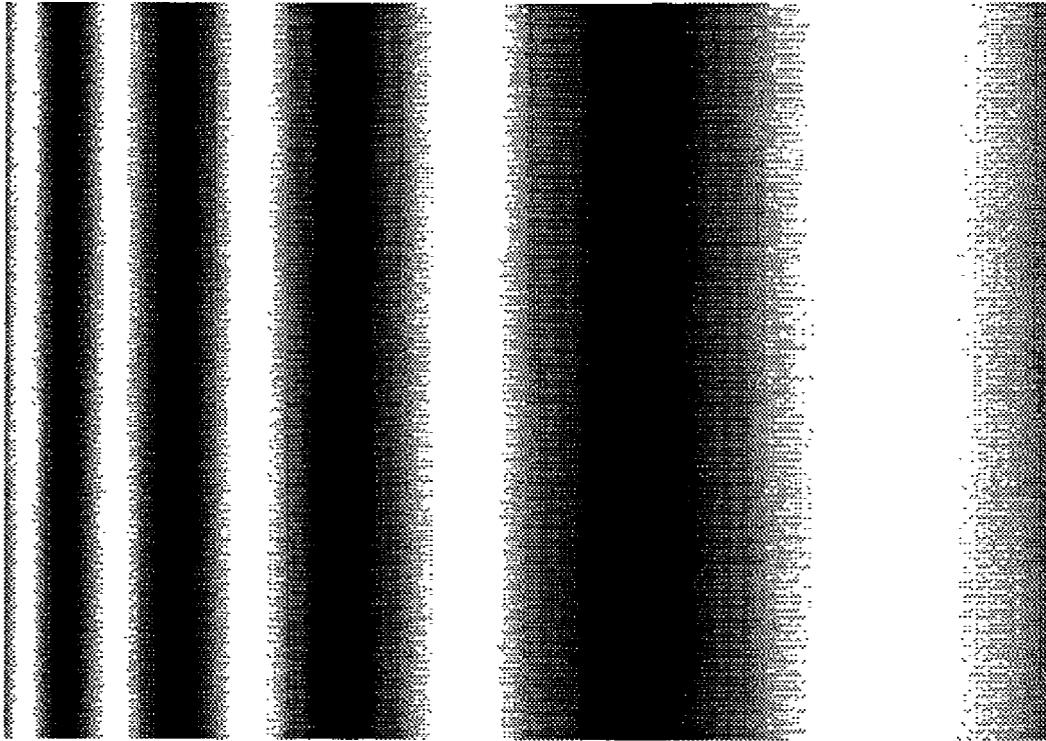


Figure 4. The higher spatial frequencies appear to be farther away than the lower spatial frequencies.

called the size-disparity correlation. An illustration of this effect is shown in Figure 4, where the low-spatial-frequency region of the image appears closer than its high-spatial-frequency region. This property also contributes to perceptions of depth from monocular perspective gradients, one of the key demonstrations of ecological psychology (Gibson, 1950).

In contrast to this property, Brown and Weisstein (1988b) have demonstrated that if regions filled with relatively higher spatial frequency sinusoidal gratings are adjacent to regions containing relatively lower spatial frequency gratings, the regions with the higher frequency appear closer in depth than those containing the lower frequency, as illustrated in Figure 5.

A comparison of the opposite dependence between spatial frequency and depth in Figures 4 and 5 shows that whether a spatial frequency difference signals "near" or "far" depends upon the global organization of the image, notably how the image is segmented by boundaries, not merely upon a spatial frequency difference per se.

Related studies clarify that the Weisstein effect can play a powerful role in the figure-ground separation of images capable of supporting reversible figure-ground configurations. Whether the configuration is a disk-annulus (Wong & Weisstein, 1982, 1983), a bipartite field, a Maltese cross, or a Rubin (1921) faces/vase (Klymenko

& Weisstein, 1986; Figure 5b), the regions containing the relatively higher spatial frequency appeared as a figure more often than did the regions containing the relatively lower spatial frequency. Correspondingly, the figure appeared closer in depth than did the ground. Pentland (1985) has also shown that the more an occluded object is blurred, the more the occluding object appears to be closer in depth.

Brown and Weisstein (1988a) noted that a number of authors have proposed that "figure and ground analysis is carried out by systems with different information processing characteristics" (p. 56). Such a view is hard to support because the relative depths and sizes of figure and ground regions may vary between wide limits, and a switch between figure and ground percepts in response to reversible figures would require a switch to occur between these systems without any change occurring in image characteristics. Thus, the same image parameters which may be interpreted as a figure at one moment may be interpreted as a ground during the next moment. Likewise, a region that may be perceived as a figure in one image may be perceived as a ground in another image. For example, the "low"-spatial-frequency regions in Figure 5 may be the "high"-spatial-frequency regions in a different image which uses an even lower spatial frequency in the alternating regions. No simple system

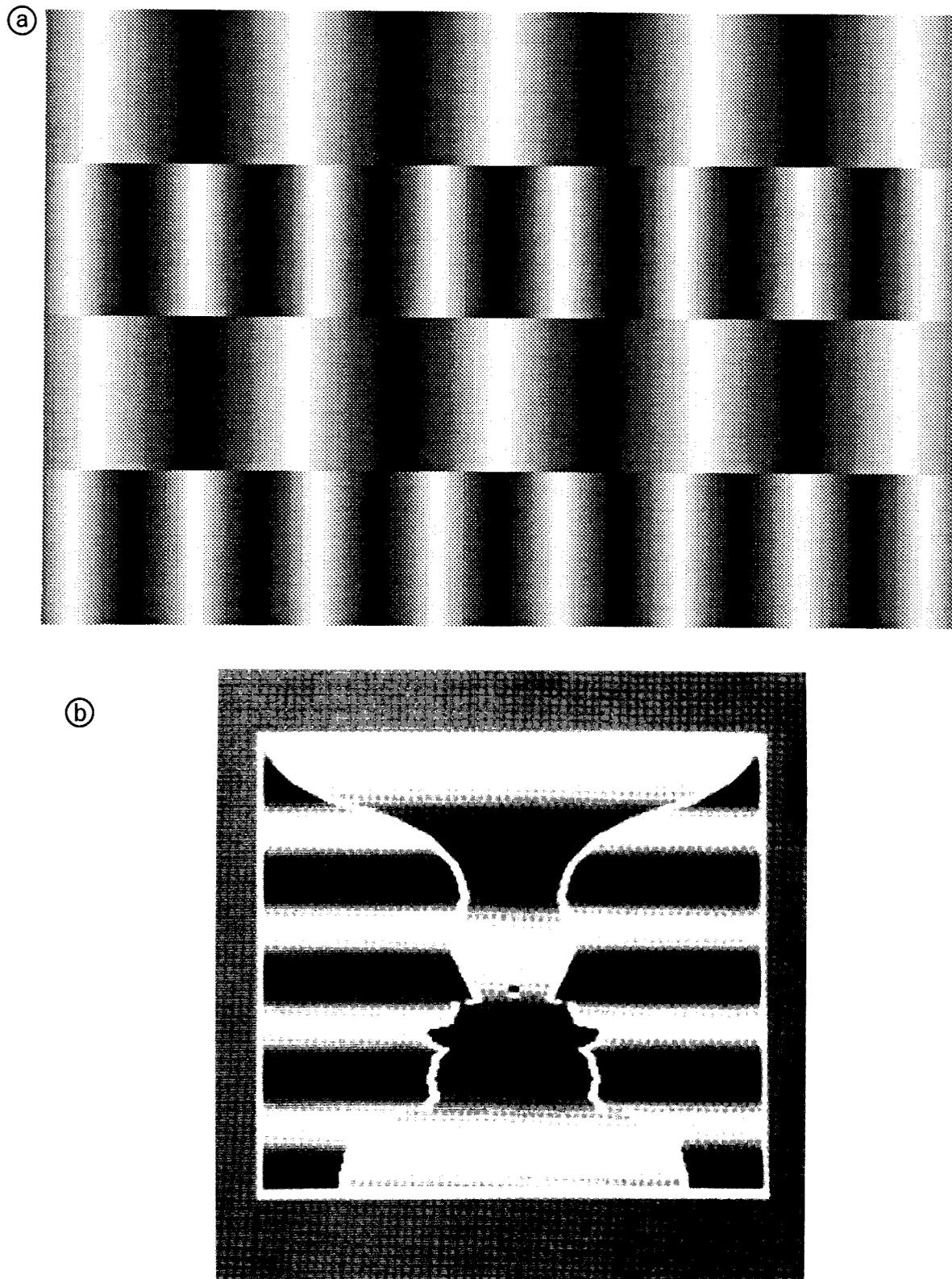


Figure 5. The higher spatial frequencies appear to be closer than the lower spatial frequencies. (a) From "A spatial frequency effect on perceived depth" by J. M. Brown and N. Weisstein, 1988, *Perception & Psychophysics*, 44, 157-166. Copyright 1988 by the Psychonomic Society. Adapted by permission. (b) From "Spatial frequency differences can determine figure-ground organization" by V. Klymenko and N. Weisstein, 1986, *Journal of Experimental Psychology: Human Perception & Performance*, 12, 324-330. Copyright 1986 by the American Psychological Association. Reprinted by permission.

dichotomy in terms of locally defined properties, such as spatial frequency, can cope with this type of global context sensitivity.

These data challenge theories to explain how the expected relationship between spatial frequency and depth, as shown in Figure 4, may be reversed by boundary segmentation processes, as shown in Figure 5, to influence which parts of an image or scene will appear as figure and which as ground. Relative depth may also be influenced by factors other than spatial frequency, notably binocular disparity, which the spatial frequency effect can override (Brown & Weisstein, 1988b). Such data show that the relationship between spatial frequency, binocular disparity, and relative depth is not captured by such models as that of Marr and Poggio (1979), which restrict their attention to the early processing of stereo information. One task of the present theory is to further develop the mechanisms, outlined in Grossberg (1987c), that distinguish early processing of stereo disparity from later processing of surface depth and form.

#### 4. 3-D Percepts of Occluded and Occluding Figures in 2-D Pictures

The spatial organization of occluding and occluded objects also has a powerful influence on depth perception,

such that image regions that correspond to partially occluded objects may appear to lie behind the occluding objects. This is true during inspection of 2-D pictures as well as during inspection of 3-D scenes. Inspection of the occluded B shapes in Figure 6b (Bregman, 1981; Kanizsa, 1979) illustrates that occluded objects in a 2-D picture may function perceptually as image "figures" even though they are not perceived as the nearest objects in the scene. Thus, although spatial frequency and/or binocular disparity may cause figures to pop out of a scene by making them appear closer, this is not the only way for figures to be defined. A comparison of Figures 6b and 6c shows that the existence of the black sinewy shape in front of the occluded Bs is needed in order for them to be readily recognized as Bs.

How does a 2-D image create a 3-D percept of occluding objects in front of occluded objects, as in Figure 6b? How are the occluded objects recognized in Figure 6b but not Figure 6c even though they are equally well seen in both? A comparison of Figures 6b and 6c illustrates that properties of form, color, and depth interact to generate a percept, and that this interaction may, as in Figure 6b, or may not, as in Figure 6c, generate a 3-D representation of a 2-D image. This 3-D representation enables the occluded parts of the B shapes to be completed for pur-

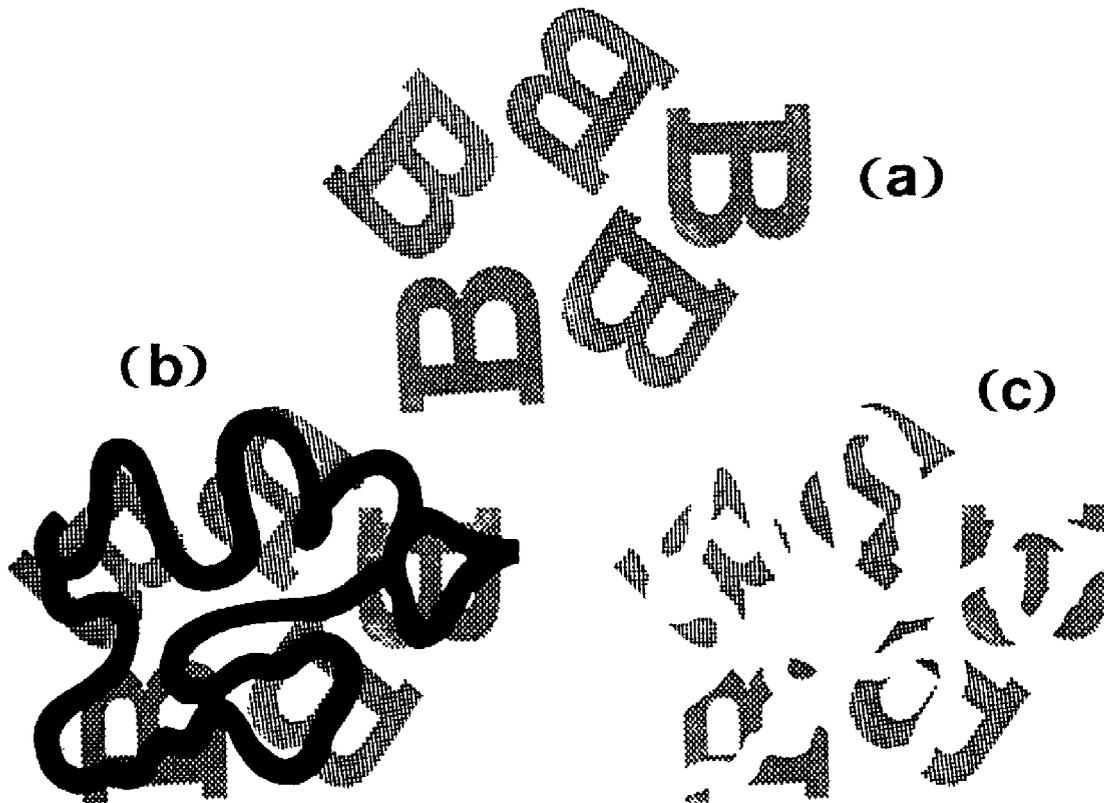


Figure 6. Role of occluding region in recognition of occluded letters: (a) Uppercase "B" letters; (b) same, except partially hidden by a black snake-like occluder; (c) same, except that the occluder is white, and therefore merges with the remainder of the white background. Although the exposed portions of the letters are identical in (b) and (c), they are much better recognized in (b). From "Stereoscopic depth: Its relation to image segmentation, grouping, and the recognition of occluded objects" by K. Nakayama, S. Shimojo, and G. H. Silverman. 1989, *Perception*, 18, 55-68. Copyright 1989 by Pion Ltd. Reprinted by permission.

poses of recognition in response to Figure 6b but not Figure 6c, even though the occluded regions are not seen in either Figure 6b or Figure 6c.

Nakayama, Shimojo, and Silverman (1989) have argued, from such a comparison and related experiments, that partially occluded objects "are enclosed by two types of bounding contour: those inherent to the object itself (intrinsic) and those defined by occlusion (extrinsic). . . . Extrinsic contours . . . must be removed prior to a stage of template matching. . . . The common border is attached to and regarded as intrinsic to the closer region. . . . Intrinsic borders aid in the segmentation of image regions and thus prevent grouping, whereas extrinsic borders provide a linkage to other extrinsic borders and facilitate grouping" (p. 55). In summary, the common boundaries of the black occluder and the gray B shapes are somehow removed from the B representations and assigned to the occluder.

Nakayama, Shimojo, and Silverman (1989) traced the basis for this distinction to "the natural constraints of the real world [which] dictate that the border always 'belongs to' the region corresponding to the closer object. Similarly, it does not belong to the farther object and is thus extrinsic to it" (p. 58). Such a description is intuitively appealing, but it does not explain the phenomenon. It calls "natural constraints" those which exist and which must therefore be explained, without characterizing how "natural constraints" are functionally and mechanistically distinguished by the brain. A key insight of the Nakayama, Shimojo, and Silverman (1989) argument will nonetheless be supported in the subsequent analysis—namely, that the boundaries that are shared by the gray B shapes and the black occluder are detached from the remaining B boundaries. The shared boundaries are used to generate a boundary segmentation and filled-in surface representation of the black occluder "in front of" the surface on which the B fragments lie. When the remaining B boundaries are freed from the shared boundaries, they can generate a more complete boundary segmentation of whole B letters. At a later processing stage, the boundaries of the black occluder, including the shared boundaries, are *reattached* to the B shapes in the filling-in domains to prevent the gray color of the Bs from flowing "behind" the black occluder and thereby rendering it transparent, much as the nearer B boundary in the Da Vinci stereopsis display of Figure 2 prevents filling-in of the surface BD into the region AB using the mechanism depicted in Figure 3b.

Thus, the Nakayama et al. claim is both supported and contested by the present theory. In either case, an argument based upon "natural constraints" cannot explain many depthful percepts, including the Weisstein effect, or various percepts described below which involve illusory contours, which I will explain using the same mechanisms. In all these cases, a unified explanation derives from an analysis of how a continuous surface representation at a fixed depth can be synthesized from spatially discontinuous image signals that may not even activate binocular disparities that represent the same depths.

In the case of the Weisstein effect, an interaction between the boundary segmentations of *multiple spatial scales* generates a 3-D percept from a 2-D image. In the case of the Bregman-Kanizsa Bs, an interaction between the boundary segmentations of *differently colored regions* generates a 3-D percept from a 2-D image. We need to analyze how the Bregman-Kanizsa form-color interaction selectively activates some spatial scales more than others, and thereby generates a 3-D percept in much the same way as in the Weisstein effect. In both cases, we need to understand how selective activation of some scales more than others creates the basis for a percept of relative depth, and how this depth difference may be used to prevent filling-in of occluded regions "behind" occluding regions.

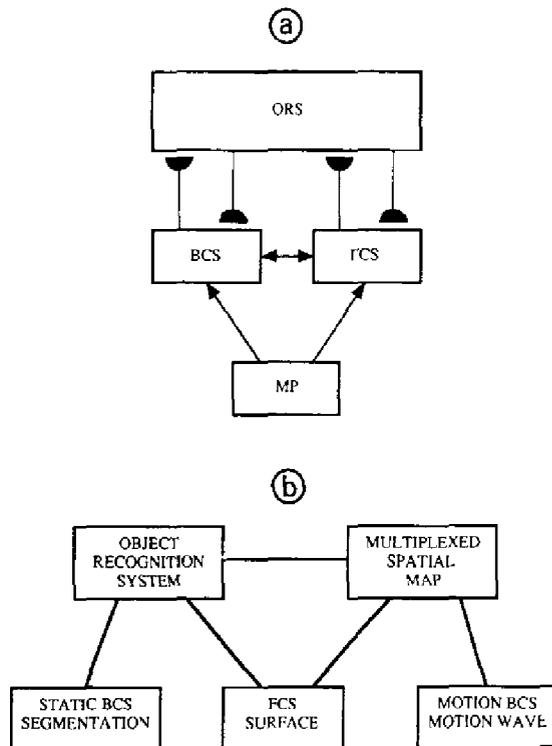


Figure 7. (a) Completed boundaries within the boundary contour system (BCS) can be recognized within the visual object recognition system (ORS) via direct BCS - ORS interactions whether or not they are seen in the feature contour system (FCS) by separating two regions with different filled-in brightnesses or colors. The FCS - BCS interactions are introduced in this article. The monocular preprocessing (MP) stage is defined in the text. (b) The reciprocal interactions of the ORS with BCS and FCS are supplemented by reciprocal interactions with an attentive spatial map. These object-based and spatial-based interactions are used to coordinate attentive object recognition, spatial orientation, and visual search. Expressed somewhat more technically, the static BCS and FCS interact reciprocally with the ORS for purposes of visual object recognition. The FCS and motion BCS interact reciprocally with a multiplexed spatial map for purposes of spatial attention. Both systems interact together to define spatially invariant ORS recognition categories and to control visual search. See text for details.

### 5. Ocluded Boundary Completion and Recognition Without Filling-In

Even if the shared boundaries between occluder and B shapes in Figure 6b are somehow deleted, how does an observer so quickly recognize the incomplete B figures? The boundary completion process of the present theory is capable of generating illusory contours between the (approximately) colinear line ends of the incomplete B figures (Grossberg & Mingolla, 1985a, 1985b, 1987a). This property of illusory contour completion raises a central question in visual perception for which the theory offers an answer, namely: if illusory contours complete the B shapes and thereby enhance their *recognition*, why do we not *see* these illusory boundaries in the sense of detecting a perceived brightness or color contrast at their locations?

Figure 7 schematizes part of the answer. A boundary that is completed within the segmentation system (denoted BCS) does not generate visible contrasts within the BCS. In this sense, *all boundaries are invisible*. Visibility is a property of the surface filling-in system (denoted FCS). The completed BCS boundary can directly activate the ob-

ject recognition system (ORS) whether or not it is visible within the FCS. Within the present theory, the ORS is predicted to include the inferotemporal cortex (Mishkin, 1982; Mishkin & Appenzeller, 1987; Schwartz, Desimone, Albright, & Gross, 1983), whereas the FCS visible surface representation is predicted to include area V4 of the pre-striate cortex (Desimone et al., 1985; Zeki, 1983a, 1983b).

In summary, a boundary may be completed within the BCS, and thereby improve pattern recognition by the ORS, without necessarily generating a visible brightness or color difference within the FCS. This key insight of the theory has made it possible to explain many perceptual properties that are otherwise mysterious. Figure 8 provides several illustrations of these processes at work. In Figures 8a and 8b, respectively, the vertical illusory boundary and the circular illusory groupings are vivid even though they do not correspond to large perceived contrast differences. Figure 8c illustrates the type of display that Biederman (1985, 1987) has used to study recognition of incomplete and briefly presented line drawings that are derived from a familiar line drawing by deleting some of the lines. In some cases, completing deleted lines by colinearly con-

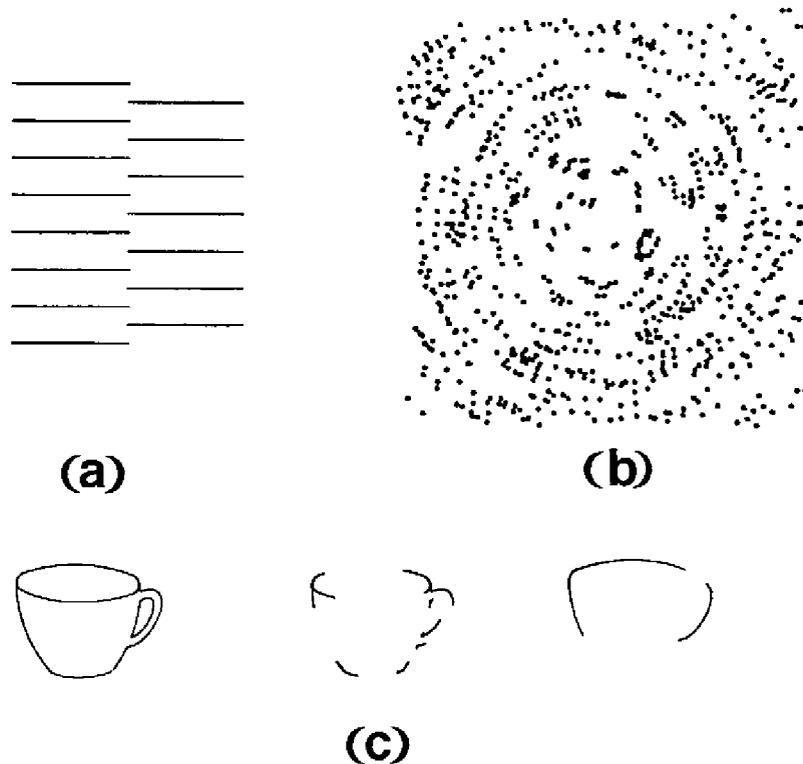


Figure 8. (a) The vertical line is easily recognized in the absence of a vertically oriented contrast difference. (b) A Glass pattern. The emergent circular pattern is recognized without being seen. From "Pattern recognition in humans: Correlations which cannot be perceived" by L. Glass and E. Switkes, 1976, *Perception*, 5, 67-72. Copyright 1976 by Pion Ltd. Reprinted by permission. (c) A line drawing (left column) with recoverable (middle column) and unrecoverable (right column) deletions. Completion of recoverable deletions facilitates recognition even though they are not seen. From "Human image understanding: Recent research and a theory" by I. Biederman, 1985, *Computer Vision, Graphics, and Image Processing*, 32, 29-73. Copyright 1985 by Academic Press. Reprinted by permission.

necting line ends in the degraded drawing restores the original drawing. In other cases, an unfamiliar drawing is generated by colinear completion. Biederman's results show that when illusory contours complete a familiar object, recognition is better than when illusory contours complete an unfamiliar figure. In the present theory, these boundary completion processes are assumed to occur within the BCS and the facilitation and interference effects on recognition are mediated by the BCS  $\leftrightarrow$  ORS interaction, even if the BCS  $\rightarrow$  FCS interaction does not lead to visibility of the illusory contours. Grossberg (1987a, Section 20) discusses how parametric properties of Biederman's data may be explained by BCS mechanisms.

The distinction between "recognition" by the ORS and "seeing" by the FCS is not, however, sufficient to explain why the occluded regions of a B, after their boundaries are completed, do not trigger filling-in of visible contrasts behind the black occluder. This property requires active explanation because such filling-in does sometimes occur, as during transparency phenomena (Beck, Prazdny, & Ivry, 1984; Metelli, 1974a, 1974b; Metelli, DaPos, & Cavedon, 1985; Meyer & Senecal, 1983). As noted in Section 2I, boundaries of a nearer surface are added to the boundaries of a farther surface within the FCS to prevent filling-in of the gray B color behind the black occluder. This prevents the black occluder from looking transparent. It does not explain, however, why some surfaces do look transparent. This property is explained below, as are other properties of transparency.

## 6. Occluded Boundary Completion and Recognition With Filling-In: 3-D Neon Color Spreading

An explanation of the connection between transparency, illusory contours, and neon color spreading was outlined in Grossberg (1987b, Section 21). This explanation has been further supported by experiments by Nakayama et al. (1990), who have studied a set of displays that seriously challenge traditional theories of stereopsis and 3-D vision. Their experiments analyzed how a change in the relative depth of occluding and occluded figures correlates with whether or not filling-in will occur within illusory boundaries, and thereby render them visible. One of their key demonstrations adapted the Ehrenstein pattern used by Redies and Spillmann (1981) to demonstrate neon color spreading in the 2-D case (Figure 9a). In this percept, the gray color in the cross fills-in a circular region that is defined by illusory contours which join the four white/gray edges at the ends of the cross. 2-D percepts of neon color spreading were explained using monocular BCS/FCS interactions in Grossberg and Mingolla (1985b) and Grossberg (1987b, Section 31). This explanation, in turn, influenced the Nakayama et al. experiments on 3-D neon color spreading. These 3-D percepts are explained herein using the theory's new computational rules. The main 3-D experiments are summarized below.

In order to study correlations between relative depth and neon color spreading, Nakayama et al. (1990) generated two sets of Redies-Spillmann images for use in a stereogram (Figure 9b). Each image consisted of a gray

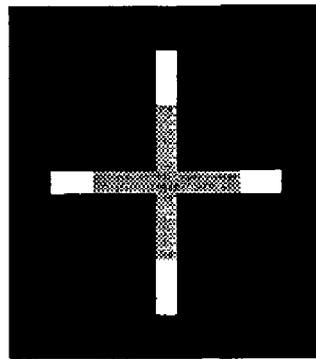
(or red) cross abutting colinear white bars on a black background. The two images differed only in the disparity of the horizontal inner limbs of the gray cross relative to the cross as a whole; that is, they differed in the relative positions of the vertical white/gray edges of the horizontal limbs in the two crosses. This simple manipulation had a dramatic effect on the percept.

When the vertical white/gray edges determine a nearer disparity than the rest of the image, they appear to be bounded by a transparent circular gray disk that connects them to the horizontal white/gray edges. In particular, a circular illusory contour bounds neon color spreading from the cross into the disk. The surface on which the cross lies can be seen through the transparent gray disk and appears to lie at a further distance. An observer can both recognize and see the circular illusory contour by virtue of the gray disk that it bounds. When the vertical white/gray edges determine a further disparity, the black surface appears to be opaque and to have a cross-shaped hole in it. The observer sees a cross-shaped gray region at a further distance within this hole, and recognizes this region to be part of a circular disk that is partly occluded by the opaque black surface. In this case, an observer can recognize the circular boundary of the disk, but can only see the unoccluded cross-shaped gray region as a surface. As in the discussion of Section 5, some completed boundary representations can be recognized but not seen.

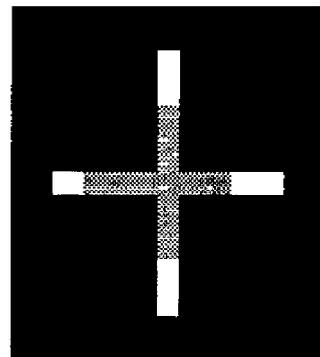
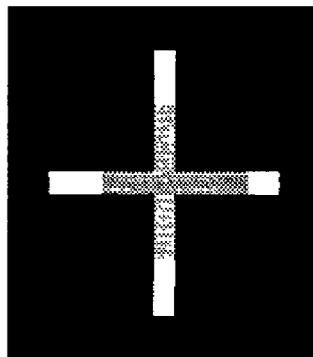
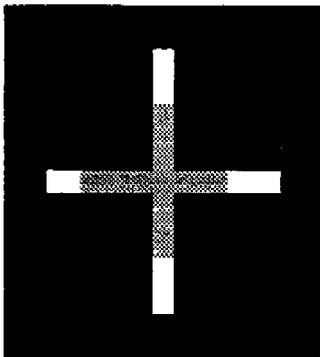
This experiment demonstrates some paradoxical properties of 3-D vision using particularly simple and evocative images. One important property is that a continuous surface percept may be generated by a sparse and discontinuous set of depth cues. In particular, the relative depth of the filled-in circular disk is determined by the binocular disparity of only two vertical white/gray edges relative to the binocular disparities of other image contours.

A second key property is that a continuous surface-in-depth can be generated by joining contours with different binocular disparities. The circular disk is bounded by vertical and horizontal white/gray edges with different disparities. In particular, the horizontal white/gray edges have zero binocular disparity. The theory explains how this happens using the same mechanisms as those in the discussion of Da Vinci stereopsis in Section 2E and Figure 3. Boundaries corresponding to nonzero disparities (the vertical white/gray edges) are added to boundaries corresponding to near-zero disparities (the horizontal white/gray edges) before they cooperate to initiate boundary completion of a circular, and thus connected, boundary. This circular boundary contour contains the filling-in of gray color from the cross.

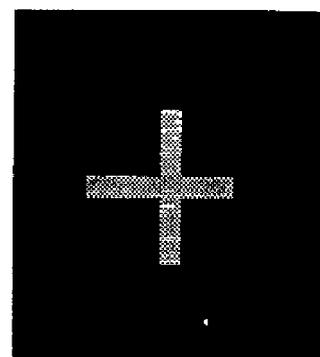
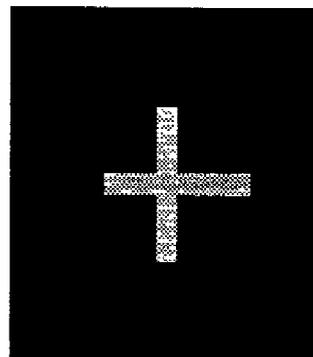
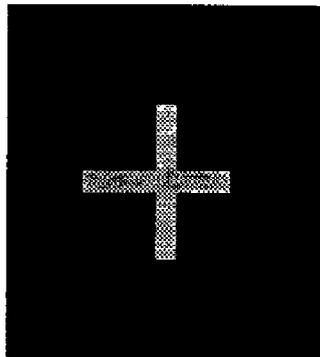
A third property concerns the asymmetric percepts of near and far objects. When the circular illusory boundary lies in front of the boundaries that define the background, gray color from the cross can fill-in the circle without obstruction. When the circular illusory boundary lies behind the boundaries of the cross, the cross boundaries prevent gray color from spreading beyond the cross itself. The theory traces this asymmetry between near and far to the same mechanism that prevents the



(a)



(b)



(c)

Figure 9. (a) The Ehrenstein pattern studied by Redies and Spillmann (1981) can induce a percept of neon color spreading as part of a 2-D percept. (b) 3-D version of the Redies and Spillmann display wherein the leftmost or the rightmost two images are viewed by the left and right eyes in a stereoscope. (c) When the white limbs of the cross are removed, totally different 3-D percepts are seen. See text for details. From "Transparency: Relation to depth, subjective contours, luminance, and neon color spreading" by K. Nakayama, S. Shimojo, and V. S. Ramachandran, 1990, *Perception*, 19, 497-513. Copyright 1990 by Pion Ltd. Reprinted by permission.

nearer surface AB in the Da Vinci stereopsis display of Figure 2 from appearing transparent due to filling-in of the farther surface, BD, behind AB. Boundaries that correspond to nearer objects—in particular objects with larger disparities—add to the boundaries that correspond to farther objects—in particular objects with smaller disparities—to prevent all nearer surfaces from looking transparent, as in Figure 3b. This mechanism is called *BF intercopies* because multiple BCS boundary representations, or copies, each corresponding to a different disparity range, converge on FCS filling-in domains in a partially ordered manner. Boundary segmentations that correspond to a given depth hereby obstruct the filling-in of surface representations that correspond to that depth and all farther depths.

A similar addition of boundaries from near to far surfaces was invoked in Section 5 to explain why the gray Bregman-Kanizsa B shapes do not fill-in behind their black occluders. In this percept, the edges of the occluder and the B shapes do not lie at different depths from the observer. One of the achievements of the present theory is to explain how this can happen in response to a 2-D picture as an epiphenomenon of the computations needed to fill-in perceptually appropriate 3-D surfaces when disparity cues are available from a 3-D scene.

In an interesting variation of the 3-D demonstration, the white outer limbs of the cross were removed (Figure 9c). When the horizontal limb of the gray cross possessed the crossed disparity relative to the vertical limb that led to a transparent disk percept in Figure 9b, a vertical bar was perceived in back of a horizontal bar. Moreover, the subjective contours that completed the horizontal bar were readily visible. The striking differences between the percepts elicited by the corresponding images in Figures 9b and 9c are consistent with early observations of Wallach and his colleagues that the “configuration of the two monocular patterns,” and not merely binocular disparities, helps to determine the final depth percept (Wallach & Bacon, 1976; Wallach & Barton, 1975; Wallach & Lindauer, 1962). How such configurational properties interact with binocular disparities to generate 3-D surface percepts is analyzed below.

Nakayama et al. (1990) drew some conclusions from these percepts that are clarified and modified by this analysis. They claimed that “transparency mechanisms can suppress subjective contours.” Instead, I will analyze how the “configuration of the two monocular patterns” can lead to different combinations of boundary completion and filling-in, whether “real” or “illusory,” to form consistent 3-D surface percepts. In particular, when one perceives a transparent disk-like object in front of the inner cross, this percept of transparency does not “suppress subjective contours.” It is true that the horizontal boundaries elicited by the cross stereogram in Figure 9c are not visible in Figure 9b. On the other hand, the percept of transparency that is elicited by the Redies-Spillmann stereogram in Figure 9b is largely *caused* by the subjective contour that encloses the disk. It is suggested below how this subjective contour helps to inhibit the boundary sig-

nals generated by the gray cross where it abuts the subjective contour, thereby enabling gray color to flow out of the gray cross and to fill-in the transparent surface that is seen within the subjective contour.

### 7. 3-D Kanizsa-Varin Percepts

Another interesting 3-D variation of a familiar 2-D percept uses the Kanizsa square. Here the disparity of the vertical boundaries in the “pacman” figures of the two images was varied (Figure 10). In the crossed-disparity case, which corresponds to closer objects, the familiar subjective contours that frame the square were seen as greatly enhanced and the Kanizsa square was seen in front of four disks which were partially occluded by the opaque Kanizsa square. An observer can recognize that the pacman boundaries were completed into disks behind the square surface, even though only the pacman regions were seen as visible surfaces (Figure 10a). When the disparity was reversed, an opaque surface was perceived through whose four (almost) circular windows were seen the four corners of an occluded square (Figure 10b). In this percept, the illusory contours that completed the four circular windows were rendered visible by filling-in the occluding surface at the nearer depth. The Kanizsa square was recognized behind the occluding surface, but only its four pacman regions could be seen as visible surfaces through the four circular windows.

Concerning the latter percept, Nakayama et al. (1990) stated that “the vertical and horizontal contours are abolished. . . . Instead, the configuration may be seen as four circular windows through which one sees the four corners of occluded square” (p. 499). How can an observer be aware of an occluded square whose vertical and horizontal contours have been “abolished”? I explain the latter percept by showing how an illusory square is generated in back of the occluding surface without triggering the filling-in that could render it visible. My explanation of the former percept suggests how an illusory square is

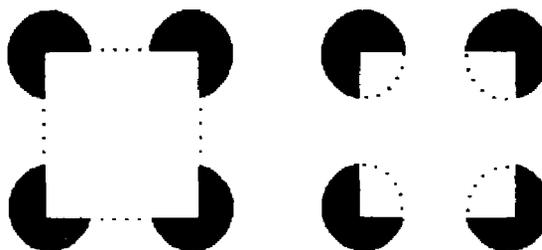


Figure 10. Differing percepts of a Kanizsa square when it is seen (a) in front of or (b) behind the four pacman figures. It appears either (a) as an opaque square surface that partially occludes four circular black disks, or (b) as a partially occluded square behind an opaque surface with four circular apertures. From “Transparency: Relation to depth, subjective contours, luminance, and neon color spreading” by K. Nakayama, S. Shimojo, and V. S. Ramachandran, 1990, *Perception*, 19, 497-513. Copyright 1990 by Pion Ltd. Reprinted by permission.

seen in front because it generates visible filling-in. Thus my explanation depends upon the two theoretical hypotheses: the ORS may recognize, or be "aware" of, BCS boundaries that are not *seen* within the FCS, and hence may seem to be "abolished" (see Figure 7); and the BCS boundaries of nearer segmentations restrict the filling-in of farther surfaces (Figure 3b). The theory has also been used to analyze why a Kanizsa percept caused by a 2-D display often seems closer to the observer than its bounding image (Grossberg, 1993).

Nakayama et al. (1990) also studied a 3-D version of the Varin (1971) display in which the pacman "mouths" of the Kanizsa square were filled with a gray (or colored) sector that completes them into disks. When the disparity was crossed, transparency and neon spreading of color into the square region was apparent. With reversed disparity, transparency was absent and little or no neon spreading was visible. Instead, the square was recognized as an occluded surface behind an opaque occluding surface with four circular apertures through which the four gray wedge-shaped corners of the occluded square could be seen. Grossberg (1987c, Section 21) showed how, in the case of a 2-D Kanizsa-Varin display, the subjective square boundary could inhibit the boundary signals generated by the colored wedges where they abut the subjective square, thereby enabling color to flow out of the wedges and to fill-in a transparent surface within the square. This explanation noted the close relationship that exists between conditions that lead to neon color spreading and the Metelli (1974a, 1974b) relations for transparency. The present extension of the theory is used to explain how these mechanisms work in the 3-D case to generate the two different surface percepts described above.

A variation of the 3-D Varin display that is analogous to the difference between the images in Figures 9b and 9c was also studied. In this variation, the outer "pacmen" were removed, leaving only the inner Varin wedges. Instead of a percept being seen of a transparent and nearer surface over the region of the wedges, the wedge region was seen as being the most distant, viewed through the opening of a square aperture. Nakayama et al. noted that the removal of the transparent surface was related to this reversal of perceived depth. This result provides dramatic support for a key property of the 3-D theory in Grossberg (1987c) on which the present article builds, namely that selective filling-in within some FCS domains, but not others, subserves not only brightness and color percepts, but also other surface properties, such as depth and form. The processes of boundary completion and filling-in that are used to help explain this result below emphasize the profound difference that exists between the relatively early mechanisms of stereopsis and a fully elaborated percept of surface depth and form.

### 8. Surface Induction by Sparse Inducers

Another display that was analyzed by Nakayama et al. (1990) also highlights this difference. See their Figure 11 and related work by Takeichi et al. (1992). Two nearly

identical stereograms were studied. In both, a white rectangle was sprinkled with random dots, all of which were in the same disparity plane as the rectangle. In the middle of each white rectangle there was also a gray rectangle that had crossed disparity. The two stereograms differed only by the binocular disparity of a single white dot in the middle of the gray region. When the white dot had the disparity of the gray region, the gray region looked like an opaque surface. When the white dot had the disparity of the white rectangle, the gray region looked like a transparent surface. This striking difference can be explained in terms of the different patterns of filling-in that are triggered by the isolated dot in the two cases. In particular, suppose that the dot's boundaries have the disparity of the boundaries of the gray region. Then they trigger filling-in of both their gray exterior and their white interior within the filling-in domain corresponding to the closer disparity. When the dot's boundaries have the disparity of the white region, they trigger these filling-in events within the filling-in domain corresponding to the farther disparity. In both cases, the boundary of the gray region triggers filling-in of its interior region within the filling-in domain of the closer disparity. The existence of two filling-in events over the same gray region within filling-in domains corresponding to different disparities contributes to the percept of transparency.

This explanation further develops the argument whereby the "Julesz 5% solution" stereogram (Julesz, 1971, p. 336) was explained in Grossberg (1987b, Section 14). In the Julesz stereogram, each image contains 5% black dots on 95% white background. The dots are divided into two subsets, one whose dots have a large disparity with respect to the image frames, and the other whose dots have zero disparity. In the binocular percept, the boundaries of dots with large disparities cause the black dots and white surround that they enclose to be seen as a planar surface lying in front of another planar surface that contains the zero disparity black dots and the white region that they enclose.

### 9. What is a Visual Illusion?

My explanation of these percepts does not support the claim of Nakayama, Shimojo, and Ramachandran (1990) that "color spreading cannot be a necessary condition for transparency because transparency occurred without color spreading" (p. 505). In contrast, I claim that "color spreading" or filling-in is controlled by the disparity of the dot, which generates a BCS boundary segmentation that triggers selective filling-in of an FCS surface representation at the corresponding relative depth from the observer. One does not "see" the color spreading in the same familiar sense that one does not "see" the filling-in that supports most surface percepts. The filling-in does occur. I claim, but it generates a familiar juxtaposition of BCS emergent segmentations and FCS filled-in surfaces. Only when the relationship between these BCS and FCS properties is unexpected, as during neon color spreading, do many people realize that they "see" the

color spreading. From this perspective, one can *define* a large class of visual illusions as *unexpected combinations of BCS emergent segmentations and FCS filled-in surfaces*.

### 10. Visual Phantoms for Brightness and Target Visibility

Another interesting probe of the relationship between illusory contours and filling-in during figure-ground separation makes use of displays that generate moving visual phantoms (Tynan & Sekuler, 1975). Moving visual phantoms involve the perception of contours and surfaces in a homogeneous region. Phantoms may, for example, be induced when a vertically oriented black-and-white square-wave grating moves horizontally above and below a black homogeneous horizontal strip. When phantoms are not visible, the homogeneous region appears as a uniform, opaque surface that occludes the inducing grating. When phantoms are visible, the apparent depth of the occluder changes and the black grating regions induce complete vertical stripes moving in front of the black horizontal strip. The phantom regions also appear darker than the nonphantom regions (Brown & Weisstein, 1988a). If a gray-and-black square-wave grating is used to induce phantoms across a gray horizontal strip, then vertical gray stripes are completed in front of the gray horizontal strip and look lighter than the horizontal strip (Brown & Weisstein, 1988a). These experiments of Brown and Weisstein are consistent with previous studies showing that the perceived brightness of a region is influenced by its perception as figure or ground (Rubin, 1921, 1958). These phantom results may be explained by using the same mechanisms that are used below to explain why we tend to see the higher spatial frequencies pop out as a figure in the Weisstein effect (Figure 5). The brightness assimilation effects may be explained by using the same mechanisms that are used below to explain 3-D neon color spreading and transparency (Section 6).

Brown and Weisstein (1988a) also showed that discrimination of the orientation of a target line was better when the line fell within a phantom region than when it fell within a nonphantom part of the occluding horizontal strip. This result extended earlier results of Wong and Weisstein (1982, 1983), which reported superior orientation discrimination within regions that are perceived to be figures than within regions that are perceived to be ground. In these experiments, the Rubin (1921, 1958) faces/vase display was used. A target in the middle of the picture was discriminated better when it was perceived as a vase than when it was perceived as the background of the faces. Wong and Weisstein also observed that sharp targets are detected better against a figure, and blurred targets are detected better against a background. These results are explained below using the same mechanisms as in the explanation of the Weisstein effect.

## PART II

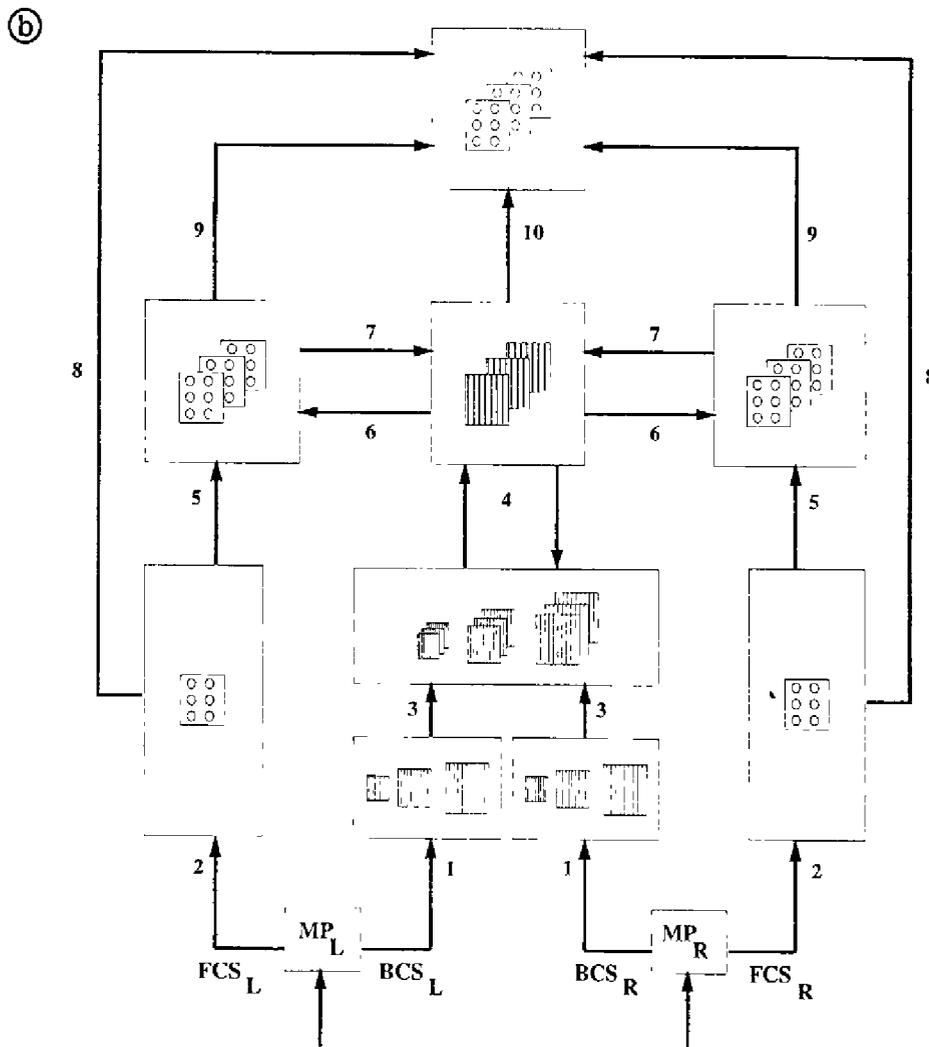
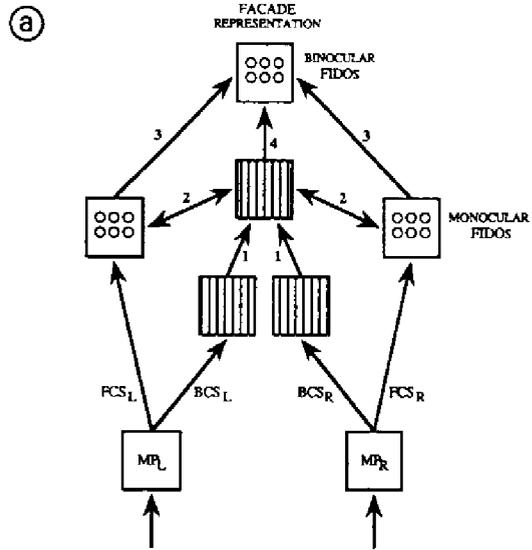
### 11. Introduction to FACADE Theory

Perceptual properties such as those summarized above illustrate how 3-D segmentations and surface representa-

tions are formed, and how visual figures pop out from other figures and their backgrounds. The theory that is now described provides a unified explanation of these and other percepts. It develops an earlier theory of 3-D pre-attentive vision that was introduced in Grossberg (1987b, 1987c). This theory has been called the FACADE theory because it suggests how visual representations of Form-And-Color-And-DEpth, or FACADES, are generated in area V4 of the prestriate visual cortex (Figure 11a). FACADE theory describes the neural architecture of two parallel subsystems, the boundary contour system (BCS) and the feature contour system (FCS). The BCS generates an emergent 3-D boundary segmentation of edges, texture, shading, and stereo information at multiple spatial scales (Carpenter, Grossberg, & Mehanian, 1989; Cruthirds, Gove, Grossberg, & Mingolla, 1991; Grossberg, 1987b, 1987c, 1990; Grossberg & Marshall, 1989; Grossberg & Mingolla, 1985a, 1985b, 1987a, 1987b; Grossberg & Somers, 1991, 1992). The FCS compensates for variable illumination conditions and fills-in surface properties of brightness, color, and depth among multiple spatial scales (Cohen & Grossberg, 1984; Grossberg, 1987b, 1987c;

---

Figure 11. (Facing page) (a) Macrocircuit of monocular and binocular interactions of the boundary contour system (BCS) and the feature contour system (FCS) drawn to facilitate comparison with Figure 1 in Grossberg (1987c): Left eye and right eye monocular preprocessing stages (MPL and MPR) send parallel pathways to the BCS (boxes with vertical lines, designating oriented responses) and the FCS (boxes with three pairs of circles, designating opponent colors). The monocular signals BCS<sub>L</sub> and BCS<sub>R</sub> activate simple cells which, in turn, activate bottom-up pathways, labeled 1, to generate a binocular boundary segmentation using the complex, hypercomplex, and bipole cell interactions of Figure 12. The binocular segmentation generates output signals to the monocular filling-in domains, or FIDOS, of the FCS via pathways labeled 2. This interaction selects binocularly consistent FCS signals, and suppresses the binocularly inconsistent FCS signals. Reciprocal FCS → BCS interactions enhance consistent boundaries and suppress boundaries corresponding to further surfaces. The surviving FCS signals activate the binocular FIDOS via pathways 3, where they interact with an augmented binocular BCS segmentation to fill-in a multiple-scale surface representation of Form-And-Color-And-DEpth, or FACADE. Processing stages MPL and MPR are compared with LGN data; the simple-complex cell interaction with V1 data; the hypercomplex-bipole interaction with V2 and (possibly) V4 data, notably about interstripes; the monocular FCS interaction with blob and thin stripe data; and the FACADE representation with V4 data (see Figure 1). Additional interactions from FCS to BCS along pathways labeled 2, 3, and 4, and among FCS and BCS copies, are described in the text. (b) A finer representation of FACADE interactions: Output signals from MPL and MPR activate BCS simple cells with multiple receptive-field sizes via pathways 1. MPL and MPR outputs are also transformed into opponent FCS signals via pathways 2. Pathways 3 generate multiple cell pools that are sensitive to multiple disparities and scales. BB intrascales are at work among the resultant cells, as are the first and second competitive stages. Pathways 4 combine the multiple scales that correspond to the same depth range into a single BCS copy via BB interscales. Multiple copies that correspond to different (but possibly overlapping) depth ranges exist. Pathways 5 are the monocular FF intercopies. Pathways 6 are the BF intercopies. Pathways 7 are the FB intercopies. Pathways 8 are the excitatory binocular FF intercopies. Pathways 9 are the inhibitory binocular FF intercopies. Pathways 10 are the BF intercopies. See text for further details.



Grossberg & Mingolla, 1985a; Grossberg & Todorović, 1988; Grossberg & Wyse, 1991, 1992).

The BCS has been used to analyze and predict neurobiological data concerning the parvocellular processing stream from the lateral geniculate nuclei (LGN) through cortical area V4 via the interblob and interstripe networks of cortical areas V1 and V2, respectively (see Figure 1). The FCS has been used to analyze and predict data concerning the parvocellular processing stream from the LGN through cortical area V4 via the blob and thin stripe networks of V1 and V2 (Figure 1). Interactions between the BCS and FCS give rise to FACADE representations that are predicted to occur in area V4. In vivo, these cortical processing streams multiplex combinations of position, orientation, size, disparity, and color selectivity (Figure 1) which are clarified by BCS and FCS computational properties. Remarkably, BCS and FCS properties are computationally complementary (Grossberg, Mingolla, & Todorović, 1989), a fact which suggests that the two cortical streams are intimately related, rather than comprising independent modules, and may arise through a process of global symmetry-breaking during morphogenesis. The magnocellular processing stream from LGN to cortical area MT via lamina 4B and thick stripe networks of cortical areas V1 and V2 (Figure 1) are analyzed elsewhere in terms of a motion BCS (Grossberg & Mingolla, 1993; Grossberg & Rudd, 1989, 1992). To distinguish the BCS discussed here from the motion BCS, it will be called the static BCS. The motion BCS is not the focus of the present article, but its relevance to attentional processing will be discussed in Section 49.

Many experimental and modeling studies published subsequent to the original BCS and FCS articles provide further support for BCS and FCS properties. These include studies of texture segregation (Beck, Graham, & Sutter, 1991; Beck, Rosenfeld, & Ivry, 1990; Cruthirds, Grossberg, & Mingolla, 1993; Graham, Beck, & Sutter, 1992; Sutter, Beck, & Graham, 1989), border effects on color detection (Eskew, 1989; Eskew, Stromeyer, Picotte, & Kronauer, 1991), visual phantoms (Brown & Weisstein, 1988a), 3-D surface formation from 2-D textures (Buckley et al., 1989; Todd & Akerstrom, 1987), interactions between filling-in of brightness or color and illusory contour formation (Dresp, Lorenceau, & Bonnet, 1990; Field, Hayes, & Hess, 1993; Kellman & Shipley, 1991; Leshner & Mingolla, 1993; Nakayama, Shimojo, & Ramachandran, 1989; Prinzmetal, 1990; Prinzmetal & Boaz, 1989; Ramachandran, 1992; Shipley & Kellman, 1992; Takeichi, Shimojo, & Watanabe, in press; Watanabe & Sato, 1989; Watanabe & Takeichi, 1990), interactions between depth, emergent segmentation, and filling-in (Meyer & Dougherty, 1987; Nakayama, Shimojo, & Ramachandran, 1990; Takeichi, Watanabe, & Shimojo, 1992; Watanabe & Cavanagh, 1992), orientation-specific luminance aftereffects (Mikaelian, Linton, & Phillips, 1990), transient dynamics of filling-in (Arrington, 1992, 1993; Paradiso & Nakayama, 1991), cortical dynamics of emergent segmentation (Peterhans & von der Heydt, 1989; von der Heydt, Peterhans, & Baumgartner, 1984), and grouping processes during visual search (Humphreys, Quinlan, & Riddoch, 1989).

In its original form, FACADE theory did not posit interactions between the different spatial scales of the BCS and the FCS, or from the FCS to the BCS. Such interactions were not needed to explain the data analyzed in previous articles. The present work shows how suitably defined interactions within and between BCS and FCS scales lead to explanations of a much wider body of perceptual and neural data about 3-D visual perception. These interactions are consistent with the previous theory and build upon it. Several investigators (e.g., Tolhurst, 1972; Watt, 1987; and Wilson, Blake, & Halpern, 1991) have described experimental evidence for the existence of interactions between scales. The present theory proposes interscale interactions that clarify the data which led to these proposals, but uses interactions which have not previously been described because their functional role depends upon BCS and FCS mechanisms for their description.

These interactions constitute a set of simple computational rules that are carried out in a prescribed order. Different sets of experiments lend greater support to some rules than they do to others. Removal of any rule prevents the theory from explaining a subset of these data. Taken together, the rules as a whole are supported by a large body of perceptual data that have as yet received no other explanation. In addition, the neural interpretation of these rules leads to a series of testable neurobiological predictions concerning the types and ordering of interactions that occur within and between the two parvocellular cortical processing streams. Although the theory cannot predict unequivocally the processing stages at which such rules may be instantiated in different mammals, it can and does suggest the earliest stages that are consistent with known data, and the ordering of stages within which the rules should be realized. These earliest possible stages are used in the neural predictions described herein.

It needs also to be emphasized that these new rules do not merely add unconstrained degrees of freedom to an already complex theory. Rather, each stage of the previous theory has been supported by evidence from a variety of perceptual and neural experiments, and no stage of the enhanced theory can be removed without undermining the theory's explanation of many additional experiments. Every stage and interaction of the theory also has a clear functional meaning that may be used to guide the design of additional experimental tests. Although the theory is more complex than less ambitious models, its complexity does not exceed that of the parvocellular cortical streams which it sets out to explain.

In previous articles, the static BCS was used to suggest a new computational model and rationale for the neural circuits governing classical cortical cell types such as simple cells, complex cells, and hypercomplex cells in cortical areas V1 and V2 (Figure 12). Functional roles for additional cell properties, such as end-stopped simple cells (Grossberg & Mingolla, 1993) and reciprocal top-down pathways (Grossberg, 1980), have been described, but are not needed to explain the data discussed herein. The theory also predicted a new cell type, the bipole cell (Cohen & Grossberg, 1984; Grossberg, 1984; Grossberg & Mingolla, 1985a, 1985b), whose properties

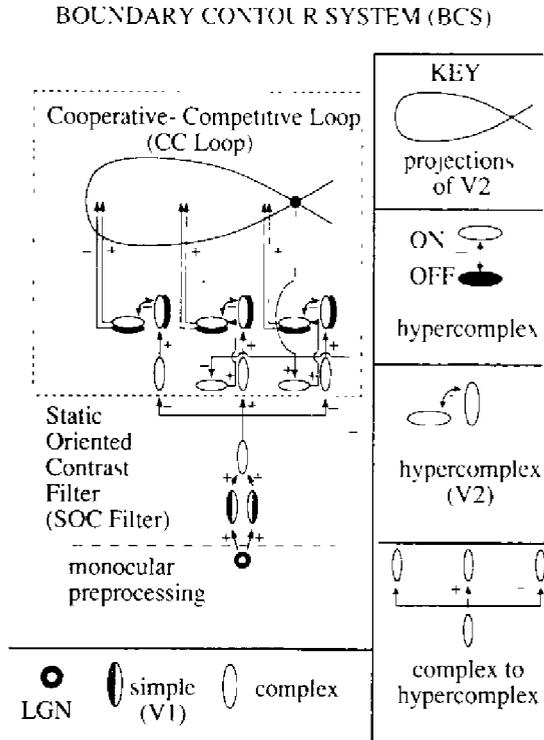


Figure 12. The monocular boundary contour system of Grossberg and Mingolla (1985b). The circuit is divided into a static oriented contrast-sensitive filter (SOC filter) followed by a cooperative-competitive feedback network (CC loop). Multiple copies of the circuit are used, each corresponding to a different range of receptive field sizes. Each copy models interactions of simple cells, complex cells, hypercomplex cells, and bipole cells.

have been supported by subsequent neurophysiological experiments (Peterhans & von der Heydt, 1989; von der Heydt et al., 1984). The interactions within the simple-complex-hypercomplex cell module define a static oriented contrast-sensitive filter, called the SOC filter. As discussed in Sections 12 and 13, this filter compensates for uncertainties of positional localization in the output of simple cells that are caused by their oriented receptive fields. It also generates output signals from the complex and hypercomplex cells that are independent of direction-of-contrast, even though simple-cell outputs are sensitive to direction-of-contrast. The interactions between bipole cells and the SOC filter define a cooperative-competitive feedback network, called the CC loop, which generates featural bindings, or emergent boundary segmentations, from combinations of edge, texture, shading, and stereo image properties. Consistent combinations of image data generate fused segmentations with coherent properties. Inconsistent combinations lead to suppression and rivalry. The FCS characterizes how on-cells and off-cells, interacting within shunting on-center off-surround networks, compensate for variable illumination. The output signals from these networks activate networks wherein electrotonically coupled cells diffusively fill-in representations of surface brightness, color, form, and depth within domains defined by BCS boundary signals (Figure 13).

This architecture's emergent properties have led to a unified explanation and predictions of a data base about psychophysics, visual perception, and cortical organization that has not yet been explained by any alternative perceptual theory. To achieve this explanatory range, a qualitatively new computational vision theory needed to be developed. Thus, this neural architecture is not merely a more efficient way to represent prior vision theories.

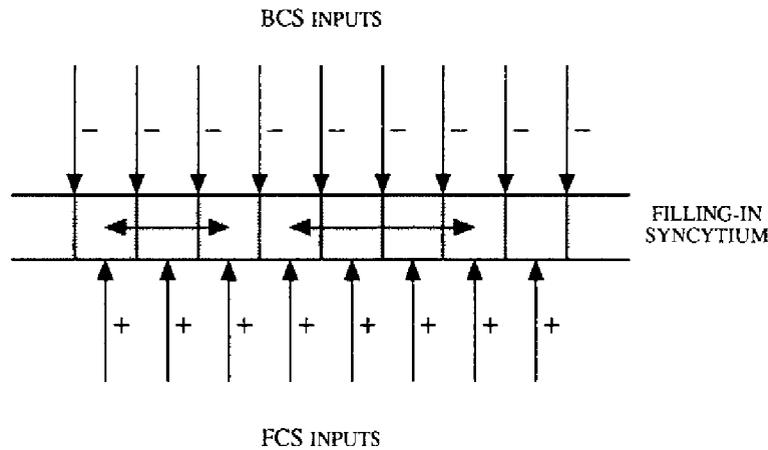


Figure 13. A monocular syncytium within the feature contour system (FCS). The feature contour signals are output signals from a shunting on-center off-surround network that discounts the illuminant. These signals activate cells that permit rapid electrotonic diffusion of activity, or potential, across their cell membranes, except at the membranes that receive boundary contour signals. The gap junctions at these membranes respond to the BCS signals with an increase in resistance that decreases diffusion across them. Thus, FCS signals rapidly fill-in across syncytium cells until they reach a BCS boundary or are attenuated by their spatial spread.

Rather, it articulates several basic new uncertainty principles and introduces a neural architecture that resolves these uncertainties through its parallel and hierarchical interactions.

### 12. The Heterarchical Resolution of Complementary Uncertainties

These uncertainty principles were discovered through a sustained analysis of the sensory uptake process. Such an analysis shows that there exist fundamental limitations of the visual measurement process at each stage of neural processing. When a neural processing stage eliminates one type of uncertainty in the input patterns that it receives, it often generates a new type of uncertainty in the outputs which it passes along to the next processing stage. Uncertainties beget uncertainties. It is not true that informational uncertainty is progressively reduced by every stage of neural processing.

Thus, low-level computations within the FCS reduce uncertainty due to variable illumination conditions, but create new uncertainties about surface brightnesses and colors. These uncertainties are resolved at a higher FCS level by a process of surface filling-in. Likewise, low-level computations within the BCS reduce uncertainty about boundary orientation, but create new uncertainties about boundary position that are resolved at a higher BCS level by a process of boundary completion. *FACADE* theory describes how the visual system as a whole can compensate for such uncertainties using both parallel and hierarchical stages of neural processing. Thus, the visual system is designed to achieve *heterarchical compensation for uncertainties of measurement*.

### 13. Emergent Segmentation: Boundary Completion and Regularization in High Image Noise

BCS operations occur automatically and without learning or explicit knowledge of input environments. A perceptual process is said to be *preattentive* if it occurs rapidly and automatically without recourse to stored templates or learned expectations. Thus, the emergent segmentations generated by the model are not the result of training on image exemplars. Nor do the equations embody a priori assumptions about such variables as direction of illumination or the shapes of objects to be encountered. Instead, the model embodies a number of circuits specialized to perform emergent, context-sensitive segmentations of a wide variety of images. By *emergent* segmentation is meant a partition of an image into boundary structures that may have no direct corollary in differences in gray level of the image itself (Figure 8). Boundaries perceived in this way are often referred to as "illusory" when seen by humans. The importance of one part of an image is evaluated by what surrounds that part of the image. Image contrasts that may represent noise in one image context may represent features in another. Every scenic input provides its own context, which the BCS uses to organize local contrasts. The proper weighting of signal against context is achieved by using properties of *self-scaling* and

*self-similarity* at several levels of the BCS network. Moreover, the BCS is sufficiently flexible, where necessary, to maintain several potential groupings simultaneously, and sufficiently rapid, when realized in hardware, to quickly converge on the most favored grouping for a given visual scene. *Regularization* refers to the smooth completion of structure at a given scale despite noisy disruptions of the signal at a smaller scale.

Figure 14 illustrates the BCS's ability to detect, complete, and regularize sharp boundaries over long distances in the presence of severe noise, a type of capability useful in penetrating camouflage. Smaller spatial scales would generate the boundaries of individual black and white compartments. The BCS needs no external temperature parameter or a priori cost function to control this self-scaling segmentation process, as in simulated annealing or other region-growing techniques. Instead, real-time cooperative-competitive nonlinear feedback interactions in the CC loop regulate a rapid convergence to an equilibrium configuration that automatically self-calibrates its criteria for grouping and segmentation according to the distribution of signal and noise in a particular image.

Figure 14 showed only the ability of the BCS to complete colinear boundaries in noise. In more complex examples of textural segregation, feature linking may be colinear, perpendicular, or diagonal to the orientations of the inducing image edges (Beck, Prazdny, & Rosenfeld, 1983). Figure 15 shows a simulation of this BCS competence. Figure 16 illustrates the BCS's capability for multiple scale segmentation. Figure 16a shows a curved textured surface. The equilibrated CC loop outputs (Figure 16b) are not simply filterings of Figure 16a contrasts. They detect the *coherence* of oriented contrasts at a given spatial scale in the form of a dense *boundary web* of oriented activity (Grossberg, 1987b; Grossberg & Mingolla, 1987a). In Figure 16, the CC loop tracks the coarse circular oriented statistics of the image. Such a boundary web forms a mesh of small compartments in the FCS to which filling-in is restricted, thereby giving rise to a smoothly shaded representation. In a multiple-scale version of the BCS, multiple populations of simple cells with different receptive field sizes input to distinct SOC filters with their own CC loops, each of which generates a different boundary web in response to the image. The totality of these multiple-scale boundary webs embodies a code for 3-D surface form. The worst correlation between human psychophysical judgments of 3-D shape-from-texture and theoretical BCS predictions based upon images such as in Figure 16a was .985 (Todd & Akerstrom, 1987).

### 14. Filling-In and Brightness Perception

A large number of paradoxical brightness and color percepts have been explained using interactions between BCS segmentations and FCS filling-in. Figure 17 summarizes a simulation of several basic brightness properties: discounting the illuminant and brightness constancy (Figures 17a and 17b), brightness contrast (Figure 17c), and brightness assimilation (Figure 17d).

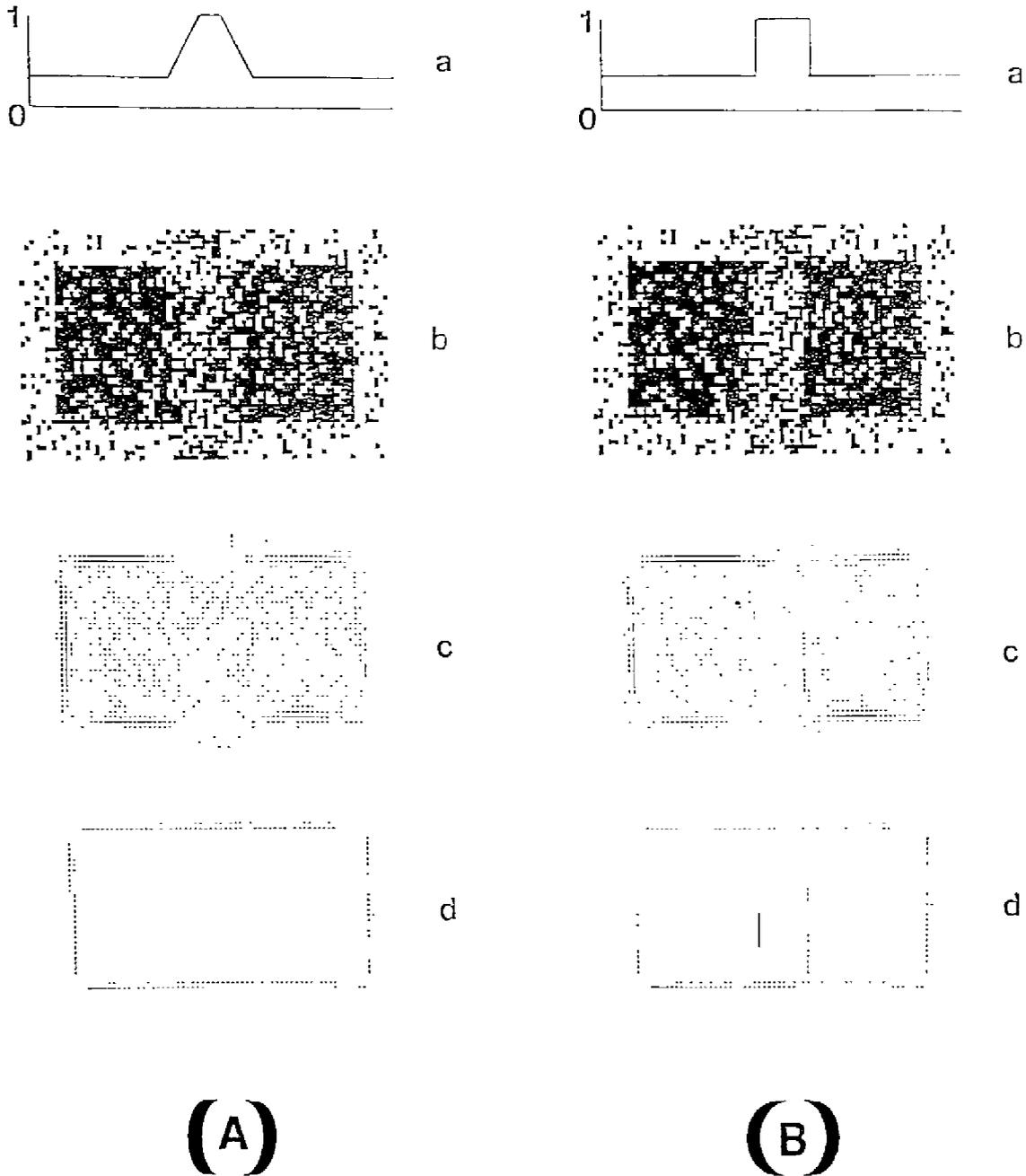


Figure 14. (A) (a) Distribution of noise in horizontal dimension of image; (b) binary image of a rectangle corrupted by noise whose distribution, as in (a), varies continuously; (c) responses of oriented contrast detectors to the image; (d) equilibrated responses of cooperative feedback cells of the boundary contour system (BCS). The rectangle is recovered and the ramped increase of noise in the middle of the figure is ignored. (B) (a) Distribution of noise in horizontal dimension of image; (b) binary image of a rectangle corrupted by noise whose distribution, as in (a), varies abruptly; (c) responses of oriented contrast detectors to the image; (d) equilibrated responses of cooperative feedback cells of the BCS. The rectangle is recovered and the abrupt increase of noise in the middle of the figure supports a vertical boundary segmentation.

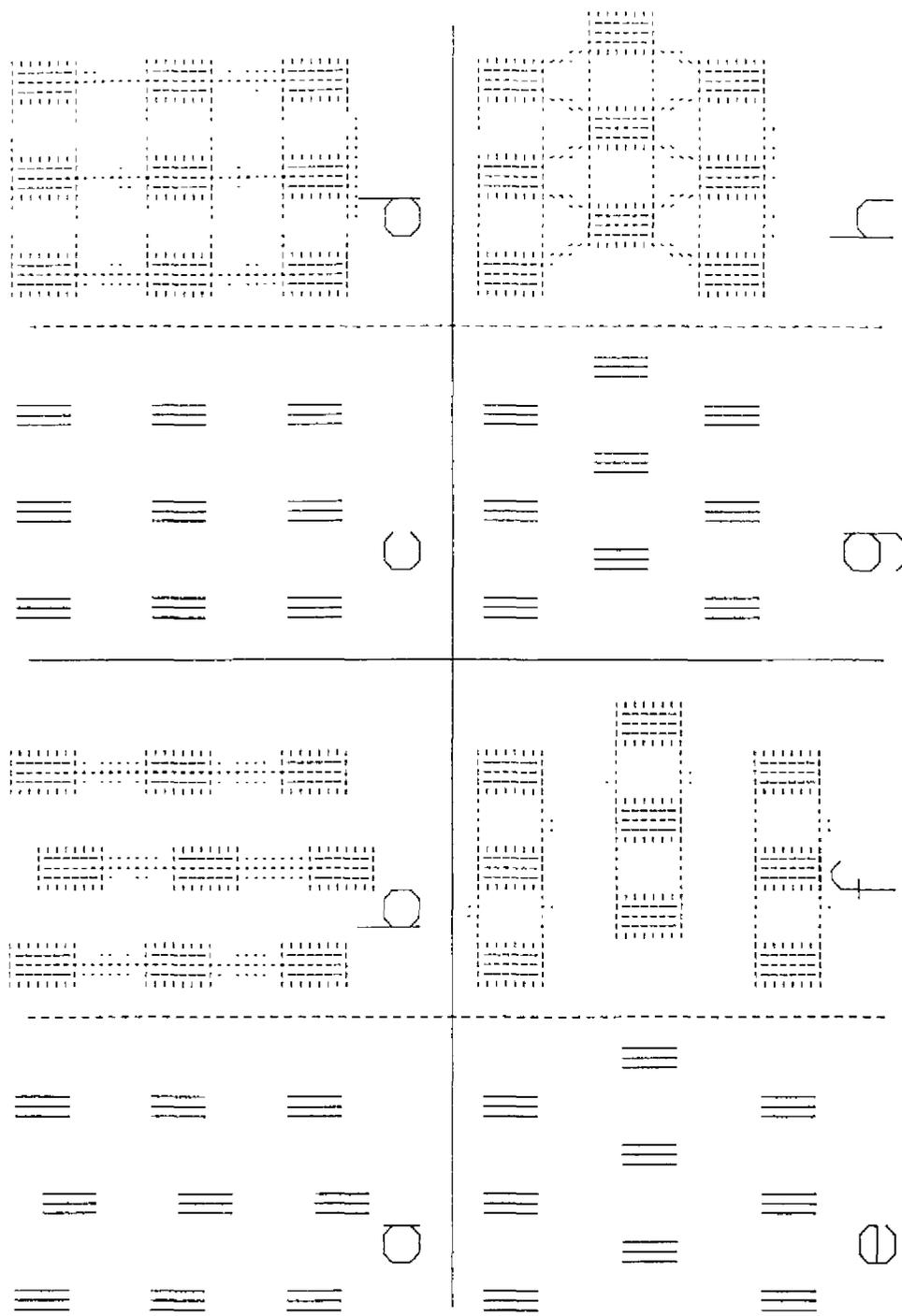
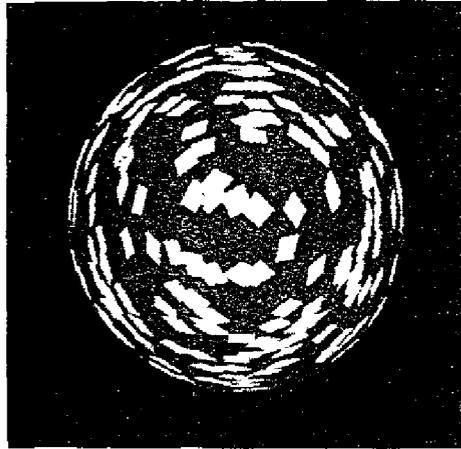


Figure 15. Computer simulations of processes underlying textural grouping. The length of each line segment is proportional to the activation of network node responsive to 1 of 12 possible orientations. Parts (a), (c), (e), and (g) display the activities of oriented cells which input to the cooperative-competitive (CC) loop. Parts b, d, f, and h display equilibrium activities of oriented cells at the second competitive stage of the CC loop. A pairwise comparison of (a) with (b), (c) with (d), and so on, indicates the major groupings sensed by the network. These simulations demonstrate that an emergent segmentation can form colinear to [as in (b) and (d)], perpendicular to [as in (d), (f), and (h)], or diagonal to [as in (b)], the inducing vertically oriented lines by merely changing the relative positions of these lines. See text for details. From "Neural dynamics of surface perception: Boundary webs, illuminants, and shape-from-shading" by S. Grossberg and E. Mingolla, 1987, *Computer Vision, Graphics, & Image Processing*, 37, 116-165. Copyright 1987 by Academic Press. Reprinted by permission.

(a)



(b)

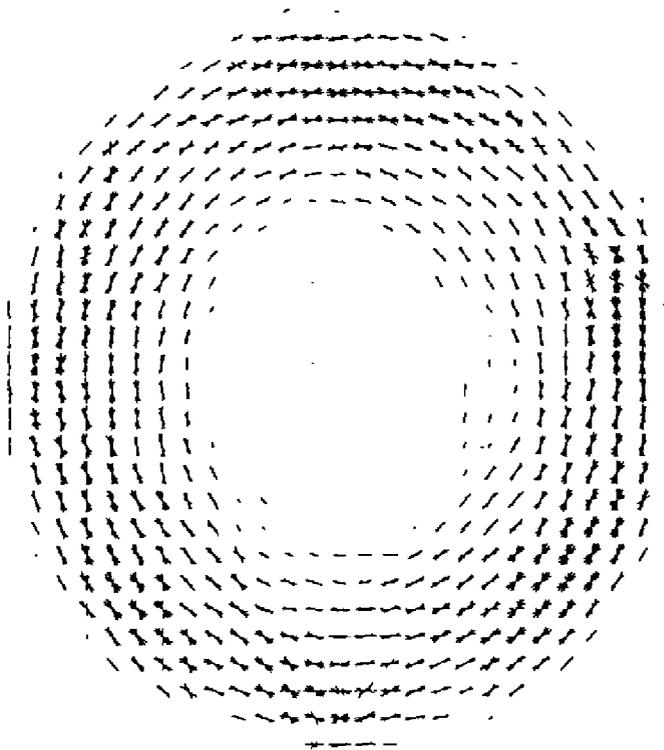
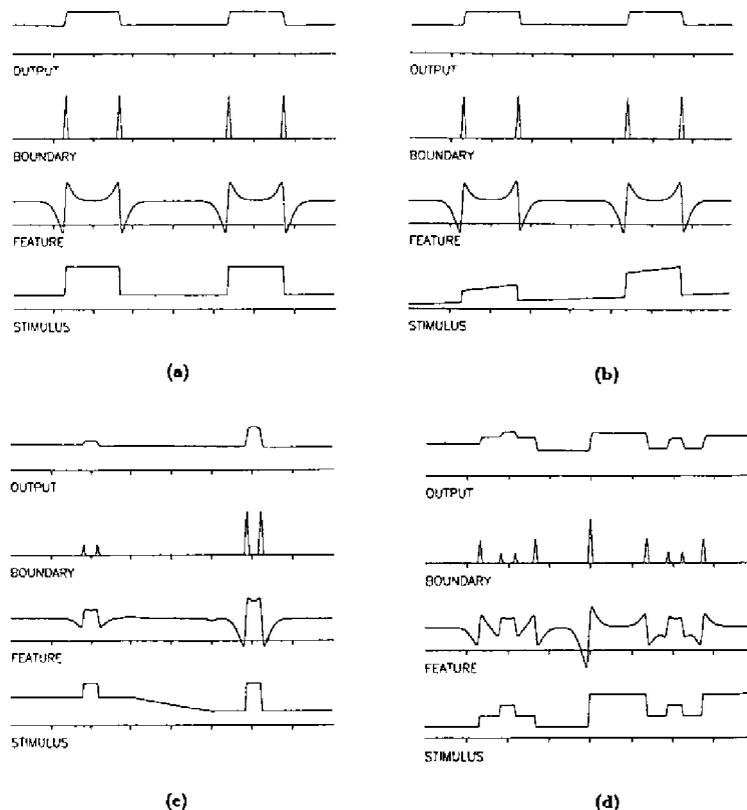


Figure 16. (a) A textured, curved surface. From "Perception of three-dimensional form from patterns of optical texture" by J. T. Todd and R. A. Akerstrom, 1987, *Journal of Experimental Psychology: Human Perception & Performance*, 13, 242-255. Copyright 1987 by the American Psychological Association. Adapted with permission. (b) Boundary web response of cooperative feedback cells of the CC loop to the inputs from Figure 16a. From "Neural dynamics of surface perception: Boundary webs, illuminants, and shape-from-shading" by S. Grossberg and E. Mingolla, 1987, *Computer Vision, Graphics, & Image Processing*, 37, 116-165. Copyright 1987 by Academic Press. Reprinted by permission.



**Figure 17.** Simulations of feature contour system (FCS) interactions in response to images with a one-dimensional symmetry: the luminance profile (STIMULUS) in (b) is tilted with respect to that in (a) due to an asymmetric light source, but the filled-in percept (OUTPUT) is the same as that in (a), illustrating a discounting of the illuminant and brightness constancy. Although the small patches have equal luminance in (c), their filled-in percepts are different, in the direction opposite to that of their backgrounds, illustrating brightness contrast. Although the small inner patches have equal luminance in (d), the filled-in percept of the right patch is darker than that of the left patch, in the direction of their surrounding patches, thus illustrating brightness assimilation. (Reprinted from Grossberg & Todorović, 1988.)

The following properties are relevant to our present expository needs. The feature contour pattern (denoted FEATURE in Figure 17) distorts the luminance pattern of the image (denoted STIMULUS in Figure 17) in order to discount the illuminant, as in Figures 17a and 17b. This is true because the off-surround in an on-center off-surround network that discounts the illuminant must decrease across space faster than the light source gradient that it is discounting. The distorted FEATURE pattern is used to transform the luminance pattern (STIMULUS) into a brightness pattern (OUTPUT). Boundaries (denoted BOUNDARY) are computed from the contrastive cusps that are formed in the FEATURE pattern as a result of the discounting process. These boundaries define syncytia, or Filling-In Domains, wherein FEATURE-induced activations diffuse laterally until they hit boundary obstructions.

A major conclusion from these results is that the visual system compensates for the distortions caused by discounting the illuminant by filling-in a surface representation

at a later processing stage. This surface uses the discounted FEATURE pattern as inputs. Sometimes the filling-in process succeeds in recovering a brightness representation that veridically represents the reflectances of the luminance patterns, as in brightness constancy (Figure 17b). At other times, new distortions are caused, such as in brightness contrast (Figure 17c) or, its apparent opposite, brightness assimilation (Figure 17d). The Craik-O'Brien-Cornsweet effect and many other challenging brightness percepts have also been simulated (Cohen & Grossberg, 1984; Grossberg & Todorović, 1988). These simulations compared brightness data with steady-state network activations after filling-in equilibrated. Paradiso and Nakayama (1991) carried out masking experiments to test the physical reality of the filling-in process by trying to catch it "in the act." On the basis of these results, they expressed concern about the ability of the Grossberg and Todorović (1988) model to explain transient masking data. Arrington (1992, 1993) carried out simulations of the

transient dynamics of the Grossberg and Todorović (1988) model in the Paradiso and Nakayama (1991) conditions. He demonstrated that the model “is very good at predicting the brightness percepts in a wide variety of masking experiments reported by Paradiso and Nakayama (1991).”

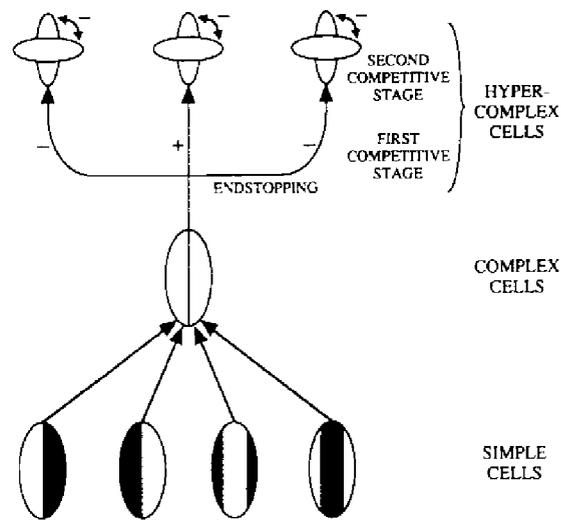
BCS/FCS interactions have been used to process complex images derived from artificial sensors—such as laser radars, synthetic aperture radars, magnetic resonance brain images, infrared sensors, and high-altitude photographs—for which it is difficult, if not impossible, to control detection conditions or to define rules that characterize image properties (Cruthirds et al., 1992; Lehar, Howells, & Smotroff, 1990; Lehar, Worth, & Kennedy, 1990; van Allen & Kolodzy, 1987; Worth, Lehar, & Kennedy, 1992). The Grossberg and Todorović (1988) model has been realized as a VLSI chip by Andreou and Boahken (1991). Their chip design uses the retinal interpretation that was described in Grossberg (1987b, Section 25), which compared cortical filling-in data with the retinal data of Piccolino, Neyton, and Gerschenfeld (1984) concerning the dynamics of H1 horizontal cells.

### 15. End Cuts Overcome Positional Uncertainties Caused by Oriented Receptive Fields

We are now ready to review the FACADE theory concepts that are needed to extend the theory to explain 3-D data, such as those summarized in Sections 2–10. First, the concept of *end cut* is reviewed in order to explain how some boundaries get attached to figures and detached from backgrounds.

In order to effectively build up boundaries, the BCS must be able to determine the orientation of a boundary at every position. The simple cells at the first stage of the BCS thus possess orientationally tuned receptive fields, or oriented masks (*Figure 18*). These oriented receptive fields are oriented *local contrast* detectors that fire in response to properly oriented edges, textures, and shading. Orientationally tuned simple cells that are sensitive to different bands of oriented contrasts respond to each small region of the scene, as in the hypercolumn model of Hubel and Wiesel (1977).

The fact that simple-cell receptive fields are oriented reduces the number of possible groupings to which they can respond. However, receptive field elongation also creates uncertainty about the positions with respect to the receptive field of the image contrasts that fire the cell. This positional uncertainty becomes acute during the processing of image line ends, corners, and other contours that change rapidly across space with respect to receptive field size. In particular, oriented receptive fields cannot detect the ends of thin lines (Grossberg & Mingolla, 1985b) whose widths fall within a certain range, as illustrated in *Figure 19a*. This property illustrates a basic uncertainty principle which says: Orientational “certainty” implies positional “uncertainty” at the ends of scenic lines whose widths are neither too small nor too large with respect to the dimensions of the oriented receptive field. In the absence of subsequent processing within the BCS, the BCS boundary generated by such a line would contain



**Figure 18.** A simplified monocular model of the interactions that convert simple cells into complex cells and then into two successive levels of hypercomplex cells. The interactions (simple cell) → (complex cell) and (complex cell) → (hypercomplex cell) describe two successive spatial filters, which together are called the static oriented contrast-sensitive (SOC) filter. Simple cells form one filter. Their rectified outputs combine as inputs to complex cells. A second filter is created by the on-center off-surround, or end-stopping, network that generates hypercomplex cell receptive fields from combinations of complex cell outputs. Higher order hypercomplex cells further transform hypercomplex cell outputs via a push-pull competition across orientations.

a hole at the line end. Such a boundary could not prevent brightness and color signals from flowing out of the line end during FCS filling-in (*Figure 19b*). Many percepts would hereby become badly degraded by featural flow.

Later processing stages within the BCS are needed to close the holes at line ends and other positions of rapid change in the orientation of scenic contours. The process which completes the boundary at a line end is called an *end cut* (*Figure 19c*). End cuts actively reconstruct the line end at a higher cortical processing stage than the oriented receptive field. They are often, but not always, perpendicular to the line end, as they are in response to the vertical lines in *Figure 19a*. Because line ends are constructed, much as illusory contours are, at a cortical processing stage, we say that *all line ends are illusory*.

### 16. The SOC Filter: The Simple-Complex-Hypercomplex Cell Module

The processing stages that are hypothesized to generate end cuts are summarized in *Figure 18* (Cohen & Grossberg, 1984; Grossberg & Mingolla, 1985a, 1985b). First, pairs of simple cells sensitive to like position and orientation but opposite direction-of-contrast generate rectified output signals that summate at the next processing stage to activate complex cells. The target complex cells are thus sensitive to the same position and orientation, but are insensitive to direction-of-contrast. These pairs

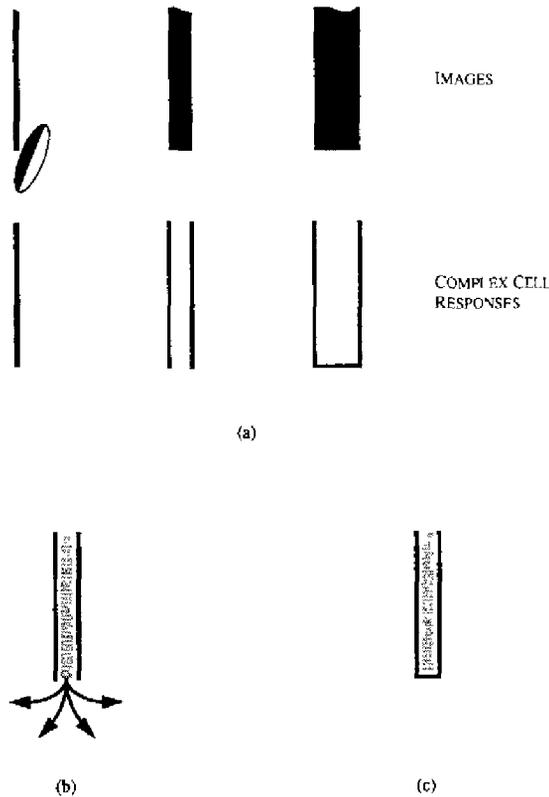


Figure 19. Responses of oriented receptive fields to lines of variable width: (a) Narrow lines and thick edges activate a connected band of oriented responses. Intermediate line widths are not detected at the line end. (b) For such incomplete boundaries, filling-in could cause diffusion of visible signals out of the line end. (c) An end cut closes the boundary at a line end.

of simple cells may be odd-symmetric and even-symmetric (Foster, Gaska, Nagler, & Pollen, 1985; Pollen, Foster, & Gaska, 1985; Pollen & Ronner, 1981, 1982; Spitzer & Hochstein, 1985). The complex cells maintain their sensitivity to *amount* of oriented contrast, but not to the *direction*, or polarity, of this oriented contrast. They pool inputs from receptive fields with opposite direction-of-contrast in order to generate boundary detectors that can detect the broadest possible range of luminance or chromatic contrasts (DeValois, Albrecht, & Thorell, 1982).

The rectified output from a complex cell activates a second filter which is composed of two successive stages of spatially short-range competitive interaction whose net effect is to generate end cuts (Figure 18). First, a complex cell of prescribed orientation excites like-oriented hypercomplex cells corresponding to its location and inhibits like-oriented hypercomplex cells corresponding to nearby locations. As a result, an on-center off-surround organization of like-oriented cell interactions exists around each perceptual location. This mechanism is analogous to the

neurophysiological process of *end stopping*, whereby hypercomplex cell receptive fields are derived from interacting complex cell output signals (Hubel & Wiesel, 1965; Orban, Kato, & Bishop, 1979). The outputs from these model hypercomplex cells activate a second competitive mechanism that inputs to model higher order hypercomplex cells. Here, at each perceptual location, cells cooperate if they represent similar orientations but compete if they represent dissimilar orientations, notably perpendicular orientations. This competition defines a tonically active, push-pull opponent process. If a given orientation is excited, then its perpendicular orientation is inhibited. If a given orientation is inhibited, then its perpendicular orientation is excited via disinhibition. In summary, the hypercomplex cells at the first competitive stage compete *across* positions *within* each orientation; the higher order hypercomplex cells at the second competitive stage compete *within* each position *across* orientations.

The combined effect of these two competitive interactions is to generate end cuts as follows (Figure 20). The strong vertical complex cell activations along the edges of a vertical scenic line inhibit the vertical hypercomplex cells just beyond the line end. These inhibited hypercomplex cells disinhibit horizontal activations of higher order hypercomplex cells at the line end to form an end cut. In summary, the positional uncertainty generated by the orientational tuning of simple cell receptive fields is eliminated by the interaction of two spatially short-range competitive mechanisms which convert complex cells into two distinct populations of hypercomplex cells.

The properties of these competitive mechanisms have predicted and helped to explain a variety of neural and perceptual data. For example, the prediction of the theory summarized in Figure 20 anticipated von der Heydt et al.'s (1984) report that cells in prestriate visual cortex respond to perpendicular line ends, whereas cells in striate visual cortex do not. The end-cut process also exhibits properties of hyperacuity which have been used (Grossberg, 1987b) to explain psychophysical data about hyperacuity (Badcock & Westheimer, 1985a, 1985b; Watt & Campbell, 1985). A version of the double-filter model in Figure 18 has been used to explain data about texture segregation (Beck, Sutter, & Ivry, 1987; Sutter et al., 1989) in a way that supports the texture analyses of Grossberg and Mingolla (1985b, 1987a). See Section 20.

### 17. The Role of End Cuts and End Gaps in Neon Color Spreading

The emergent segmentation process has also enabled FACADE theory to explain an extensive body of data about monocular neon color spreading (Grossberg, 1987b, Section 31; Grossberg, 1987c, Section 21; Grossberg & Mingolla, 1985a). This explanation is reviewed herein because it includes key ideas for explaining figure-ground separation, including the percepts reviewed in Sections 2-10, such as 3-D neon color spreading (Nakayama et al., 1990), the nearer appearance of higher spatial frequencies in suitable spatial configurations (Brown & Weisstein, 1988b),

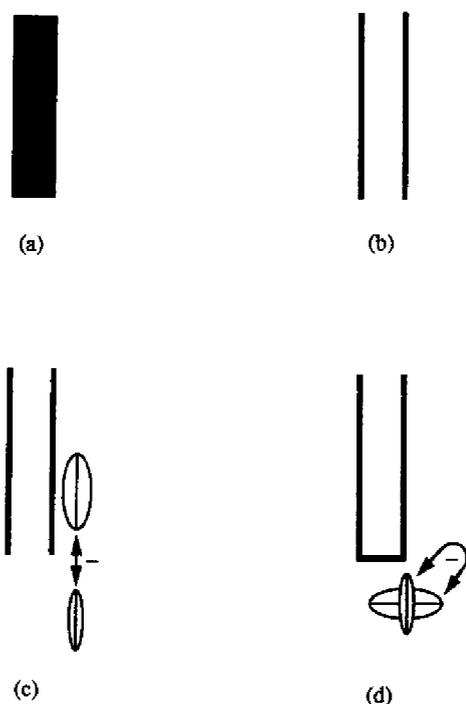


Figure 20. Creation of end cuts: (a) A line of intermediate width; (b) complex cell activations leave a gap at the line end; (c) the first competitive stage inhibits vertical hypercomplex cells at the line end; (d) the second competitive stage generates an end cut by disinhibiting horizontal higher-order hypercomplex cells at the line end.

and recognition of the occluded Bregman-Kanizsa Bs (Bregman, 1981; Kanizsa, 1979).

Consider the Redies-Spillmann display in Figure 21a. When such displays are properly prepared, the gray color of the cross fills in a disk-shaped area whose boundary intersects the white/gray vertical and horizontal edges. Grossberg and Mingolla (1985a) described how BCS segmentations and FCS filling-in operations could explain this percept. In particular, in order for neon to spread out of the cross, the contrast of the inner cross with respect to the background must be less than that of the flanking branches of the cross with respect to the background (van Tuijl & de Weert, 1979). Accordingly, in Figure 21a, the gray cross against the black background has less contrast than the white bars against the black background. For definiteness, consider a vertical branch of the cross, where the gray and white bars join. Within the BCS, the vertically oriented complex cells are more activated at the white/black contour than at the gray/black contour (Figure 21b).

Thus, the vertical hypercomplex cells at the white/black contour receive net excitation from the complex cells. However, the vertical hypercomplex cells at the gray/black contour that are close to the end of the cross receive net inhibition from the complex cells. Thus gaps begin to form in the vertical BCS boundaries at the vertical gray/black

contours near the end of the cross. In addition, inhibition of these vertical hypercomplex cells disinhibits horizontal higher order hypercomplex cells at the end of the cross (Figure 21c). Horizontal end cuts are hereby formed. A similar argument shows how horizontal gaps and vertical end cuts occur at the gray/black contours near the horizontal ends of the cross. Such gaps, called *end gaps*, are the locations at which gray color flows out of the cross into the surrounding black region during neon color spreading. Neither these gaps nor the BCS boundaries are fully formed until the long-range cooperative-competitive feedback interactions of the CC loop are activated by, and equilibrate to, inputs from the SOC filter, as summarized below.

### 18. The CC Loop: Long-Range Cooperation, Feature Binding, and Boundary Completion

Hypercomplex cell outputs from the second competitive stage input to the bipole cells of the CC loop (Figure 12). The bipole cells define a spatially long-range cooperative interaction that helps to generate sharp coherent boundary segmentations from noisy local boundary

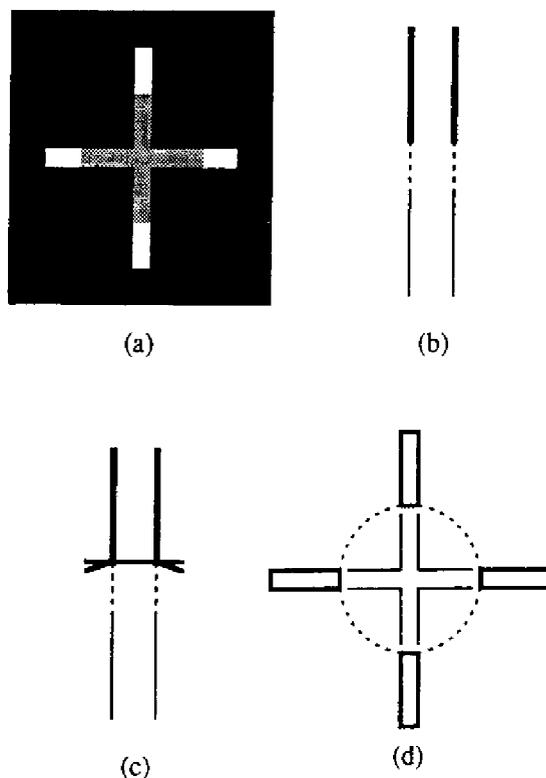


Figure 21. (a) A 2-D Redies-Spillmann display. (b) End gaps start to form at the boundaries of the gray cross that abut the boundaries of the white edges at the first competitive stage of hypercomplex cells. (c) End cuts start to form at the ends of the boundaries of the white edges at the second competitive stage of hypercomplex cells. (d) An illusory contour connecting the end cuts is created by the cooperative-competitive (CC) loop.

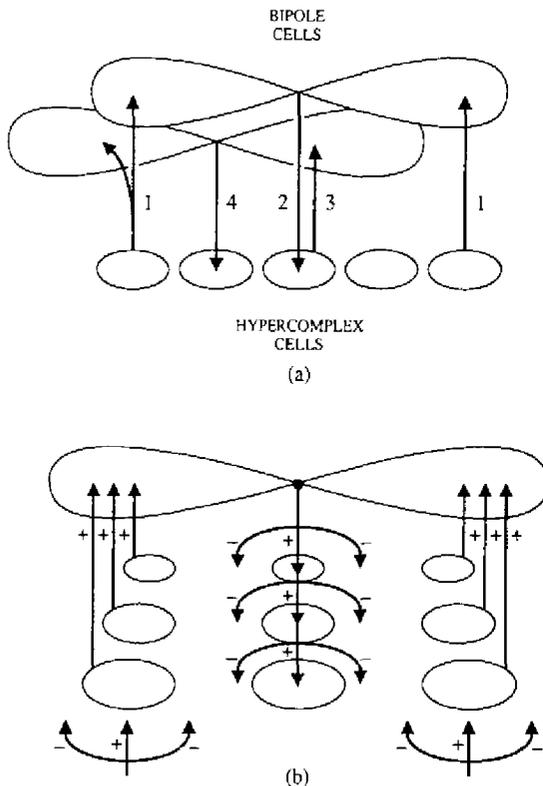


Figure 22. (a) A cooperative-competitive feedback exchange leading to boundary completion: Cells in the bottom row represent like-oriented cells at the second competitive stage whose orientational preferences are approximately aligned across perceptual space. The cells in the top two rows are bipole cells in the cooperative layer whose receptive field pairs are oriented along the axis of the competitive cells. Suppose that simultaneous activation of the pair of pathways 1 activates positive boundary completion feedback along pathway 2. Then pairs of pathways, such as 1 and 3, activate positive feedback along pathways such as 4. Rapid completion of a sharp boundary between the locations of pathways 1 can hereby be generated. (b) Multiple receptive field sizes cooperate and compete with a shared pool of bipole cells to form a 3-D boundary segmentation corresponding to a prescribed range of relative depths from the observer. This segmentation provides the best consensus of positional and orientational information from all of the interacting signals.

fragments. During this boundary completion process, hypercomplex cell outputs from (approximately) like-oriented cells that are (approximately) aligned across perceptual space cooperate via bipole cells to begin the synthesis of an intervening boundary. For example, such a boundary completion process can span the faded stabilized images of retinal veins and the retinal blind spot (Kawabata, 1984; Ramachandran, 1992). Boundary completion also completes the illusory boundaries perceived in Figure 8. This process overcomes a type of informational uncertainty that is different from the type that is overcome by end cuts (see Grossberg, 1987b, Sections 15-18, for further discussion).

The boundary completion process works as follows. As illustrated by Figure 22, pairs of similarly oriented and spatially aligned hypercomplex cells of the second competitive stage are needed to activate the intervening cooperative bipole cells. These bipole cells feed back excitatory signals to like-oriented hypercomplex cells at the first competitive stage, which, in turn, activate higher order hypercomplex cells that compete between orientations at each position of the second competitive stage. In Figure 22, positive feedback signals are triggered in pathway 2 by a bipole cell if sufficient activation occurs simultaneously in both of the feedforward pathways labeled 1 from similarly oriented hypercomplex cells of the second competitive stage. Then pathways labeled 1 and 3 can trigger feedback in the set of all bipole cells with pathways 4, both of whose receptive fields can get excited by these pathways. This feedback exchange can rapidly complete an oriented boundary between pairs of inducing scenic contrasts via a parallel exchange of bipole-gated feedback signals.

Such a boundary completion process realizes a type of real-time statistical decision theory that chooses the globally most salient boundary segmentations and suppresses less favored groupings. Each bipole cell is sensitive to the position, orientation, density, and size of the inputs that it receives from the second competitive stage. Each bipole cell performs like a type of statistical "and" gate, since it can fire feedback signals to the first competitive stage only if both of its receptive field branches, or poles, are sufficiently activated. In response to a continuous image edge, the boundary completion process generates boundary activations simultaneously along the full length of the edge. In response to a widely separated pair of inducers, such as a Kanizsa square (Figure 10), only a subset of bipole cells may be able to bridge the gap on the first feedback cycle, as in Figure 22, followed by rapid parallel completion by simultaneously acting arrays of bipole cells on the second and subsequent cycles. The inward action of the bipole cells enables boundary segmentations to form over variable distances in a self-scaling fashion that maintains boundary sharpness, speed of formation, and insensitivity to image noise. Computer simulations showing boundary completion of a circle, a reverse-contrast Kanizsa square, and an Ehrenstein figure were first reported in van Allen and Kolodzy (1987).

The existence of such bipole cells was predicted (Cohen & Grossberg, 1984; Grossberg, 1984; Grossberg & Mingolla, 1985a, 1985b) shortly before von der Heydt et al. (1984) reported analogous cell properties in prestriate visual cortex, in the same report that confirmed that prestriate cortical cells respond to perpendicular line ends, to which striate cortical cells do not, as in Figures 19 and 20. Peterhans and von der Heydt (1989) have also observed the predicted cooperative sharpening of prestriate cell responses as an increasing number of colinearly arranged perpendicular line ends are added to the image. Eckhorn et al. (1988) and Gray, König, Engel, and Singer (1989) have shown that cooperative linking operations

among similarly oriented cells can cause synchronous oscillations of cortical activity, as can also occur in the CC loop. See Sections 22 and 23.

### 19. The CC Loop and Neon Color Spreading

The CC loop completes an illusory boundary that connects the four sets of end cuts at the ends of the gray cross (Figure 21d). As these illusory boundaries form, their bipole cells supply positive feedback to the corresponding hypercomplex cells at the competitive stages. This positive feedback increases the inhibition across orientations at the second competitive stage (Figure 12). The hypercomplex cells at the white/black contours can withstand this inhibition because they vigorously activate their own bipole cells. The inhibited hypercomplex cells at the gray ends of the gray/black contours cannot. The end gaps in the BCS boundaries at the ends of the gray cross are hereby fully formed. When this emergent boundary structure is mapped into the FCS, it allows gray color to flow through the gaps and to fill-in the circular boundary that surrounds the cross. The top-down feedback from bipole cells to hypercomplex cells also inhibits hypercomplex cells that code nearby positions (Figure 12). This negative bipole feedback helps to select the best positions of an emergent boundary, while the hypercomplex competition among orientations helps to select the best orientations (Grossberg & Mingolla, 1987a).

This explanation of 2-D neon color spreading provides a foundation for explaining 3-D neon color spreading and transparency. In particular, I suggest that the end gaps that initiate neon color spreading also initiate the phenomena of boundary detachment whereby the boundaries of the black occluder in Figure 6b are detached from the boundaries of the gray occluded regions. The remaining boundaries of the gray regions can then complete the B boundaries via emergent segmentation.

Before addressing these 3-D phenomena, I review some properties of 2-D binding and segmentation in order to increase the reader's intuition as well as to supply additional evidence that supports the existence of these BCS mechanisms in the brain.

### 20. Textural Segregation and Grouping

Grossberg and Mingolla (1985a, 1985b) proposed that data on textural segregation (Beck et al., 1983) could be accounted for by the oriented filtering, competitive, and cooperative interactions of the static BCS. A number of psychophysical studies and similar computational models have since supported this intuition, and provided new data on textural segregation (Bergen, 1991; Sutter et al., 1989). The work of Sutter et al. (1989) is especially significant for its extensive psychophysics and model simulations. Sutter et al.'s *complex channel* model contains an initial stage of linear, oriented filtering, analogous to SOC filter simple cells. The model rectifies or squares these signals to generate insensitivity to direction of contrast, as at SOC filter complex cells. Lastly, the model

employs another, spatially broader linear filter analogous to the on-center off-surround interaction from SOC filter complex cells to hypercomplex cells. One notable difference is that the complex channel second filter can be at any orientation, whereas the BCS first competitive stage has the same orientation as the complex cells, but a larger spatial scale. The cooperative bipole cells of the BCS—which have no analog in the Sutter et al. model—can group signals over a variety of orientations, which may differ from the orientational preference of the SOC filter cells that excite the bipole cells.

Graham et al. (1992) obtained good fits to most of their texture data with the complex channel model, but noted that the effects of element-to-background contrast required an additional compressive nonlinearity. They suggested that this nonlinearity might occur either before or after the stages of oriented filtering in the complex channel, and concluded that the latter approach fitted their data better. This compressive nonlinearity was achieved by a cross-orientation inhibition analogous to the BCS second competitive stage. Graham et al. also remarked that “higher level processes may turn out to play a substantial role in region segregation but such processes should not be invoked until they are needed.” Such “higher level” processes seem to include the cooperative linking, or binding, of features by CC loop bipole cells into “emergent features” such as those described by Beck et al. (1983) and illustrated in Figures 8 and 9. Cruthirds et al. (1993) have achieved better model fits to texture data by including cooperative bipole cells. Leshner and Mingolla (1993) also needed a combination of competitive and cooperative BCS interactions to account for their data on illusory contour clarity.

### 21. Cortical Dynamics of Reset and Visual Persistence

Positive feedback plays an important role in such a cooperative feature binding process, as in Figure 22. The visual system needs a cooperative-competitive feedback process to rapidly choose from the infinitude of possible completions within a spatial distribution of boundary signals. Models that do not include a feedback process (e.g., Peterhans & von der Heydt, 1989) cannot make such decisions and will either generate spurious signals all over the plane or use thresholds to choose only the strongest signals (von der Heydt, Heitger, Rosenthaler, Peterhans, & Kübler, 1992). Either situation is problematic. Without a threshold, too many boundary signals will clutter the percept. With a threshold, a weak but desirable completion cannot be distinguished from a spurious weak completion. In contrast, a feedback cooperative-competitive system has self-calibrating properties (Grossberg, 1973, 1983) which dynamically change with the input array so that appropriate completions, even if weak, are well represented but spurious ones are inhibited.

While positive feedback enables coherent bindings of spatially distributed feature detectors to form rapidly, it

also creates hysteresis that could cause bindings to persist too long. As the visual scene rapidly changes, such persistence could degrade perception by causing massive smearing to occur. To correct this problem, a system is needed that can maintain coherent segmentations within unchanging parts of a scene, while it resets segmentations that correspond to moving scenic parts in a form-specific and speed-sensitive manner. Grossberg (1980, 1991) suggested how properties of the visual persistence that is described in psychophysical experiments could be caused by positive feedback in the visual cortical circuits that are responsible for the binding or segmentation of visual features into coherent visual forms. It was also suggested how the degree of persistence may be limited by circuits that reset these segmentations at stimulus offset, notably by habituated transmitter gates embedded in the cooperative-competitive circuits. Such gates enable off-cells (see Figure 12) to be activated in response to offset of inputs to the on-cells that activate and maintain the segmentations. These off-cell rebounds are used to rapidly inhibit the CC loop bipole cells and thus to shut down the corresponding segmentations. Francis, Grossberg, and Mingolla (in press) have shown how the static BCS model, including these habituated transmitter gates, can quantitatively simulate psychophysical data showing increase of persistence with spatial separation of a masking stimulus; inverse relation of persistence to flash luminance and duration; greater persistence of illusory contours than real contours, with maximal persistence at an intermediate stimulus duration; and dependence of persistence on preadapted stimulus orientation. Data concerning cortical cell responses to illusory and real contours were also analyzed, as were alternative models of feature binding and persistence properties. The simulations of data showing an inverse relation of persistence to flash luminance support the idea that reset of the boundary segmentation and collapse of the filled-in percept that it supports subserves this property. Data showing a direct relation of persistence to flash luminance, due to aftereffects or informational persistence of the stimuli, as in iconic or working memory storage processes (Coltheart, 1980; Nisly & Wasserman, 1989), are not simulated by this model.

Figures 23 and 24 summarize illustrative comparisons between psychophysical data and model simulations. Figure 23a summarizes data of Bowen, Pola, and Matin (1974) showing the inverse relation between visual persistence and stimulus duration and luminance. Figure 24a summarizes data of Meyer and Ming (1988) showing greater visual persistence of illusory contours than of real contours with a maximum persistence of illusory contours at an intermediate stimulus duration. These simulations show how every processing stage of the BCS model contributes to the explanation of these data. The habituated static BCS model has also been used to explain many psychophysical data about orientational aftereffects, binocular rivalry, and the McCollough effect (Grossberg, 1987c). These model explanations provide a conceptual bridge whereby to cross psychophysical and neural paradigms in order to design new types of experiments, much

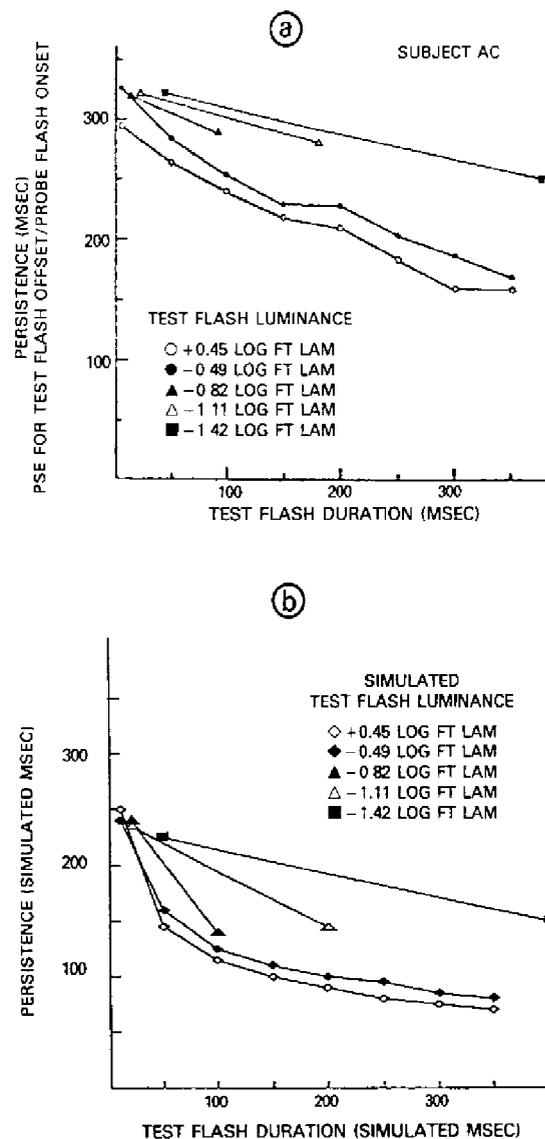


Figure 23. (a) Visual persistence is inversely related to flash luminance and flash duration. From "Visual persistence: Effects of flash luminance, duration, and energy" by R. Bowen, J. Pola, and L. Matin, 1974, *Vision Research*, 14, 295-303. Copyright 1974 by Pergamon Press. Reprinted with kind permission from Pergamon Press Ltd., Headington Hill Hall, Oxford OX3 0BW, U.K. (b) Computer simulation by Francis, Grossberg, and Mingolla (in press) of the Bowen et al. (1974) experiments using the boundary contour system (BCS).

as Nakayama et al. (1990) combined stereopsis and neon color-spreading experiments.

## 22. Synchronous Feature Binding in Visual Cortex

Further evidence for the CC loop derives from its ability to rapidly synchronize the activities of spatially distrib-

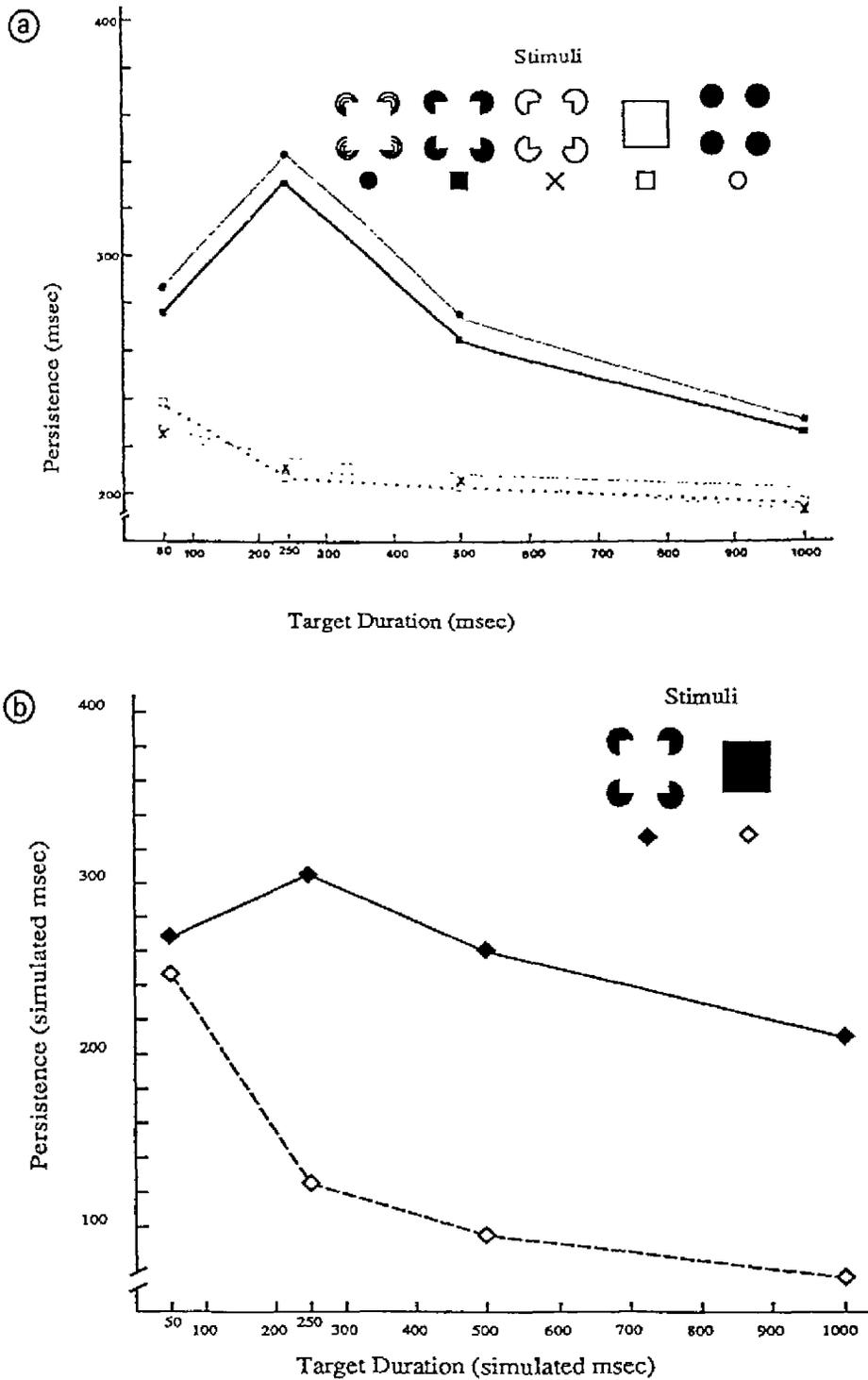


Figure 24. (a) Visual persistence of illusory contours is greater than persistence of real contours and peaks at an intermediate duration of the stimulus. From "The visible persistence of illusory contours" by G. Meyer and C. Ming, 1988, *Canadian Journal of Psychology*, 42, 479-488. Copyright 1988 by Canadian Psychological Association. Reprinted by permission. (b) Computer simulation by Francis, Grossberg, and Mingolla (in press) of the Meyer and Ming (1988) experiments using the boundary contour system (BCS).

uted cells within a single processing cycle. A theoretical rationale of such synchronization *in vivo* is suggested by considering the processing of a visual image whose various parts are registered approximately synchronously at the eyes. Subsequent processes at the multiple neural stages leading from retina to cortex may destroy this synchrony. These desynchronizing processes include spatially inhomogeneous stimulus parameters, internal cellular noise, and transmitter habituation. Thus, even if an external source of spatially distributed information is synchronous, it could be represented with spatially inhomogeneous phase leads and lags at subsequent neural processing stages. If these various parts of the processed image are not rapidly recombined into synchronized spatial patterns, then they cannot be correctly learned and recognized by the brain as part of the same event. Phase-lagging image parts from a previous image could be combined with phase-leading image parts from a later image, leading to a maladaptive reorganization of scenic parts into the "wrong" visual objects. The period of such a synchronous oscillation defines a time window during which phase leading and lagging parts of an image representation may be recombined into the correct spatial pattern code. Synchronous binding of features into coherent spatial pattern codes has also been described as part of the process whereby complex acoustic signals are grouped into unitary auditory events (Bregman, 1990; Lenneberg, 1967).

CC loop mechanisms can quickly reorganize temporally out-of-phase spatial data into a synchronous spatial pattern, as part of their emergent segmentation properties, just so long as the phase lags and leads of these data lie within a critical time interval. Moreover, model cells may be synchronized which lie much further apart than the width of an individual bipole cell receptive field. In order to be effective during real-time perception, during which a series of rapidly changing images must be correctly processed, the synchronizer must be able to act quickly, before spurious binding of incorrect image parts can occur. This property holds in the model as well as in the data.

The relevant data were collected using simultaneous, multielectrode, extracellular recordings. Two labs (Eckhorn et al., 1988; Gray et al., 1989; Gray & Singer, 1989) reported stimulus-evoked synchronized oscillations of 40–60 Hz in the primary visual cortex (Areas 17 and 18) of the cat. Coherence or synchrony of firing activity was found between cells within a cortical column (Eckhorn et al. 1988; Gray & Singer, 1989), in neighboring hypercolumns (Eckhorn et al., 1988; Gray et al. 1989), in distant hypercolumns (Gray et al. 1989), and lying in two different cortical areas (Eckhorn et al. 1988). Stimulus position, orientation, movement direction, and velocity were among the stimulus properties that yielded stimulus-evoked resonances. Synchronized oscillatory responses were frequently found at distant cortical positions when at least one of the primary coding properties was similar.

### 23. Synchrony During Preattentive Visual Coding and Attentive Recognition Learning

Our simulations of synchronized oscillations represent a confluence of several streams of research that are relevant to the distinction between seeing and recognizing that was made in Section 5. It was predicted in Grossberg (1976a, 1978b) that visual cortical codes could be expressed by resonant standing waves in which cooperatively linked cells oscillate in phase with one another. The mathematical analysis of such synchronous oscillations was begun in Ellias and Grossberg (1975). It was also noted that these standing waves could be replaced by approach to an equilibrium point, or attractor, if no "slow" variables, such as inhibitory interneurons or chemical modulators, existed in the network. Both standing waves and equilibria can, in principle, support a feature-based cortical code. The standing waves were called "order-preserving limit cycles" to emphasize that the ordering, or relative importance, of feature detector activations should persist during each coding cycle, even if their absolute activations varied through time as the oscillation unfolded.

The standing wave prediction was made in the context of a theory called adaptive resonance theory, or ART, which analyzes the role of reciprocal top-down and bottom-up corticocortical and thalamocortical adaptive filters in the development of cortical feature detectors, recognition learning, attentional processing, and memory search. Within ART, a resonant standing wave can occur when bottom-up feature-selective signals and top-down expectation signals fuse into an attentional focus. Such an attentional focus can support new learning as it gives rise to a conscious perceptual experience (Grossberg, 1980, 1982a). The predicted linkage between standing waves, attention, and conscious experience has recently attracted the interest of a large number of investigators (e.g., Crick & Koch, 1990). Eckhorn and Schanze (1991) and Grossberg and Somers (1991) have simulated standing waves using the type of bottom-up and top-down feedback interactions among adaptive filters that are used in ART circuits. Intraub (1985) has described experimental evidence consistent with ART-like temporal processing in a visual dissociation effect whereby parts of a picture may be more easily displaced to the preceding or following picture under rapid presentation conditions, if these parts are not bound together by learned expectations.

After ART was introduced in order to analyze attentive learning and recognition, subsequent research by Grossberg and Mingolla (1985a, 1985b) focused on processes of preattentive vision. As noted above, bipole cells were hereby predicted to cooperatively link perceptual features into emergent boundary segmentations via cooperative-competitive feedback signals within a CC loop. Grossberg and Somers (1991) demonstrated that CC loop circuits could also cooperatively link cells into stimulus-specific standing waves wherein cell activities might be

rapidly synchronized. As in the neurophysiological experiments, the input stimuli in these simulations were either long single-bar images or short disconnected double-bar images. The model neurons were given random initial activities. In the uncoupled case, where neurons do not interact, units receiving sufficient input exhibit stable limit-cycle oscillations, while units receiving insufficient input quickly approach a low equilibrium value (Figures 25a and 26a). Since the initial values were chosen randomly, the units oscillated in random phase with respect to one another. If all the inputs had the same amplitude, these phase relationships did not change over the time course of the simulation, since their frequencies were the same.

Using a CC loop bipole cell coupling, coherent oscillations emerged within one cycle for both the one-bar (Figure 25b) and two-bar (Figure 26b) stimuli. In the two-bar case, oscillations were induced in the slit between the two bars to create an illusory contour, and these oscillators could be almost instantly synchronized with the others. In both stimulus cases, the bipole architectures did not induce a spreading of oscillatory activity to the outer regions beyond the stimuli. Inward boundary completion without outward spreading of oscillatory activity was found to be a robust property of bipole coupling. Adaptive filter coupling also yielded rapid synchronization. It remains to analyze how the variable amount of time that it takes to reset a previous segmentation, as during visual persistence (Section 21), may influence the amount of time it takes to trigger the next segmentation and its synchronized oscillation, both in vivo and in model simulations.

Finer differences between the global structure of these oscillations may also be used to infer their different functional roles, while also testing predictions of the preattentive BCS theory and the attentive ART theory. A preattentive BCS resonance is predicted to complete across gaps in two stimulus inducers, as during the Gray et al. (1989) two-bar experiments and the perception of illusory contours (Grossberg & Mingolla, 1985a, 1985b). In contrast, an attentive ART resonance is predicted not to complete across gaps in stimulus inputs. Its top-down expectations can "confirm the hypothesis" that input features are present and can bind them into coherent recognition codes, but it does not activate new features that are not already represented in the input data (Carpenter & Grossberg, 1987a, 1987b, 1991; Grossberg, 1987b; Grossberg & Stone, 1986).

These results suggest that synchronous oscillations are not *necessary* to carry out cortical binding processes. Preattentive segmentation and attentive recognition can be achieved without them. Moreover, synchronous oscillations may be structurally or dynamically switched on just by making (say) inhibition change more slowly than excitation. On the other hand, synchronous binding does provide an extra degree of freedom for temporally realigning distributed visual data that have drifted out of phase due to multistage visual processing. Thus, although synchronous oscillations may not be necessary for binding per se, they may facilitate binding of the correct image parts into coherent spatial patterns.

#### 24. A Synthesis of Preattentive Vision and Attentive Recognition Networks

As noted in Section 16, the processes of the static BCS start out as preattentive and automatic processes. These preattentive processes may, however, influence and be influenced by the types of attentive, learned object-recognition processes that were noted in Section 23. A preattentively completed segmentation within the BCS can directly activate the object recognition system (ORS), whether or not this segmentation supports visible contrast differences within the FCS (Figure 7). The ORS can, in turn, read out attentive learned expectation, or memory priming, signals to the BCS. In response to familiar objects in a scene, the final 3-D segmentation within the BCS may thus be *doubly* completed, first by automatic preattentive segmentation processes and then by attentive learned expectation processes. This doubly completed segmentation regulates the filling-in processes within the FCS that lead to a percept of visible form.

The analysis below suggests how the occluded B boundaries in Figure 6b are completed behind the black occluder. Such completed B boundary segmentations may then be recognized within the ORS via direct BCS → ORS signals, even though they are not seen as visible surface properties at the FCS. It was suggested in Grossberg (1987b) that the reciprocal BCS ↔ ORS interactions could be attributed to ART mechanisms, including the reciprocal adaptive filter for supporting synchronized oscillations that was described in Section 23. Recent neurophysiological evidence suggesting that regions of the inferotemporal cortex that play a role in visual object recognition may embody ART-like mechanisms are reviewed in Carpenter and Grossberg (1993), Desimone (1992), and Gochin (1990).

### PART III

#### 25. Binocular Perception and 3-D Segmentation

The original static BCS model of Grossberg and Mingolla (1985a, 1985b) considered only monocular processing. Later research showed that the BCS could consistently be generalized to a binocular theory. A key design insight was derived from psychophysical data showing that human stereo vision is not based upon the matching of left and right image contrasts, as many AI vision theories had proposed. Rather, it is based upon the matching of left and right emergent segmentations (Kaufman, 1974; Ramachandran & Nelson, 1976; Tausch, 1953; Wilde, 1950). This well-known fact could not be incorporated into a computational vision theory until it was shown how, as in the BCS, emergent segmentations arose. The binocular theory showed how the monocular SOC filter could be generalized to a multiple-scale binocular filter whose outputs are automatically sorted by multiple CC loops into binocularly fused or suppressed segmentations (Grossberg, 1987c). As described below, interactions of cortical ocular dominance columns, self-organizing feature maps,

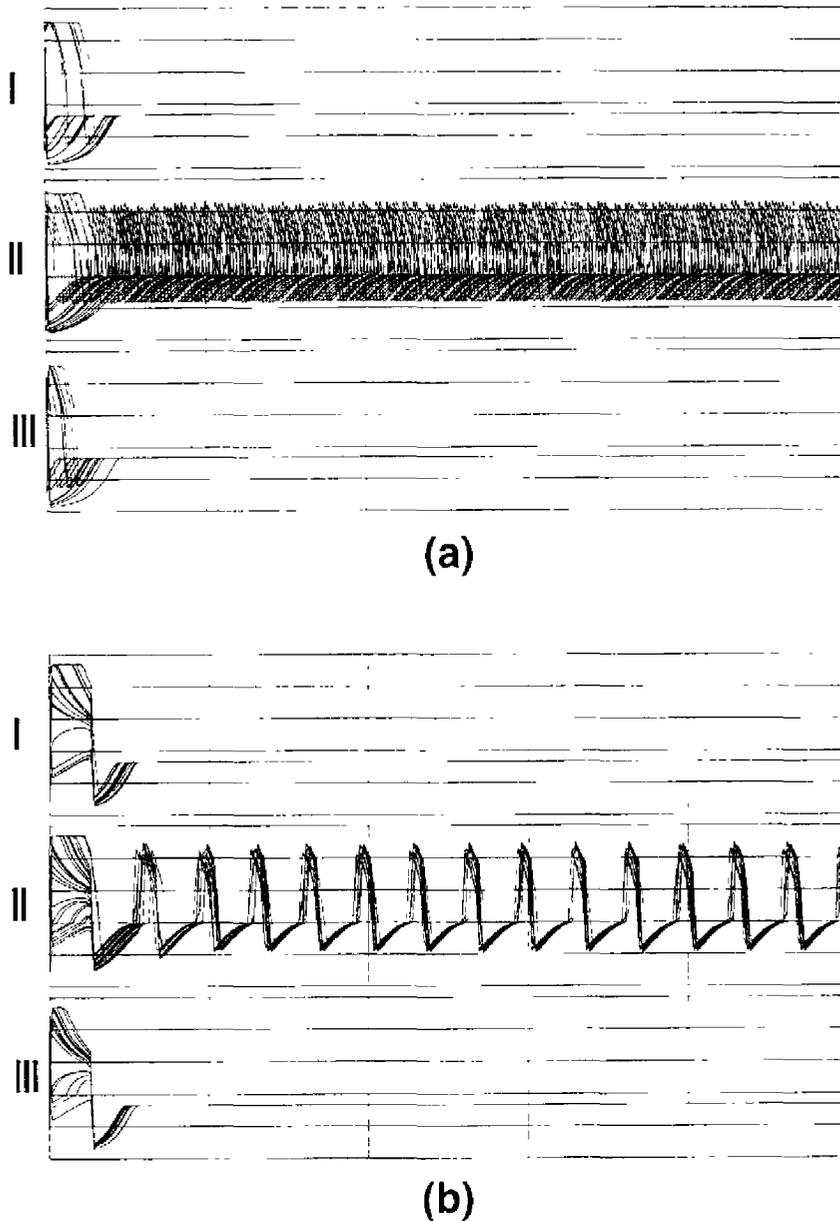


Figure 25. (a) Uncoupled model cell activities in response to a single bar input. Three windows are displayed in which plots of cell activities,  $X_i$ , versus time are overlaid. In window I, positions  $i=1$  through  $i=18$  are overlaid. In window II, positions  $i=19$  through  $i=46$  are plotted. Positions  $i=47$  through  $i=64$  are shown in window III. The positions displayed in windows I and III correspond to the image background, while window II displays activity of  $X_i$  along the bar. In this uncoupled case, the activities at positions corresponding to the background quickly approach the steady-state value, while positions along the bar oscillate in random phase. This uncoupled case represents the control simulation in response to a single bar input. (b) Bipole coupling in response to a single bar input. Using the same inputs and initial conditions which were used to generate (a), bipole coupling yielded rapid and sustained synchronization of oscillatory activity at positions along the bar without inducing oscillatory activity at positions corresponding to the background. Each bipole flank received input from six neighboring positions. From "Synchronized oscillations during cooperative feature linking in a cortical model of visual perception" by S. Grossberg and D. Somers, 1991, *Neural Networks*, 4, 453-466. Copyright 1991 by Pergamon Press. Reprinted with kind permission from Pergamon Press Ltd., Headington Hill Hall, Oxford OX3 0BW, U.K.

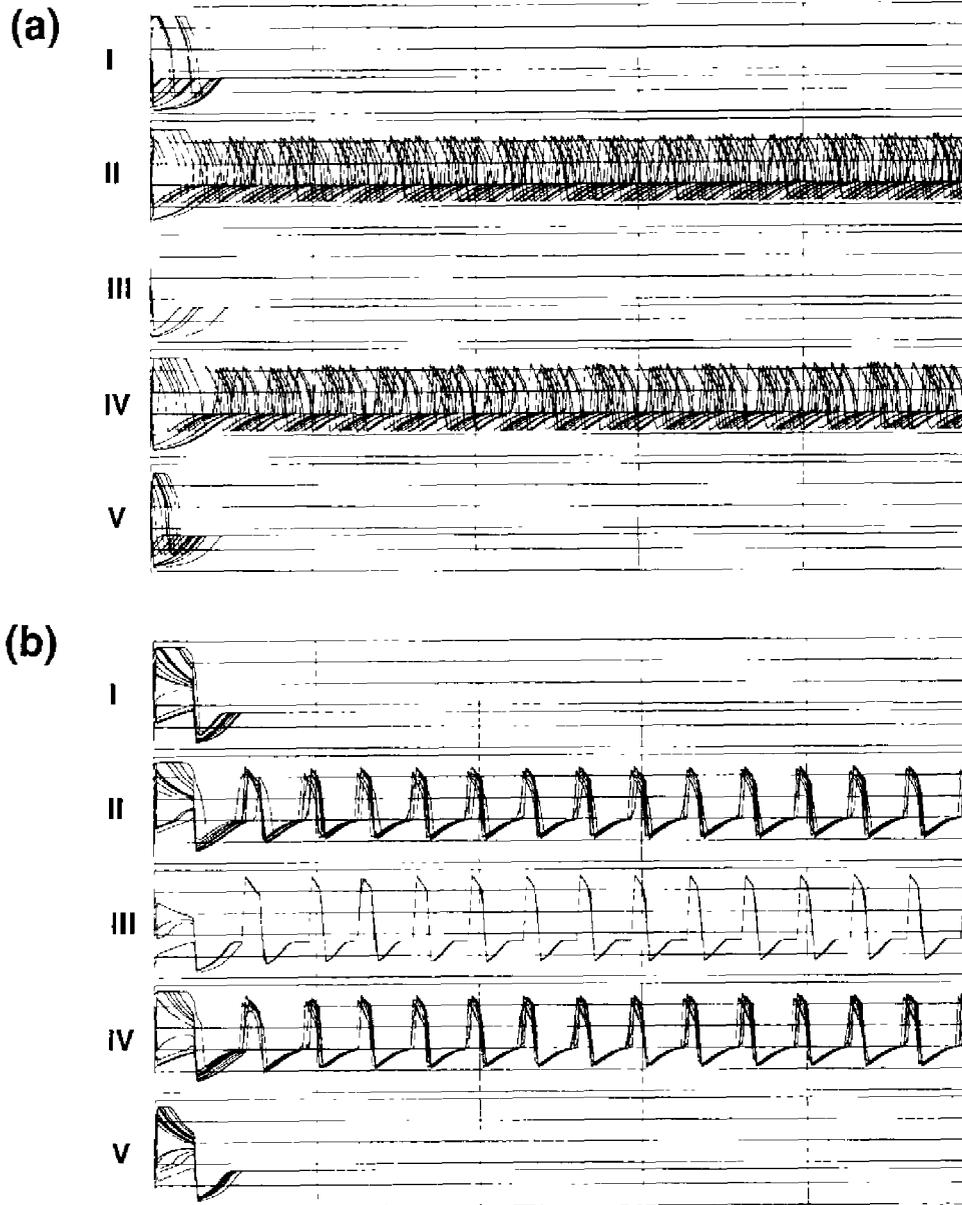


Figure 26. (a) Uncoupled model cell activities in response to a double bar input. Five windows are displayed in which plots of cell activities,  $X_i$ , versus time are overlaid. In window I, positions  $i=1$  through  $i=18$  are overlaid, while in window II, positions  $i=19$  through  $i=30$  are plotted. Windows III and IV display positions  $i=31$  through  $i=34$  and  $i=35$  through  $i=46$ , respectively, and positions  $i=47$  through  $i=64$  are shown in window V. The positions displayed in windows I and V correspond to the image background, while window III displays activity in the slit between the two bars. Windows II and IV display activities of  $X_i$  along the left and right bars, respectively. In this uncoupled case, the activities at positions corresponding to the background and the slit quickly approach the same steady-state value, while the activities at positions along both bars oscillate in random phase. This uncoupled case represents the control simulation for double-bar input. (b) Bipole coupling in response to a double-bar input. Using the same inputs and initial conditions that were used to generate Figure 26a, bipole coupling yielded rapid and sustained synchronization of oscillatory activity at positions along both bars and induced synchronous oscillatory activity at the positions between the two stimulus bars, but did not induce oscillatory activity at positions corresponding to the outer background regions. This may be interpreted as the completion of a disconnected image boundary, resulting in an illusory contour between the two bars. Each bipole flank received input from six neighboring positions. From "Synchronized oscillations during cooperative feature linking in a cortical model of visual perception" by S. Grossberg and D. Somers, 1991, *Neural Networks*, 4, 453-466. Copyright 1991 by Pergamon Press. Reprinted with kind permission from Pergamon Press Ltd., Headington Hill Hall, Oxford OX3 0BW, U.K.

and monocular BCS mechanisms enable some spatial scales to exhibit binocular fusion while other scales exhibit binocular rivalry in response to the same stimulus, and a size-disparity correlation to obtain for the maximal disparity at which a given scale can binocularly fuse monocular pairs of boundaries (Grossberg, 1987c; Grossberg & Marshall, 1989). Many data about binocular vision were comprehensible within this binocular BCS theory.

## 26. Interscale and Interstream Interactions

In its original form, this 3-D theory did not posit interactions between the different spatial scales of the BCS and the FCS, or from the FCS to the BCS. Such interactions were not needed to explain the data analyzed in previous articles. The present work shows how interactions within and between BCS and FCS scales lead to explanations of a much wider body of data about 3-D visual perception than could be handled before.

The theory posits the existence of seven types of interactions that complement, and are consistent with, previously defined BCS and FCS mechanisms (Figure 11b). These interactions clarify how the visual system can generate globally unambiguous 3-D surface representations from image data which contain several different types of local ambiguities. The main observation to make about the interactions listed below is that larger scales tend to influence smaller scales, and larger disparities tend to influence smaller disparities. Thus the new interactions tend to be *partially ordered* across scale and disparity. One illustration of this property was provided in Figure 3b to explain why filling-in of a farther surface does not always continue behind a nearer surface, thereby rendering the nearer surface transparent. These new interactions are all listed in this section to give the reader a brief overview of their significance. Later sections show in detail how each interaction helps to explain 3-D data.

The first interaction takes place among the complex cells of the BCS. Model complex cells with large receptive fields can binocularly fuse more disparities than can cells with small fields. Inhibitory competitive interactions occur between complex cells that code different disparities at the same position and size scale. These interactions are called *BB intrascales*. Typically, active BCS complex cells that code larger disparities inhibit complex cells that code smaller disparities—another example of partial ordering. This competition sharpens the disparity tuning curves of the BCS complex cells, and tends to select those complex cells whose disparity tuning best matches the binocular disparities derived from an image.

*BB interscales* are excitatory cooperative interactions from bipole cells to hypercomplex cells that code the same disparity and position, across all scales. These interactions generate multiple emergent boundary segmentations, each corresponding to a prescribed disparity range, or relative depth from the observer. Each segmentation forms the best spatial compromise between all the scales that are sensitive to its disparity range. Each such CC loop network is called a *BCS copy*. Due to the combined effect

of these cooperative interactions and of the competitive interactions of the SOC filter (Figure 12), the larger scales tend to inhibit the smaller scales within each BCS copy in the manner reported in psychophysical data (Tolhurst, 1972; Watt, 1987; Wilson et al., 1991); see Section 42. These interactions are predicted to include the cortical interstripes (Figure 1).

In the theory developed in Grossberg (1987c), each disparity-sensitive 3-D boundary segmentation, or BCS copy, interacts with a monocular FIDO, or Filling-In-Domain, of the FCS, along the BCS → FCS pathways that are denoted in Figure 11a by 2. These BCS signals select those monocular brightness and color signals, labeled  $FCS_L$  and  $FCS_R$ , that are consistent with the binocular BCS segmentation and suppress the rest. These BCS → FCS interactions are called *BF intracopies* in the present theory, because each BCS copy selects binocularly consistent monocular data from a corresponding FCS copy.

Thus the illuminant-discounted monocular FCS representation is transformed into multiple FCS copies, or monocular FIDOS, one for each BCS copy. This one-to-many transformation carries out two functions. First, it maps the monocular positions of FCS signals into the binocular, allelotropic positions of the corresponding BCS copy. It is hypothesized that the BF intracopy signals act as teaching signals to realign the FCS → FCS pathways based on their mutual correlation during visual experience. This adaptive process was used to help explain monocular McCollough effect data in Grossberg (1987c). Second, this one-to-many transformation enables monocular FCS signals that do not positionally match binocular BF intracopy signals in a given FCS copy to be suppressed. The same monocular FCS signals are selected for further processing in a different FCS copy where they do positionally match the corresponding BF intracopy signals. This one-to-many transformation is called *monocular FF intercopies*.

In addition, reciprocal interactions exist from the FCS to the BCS. They are called *FB intercopies*. These FCS output signals are derived from the filled-in FCS regions that are surrounded by connected boundaries, such as the boundaries used to discuss Da Vinci stereopsis in Section 2G. These connected regions are assumed to occur at the monocular FIDOS of Figure 11. The theory develops the hypothesis that the filled-in connected domains, which represent the monocular surface representations that are binocularly consistent, are used to build up the final 3-D surface representation at the binocular FIDOS. In particular, the filled-in connected FCS regions activate contrast-sensitive FCS → BCS pathways that generate FCS output signals at the edges of the filled-in connected regions. These outputs excite BCS cells that correspond to the same disparity and position—namely, the corresponding BCS copy—while inhibiting BCS cells that correspond to smaller disparities at that position. The FB intercopy signals hereby inhibit the BCS boundaries of any occluded region that occurs at the same positions as the boundaries of an occluding region, such as the boundaries of the gray Bs that are shared by the black occluder in the Bregman-

Kanizsa percept (Section 4). The shared B boundaries are hereby eliminated at the smaller disparity representation. The remaining B boundaries may then be colinearly completed by the CC loop at the smaller disparity.

Possible neural loci for these BF intracopies and FB intercopies are suggested by the neural interpretation of the BCS in terms of the interblob cortical stream and of the FCS in terms of the blob cortical stream. These BF and FB interactions must occur at a cortical processing stage that includes (a) oriented cortical BCS cells; (b) color-sensitive FCS cells that communicate with chromatically similar, but spatially disjoint, FCS cells; and (c) reciprocal BCS  $\leftrightarrow$  FCS interactions. The earliest possible cortical stage at which this could occur is at the blobs and interblobs of area V1. Using extracellular injections of HRP, Livingstone and Hubel (1984) reported blob-blob spatial interactions and interblob-interblob spatial interactions. However, no blob-interblob interactions were detected by this technique. Cross-correlational analyses have shown that the blob-blob interactions are color-specific, that the interblob-interblob interactions are orientation-specific, and that blob-interblob interactions do occur (Ts'o, 1989). Thus, the earliest possible cortical stage for the predicted BF intracopy and FB intercopy interaction is between the blobs and interblobs. The next possible cortical stage is between the thin stripes and interstripes. The cortical stage in question must, however, also have the property that (d) the monocular illuminant-discounted FCS signals are copied into multiple monocular FIDOS via monocular FF interscales. If the blobs themselves are not subdivided into subsets that can be selectively activated by different binocular disparities, then the thin stripes should be investigated as possible monocular FIDOS, with interstripe-to-thin-stripe pathways as the BF intracopies, blob-to-thin-stripe pathways as the monocular FF intercopies, and thin-stripe-to-interstripe pathways as the FB intercopies. The latter interpretation seems most likely at the present time.

In addition to these FF, BF, and FB interactions, *binocular FF intercopies* are predicted to occur along the pathways labeled 3 in Figure 11a. Both excitatory and inhibitory output signals are generated, as in the case of FB intercopies. The excitatory signals from each eye activate binocular FIDOS that correspond to the same disparity and position. The inhibitory signals suppress binocular FIDOS that correspond to smaller disparities at the same position. These interactions obliterate the brightness and color signals that could otherwise erroneously fill-in surface representations of occluded objects in the regions where they are occluded. The surviving excitatory signals from both eyes are binocularly matched to trigger the filling-in of the 3-D surface representation. These binocular FF intercopies occur within the blob cortical stream.

Why the *excitatory* binocular FF intercopies arise from the same source of illuminant-discounted FCS signals as the monocular FF intercopies is clarified below. In contrast, the *inhibitory* binocular FF intercopies arise from the edges of the filled-in connected regions within the monocular FIDOS, as do the FB intercopies. The excitatory

binocular FF intercopies form a one-to-many map to the binocular FIDOS. They are positionally aligned among the binocular FIDOS using BCS  $\rightarrow$  FCS boundary signals as teaching signals. These are the BF intercopies that are described below, which were used in Grossberg (1987c) to help explain data about binocular transfer of the McCollough effect. The positions of the inhibitory binocular FF intercopies are defined by the allelotropically shifted BF intracopies that define the filled-in domains whose edges activate them. The inhibitory FF intercopies also converge upon the binocular FIDOS, where they suppress FCS signals that would otherwise trigger the filling-in of occluded regions.

The final interactions are called *BF intercopies*. These are the BCS  $\rightarrow$  FCS boundary signals from a given disparity and position that add to the BCS boundaries of all smaller disparities at that position (Figure 3b), in order to prevent all nearer occluding surfaces from appearing transparent due to the filling-in of their positions by the brightness and colors of farther occluded surfaces.

We now develop a more detailed explanation of the data summarized in Sections 2-10. Sections 27-37 use BCS and FCS interactions in a pictorial way to explain these and related 3-D data. The remaining sections describe the computational principles and mechanisms that subserve these interactions in greater detail and use them to explain more data.

## 27. An Explanation of Bregman-Kanizsa Figure-Ground Separation and Completion

First let us consider how the occluded gray Bs in Figure 6 are seen and recognized on a surface behind the occluding black bands. Consider the image in Figure 27a. The white/black contrast of the occluding black band with respect to the white background is greater than the white/gray and gray/black contrasts caused by the occluded B shapes. As a result, the activation of BCS simple cells is greater at the white/black contrasts than at the white/gray and gray/black contrasts (Figures 27b and 28b). These monocular simple cells activate binocular complex cells. Since the image is viewed by both eyes at a distance, it generates a binocular disparity at each image point. This disparity increases with retinal distance from the foveation point. Larger disparities further from the foveation point and smaller disparities closer to the foveation point may all correspond to the same planar image. It is shown below how all these disparities are combined to generate a planar surface percept that corresponds to the same relative depth from the observer by using properties of the cortical magnification factor (Section 41). For present purposes, let  $D_1$  represent the set of all disparities that correspond to the planar image surface when it is binocularly viewed by an observer.

In Figures 27c and 28c, the larger receptive field size represents the largest scale that can binocularly fuse disparity  $D_1$ . Complex cells at the same position and scale compete across disparities via BB intrascales. The active cells corresponding to larger scales win the competition.

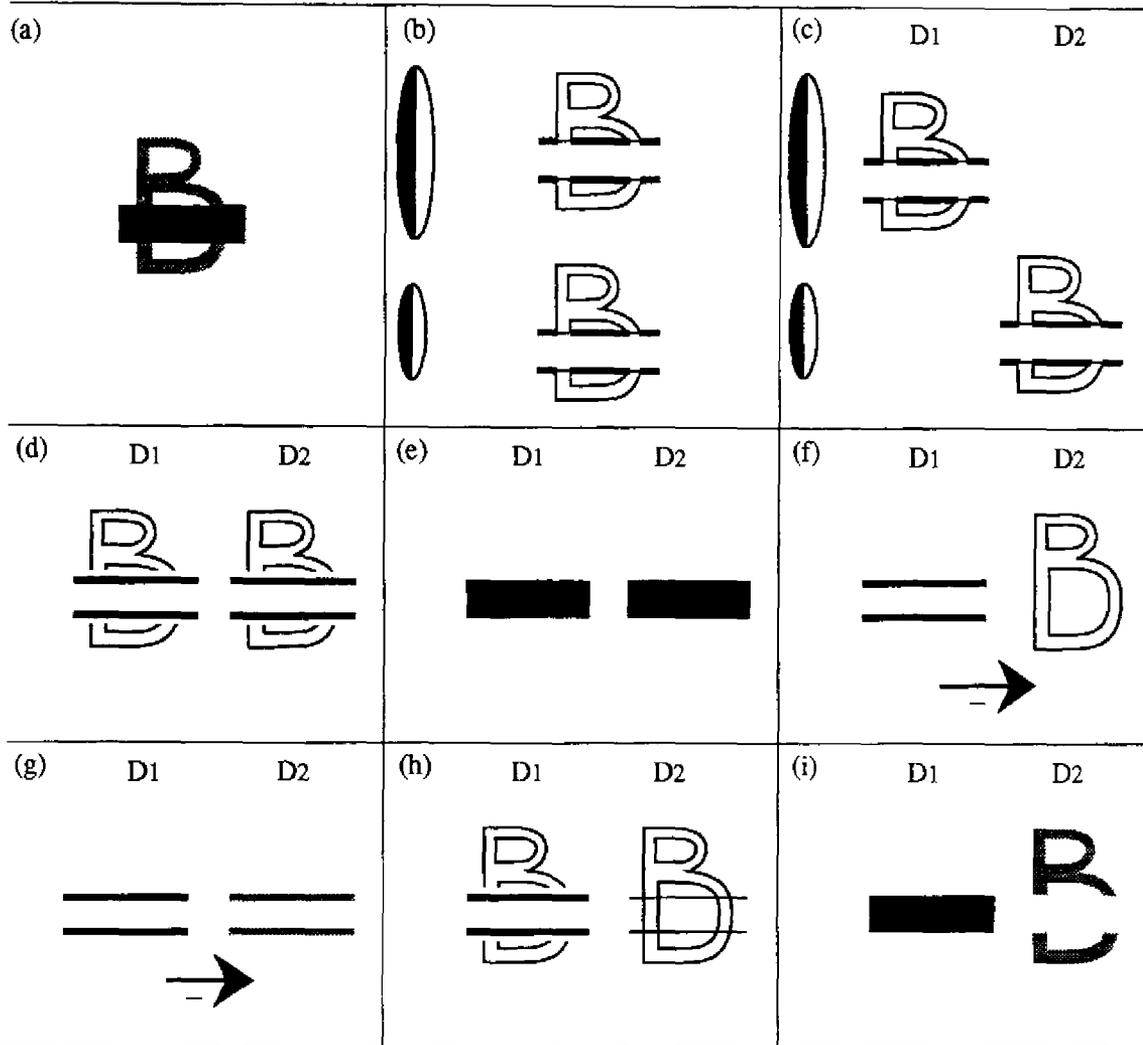


Figure 27. Bregman-Kanizsa figure-ground separation: (a) Image; (b) monocular simple cell activations in the boundary contour system (BCS); (c) complex cell activations after BB intrascale competition from disparity  $D_1$  to  $D_2$ ; (d) cooperative-competitive (CC) loop boundary segmentation at higher order hypercomplex cells after end gaps form; (e) filling-in of connected components in monocular feature contour system (FCS) syncytia; (f) FB intercopy inhibition to smaller scales and disparities, and CC loop reorganization of the B boundary; (g) FF intercopy inhibition to smaller scales and disparities; (h) BF intercopy inhibition adds boundaries to smaller scales and disparities; (i) filling-in of connected components in binocular FCS syncytia.

(Such a multiscale disparity-sensitive competition was computationally simulated in Grossberg and Marshall, 1989.) As a result of this competition, no complex cells fire at the smaller disparity,  $D_2$ , of the larger scale. On the other hand, smaller scales cannot binocularly fuse as wide a range of disparities as can larger scales. This property is due to the size-disparity correlation (see Section 38). The smaller scale in Figure 27c was chosen so that it could not fuse  $D_1$  but could fuse the slightly smaller disparity,  $D_2$ . Because disparity cells are coarsely coded before BB intrascale competition takes place, the smaller scale complex cells that are tuned to disparity  $D_2$  can re-

spond to the image contours. This can happen because there are no smaller scale complex cells that can fuse disparity  $D_1$ , and thus there is no BB intrascale competition from disparity  $D_1$  to disparity  $D_2$ . Thus, Figure 27c results from three properties: (a) a size-disparity correlation for binocular fusion; (b) coarse-coded nonzero disparity computations at binocular complex cells; and (c) competitive sharpening of disparity-sensitive complex cell responses within each scale, with larger fusible disparities winning over smaller ones.

Figures 27d and 28d show that end gaps are formed at the B boundaries as a result of CC loop feedback. Both

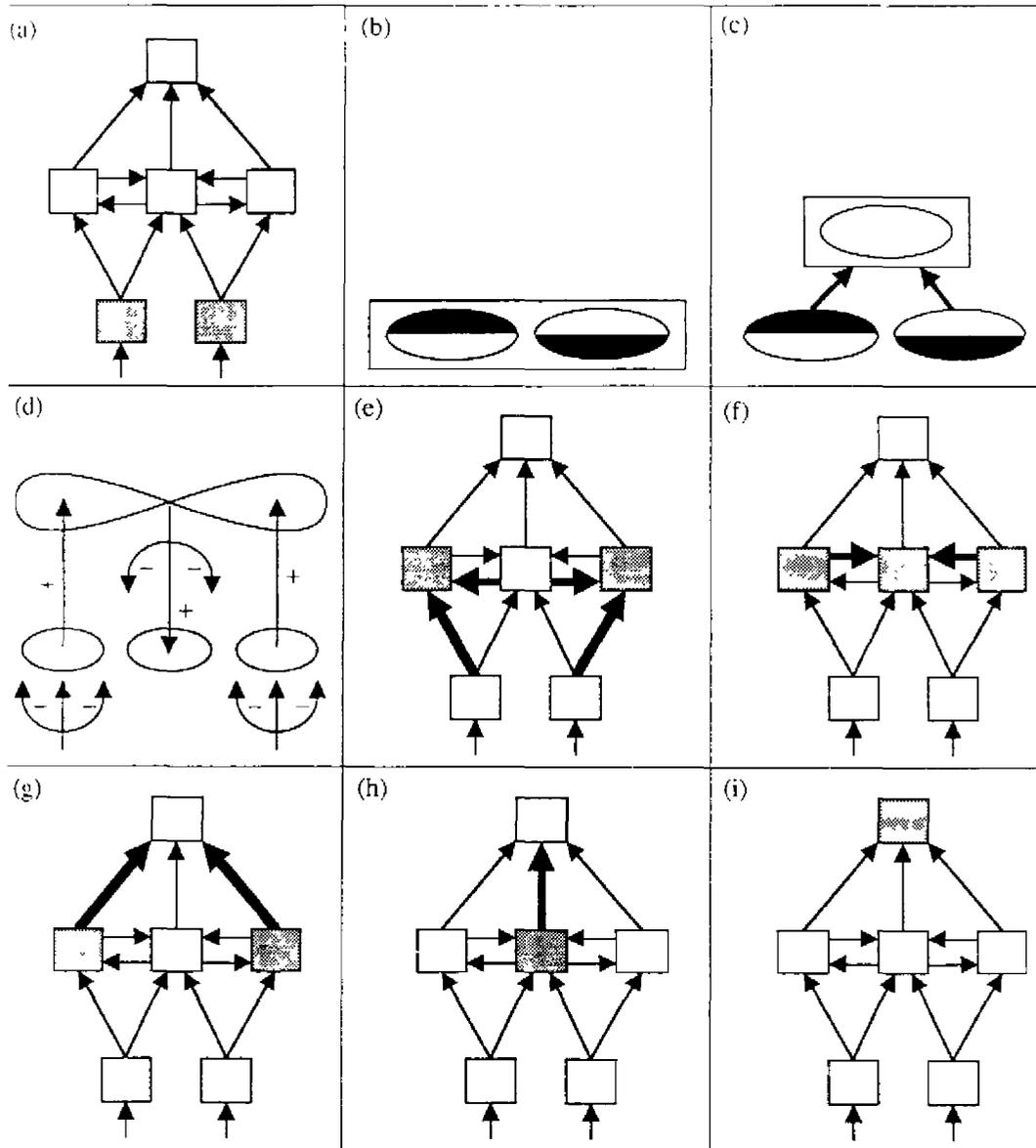


Figure 28. Active network stages during the processing of a 3-D scene: (a) Discounting of the illuminant occurs in the monocular preprocessing stages, notably the lateral geniculate nucleus; (b) simple cell activation; (c) complex cell activation; (d) emergent boundary segmentation by hypercomplex-bipole-cell feedback in the cooperative-competitive (CC) loop; (e) filling-in of the monocular synctia by monocular feature contour system (FCS) signals that are consistent with the binocular boundary contour system (BCS) segmentation; (f) FB intercopies inhibit boundaries at smaller scales and disparities; (g) FF intercopies excite filling-in of the corresponding binocular synctia and inhibit monocular FCS signals at smaller disparities; (i) the final multiscale filled-in surface representation of Form-And-Color-And-Depth (FACADE) emerges within the binocular synctia.

top-down bipole-to-hypercomplex competition between positions and hypercomplex-to-hypercomplex competition between orientations help to create these end gaps, as in Sections 17 and 19.

In Figures 27e and 28e, binocular BCS boundaries interact with monocular FCS signals via BF intracopies and monocular FF intercopies to select the monocular FCS

signals that are consistent with the binocular BCS boundaries. BCS boundaries hereby act as *filling-in generators* within the FCS; see Figure 11 and Sections 44 and 45. All other monocular FCS signals are suppressed. The selected FCS signals fill-in their respective filling-in domains, or synctia. If end gaps in the regions exist, as in Figure 27d, then the filling-in signals cross the gaps

and dissipate across space unless they are contained by other nearby boundaries, as they are in Figure 21 but not in Figure 27e. Figure 27e shows that only the boundaries of the black occluding region can contain the filling-in process during the first phase of the processing cycle.

Each filled-in connected FCS region generates contour-sensitive output signals, as in Figures 27f and 28f. Output signals are hereby generated only at the edges of the black occluder. These FCS output signals activate parallel pathways that influence both the BCS and the FCS. The BF intercopies inhibit any BCS boundaries that may exist at the same positions and orientations of smaller disparities and scales. In particular, the boundaries of the black occluder are inhibited at disparity  $D_2$ . After this happens, the incomplete B boundaries at disparity  $D_2$  can be colinearly completed by its CC loop, as in Figure 27f. These completed B boundaries generate direct BCS  $\rightarrow$  ORS signals, as in Figure 7. Thus, a completed letter B can be recognized at the ORS, even if only its unoccluded surfaces are seen at the FCS.

Why is the letter B not completely seen at the FCS? This is due to the binocular FF intercopies, which delineate the monocular surface components that are consistent with the binocular BCS segmentation. These monocular output signals are binocularly matched at the binocular FIDOS. This excitatory binocular interaction matches monocular signals that code the same position, disparity, and color. These are the FCS signals that trigger filling-in of a multiscale representation of Form-And-Color-And-DEpth at the binocular FIDOS. In addition, binocular FF intercopies inhibit all the FCS signals at their position which correspond to smaller disparities. As a result of these inhibitory FF intercopies, a surface that is filled-in at a nearer disparity cannot also be filled-in at a farther disparity unless suitably configured end gaps exist that generate a percept of transparency, as illustrated in Section 30.

The excitatory binocular FF intercopies are derived from the same FCS source as the monocular FF intercopies. These FCS signals fill-in the black occluder at disparity  $D_1$ , and the unoccluded part of the gray B at disparity  $D_2$ , as in Figure 27i. They do not fill-in gray color within the occluded region of the completed B shape in Figure 27f because they are derived from a processing stage prior to the monocular FIDO at which the complete B boundary is first input to the FCS. On the other hand, without further processing, the binocular FF intercopies would also fill-in the black occluder at disparity  $D_2$ . This additional processing is provided by the inhibitory binocular FF intercopies, which generate inhibitory signals from the monocular FIDOS to the binocular FIDOS. In particular, the inhibitory signals from the edges of the filled-in black occluder at disparity  $D_1$  of the monocular FIDO inhibit the excitatory binocular FF intercopies at the binocular FIDO that would otherwise fill-in the black occluder at disparity  $D_2$ .

Why can FCS signals from smaller disparities, such as the occluded gray shape at disparity  $D_2$ , not fill-in behind a nearer occluding surface, such as the black occluder at disparity  $D_1$ ? This is due partly to BF intercopies, which

add their boundary signals to the binocular syncytia of smaller disparities, as in Figures 27h and 28h. These BF intercopies are inhibitory signals. Inhibitory signals to an FCS syncytium create barriers to filling-in at their target cells (Cohen & Grossberg, 1984; Grossberg, 1987b; Grossberg & Todorović, 1988). As a result, in Figure 27h, complete boundaries of both the occluding band and the occluded B exist at the smaller disparity.

The BF intercopies and binocular FF intercopies of Figures 27g and 27h thus work together to generate the binocular filling-in events shown in Figures 27i and 28i. Due to BF intercopies, the B surface is filled-in at disparity  $D_2$  only where it is not occluded. Due to inhibitory binocular FF intercopies, the occluding surface is not filled-in at all at disparity  $D_2$ . The occluding surface is filled-in at disparity  $D_1$  because its FCS signals match BCS boundary signals that completely enclose them in connected regions. Because  $D_1 > D_2$ , the black occluding surface appears to be closer than the gray occluded B surface.

## 28. An Explanation of Da Vinci Stereopsis

The same mechanisms can now be used to explain the 3-D percept of the Da Vinci stereopsis image in Figure 2, with one addition: the interaction of near-zero disparity cells will be emphasized. Figure 29 outlines the main steps of the explanation. Figure 29a depicts the left (L) and right (R) eye views. It is assumed that viewing conditions enable the vertical edges A and B to be binocularly fused with disparity  $D_1$  and the vertical edges within region CD to be binocularly fused with disparity  $D_2$ , using the disparity convention of Section 27 for edges on the same planar surface. These fused boundaries are represented in Figure 29b. The larger scale is the largest scale that can just fuse  $D_1$ . The smaller scale is the largest scale that can just fuse  $D_2$ . Figure 29b shows the complex cell activations at both scales and disparities.

Consider the larger scale first. Because this scale can fuse edges A and B at disparity  $D_1$ , BB intrascales inhibit activation of  $D_2$  disparity cells by these edges. The  $D_2$  disparity cells can, however, fuse the vertical edges within region CD. Now consider the smaller scale. It can optimally fuse the CD vertical edges. It cannot fuse disparity  $D_1$ , but it can fuse  $D_2 < D_1$ . Edges A and B thus activate the  $D_2$  disparity cells, albeit less strongly. These activations are not inhibited by responses at larger disparities, because the smaller scale has no cells that are maximally tuned to these larger disparities.

None of the complex cell activations in Figure 29b form a connected boundary. This problem is overcome by using output signals from the separate pool of near-zero disparity cells. Adding the activations of near-zero disparity cells to all the pools of nonzero disparity cells, as in Figure 29c, does create some connected boundaries. Some of these near-zero activations are caused by horizontal edges. Others are caused by monocular viewing by the right eye of region BC. The image representation in Figure 29c assumes that allelotropia has deformed the binocularly

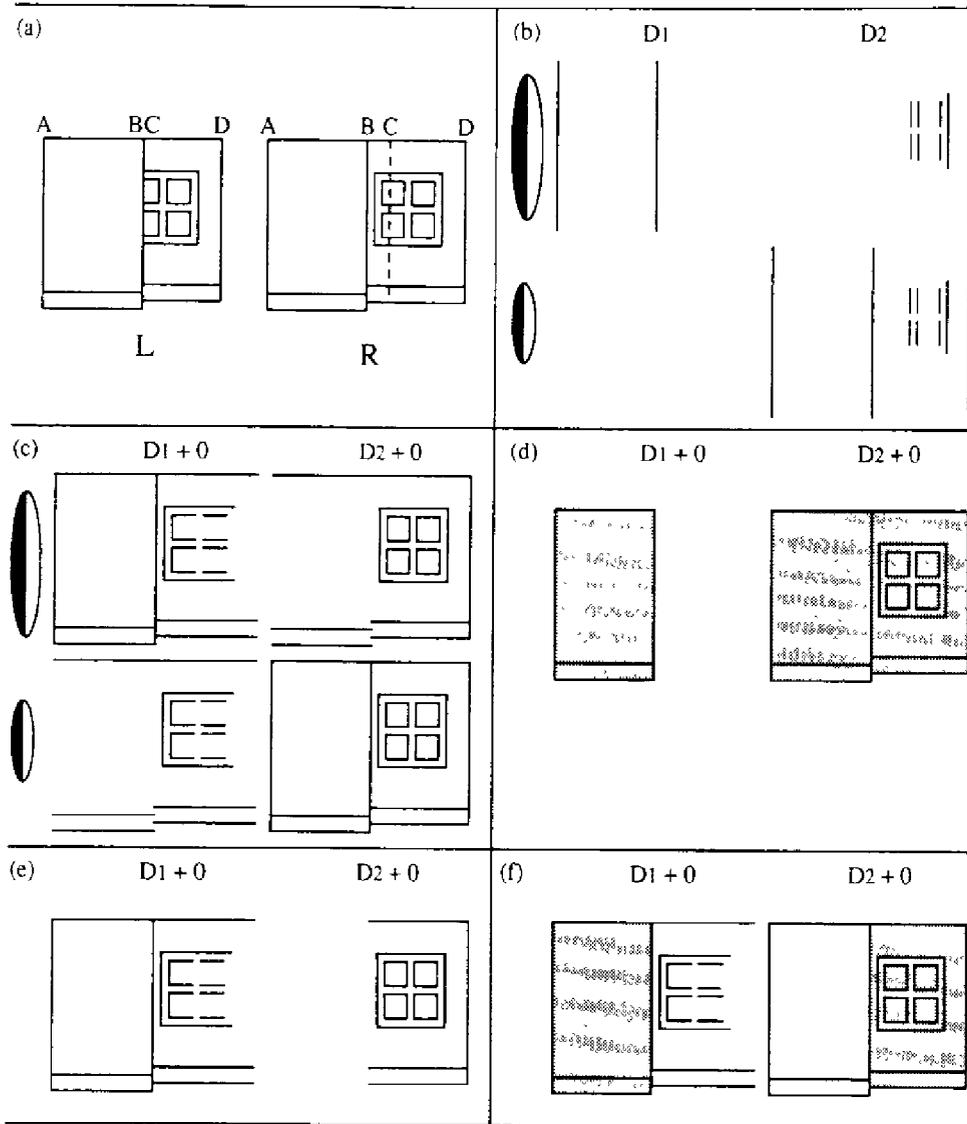


Figure 29. (a) Left and right eye views of a scene: region AB is closer than region BCD, and region BC is monocularly viewed; (b) binocularly fused complex cell responses of nonzero disparity cells at two scales and disparities; (c) combination of fused nonzero disparity responses with near-zero disparity responses to horizontal and monocularly viewed edges; (d) filling-in of connected regions; (e) deletion of boundaries at smaller scales and disparities due to FB intercopies; (f) overlay of final boundary contour system (BCS) boundary representation and filled-in surface representations at the binocular syncytia.

viewed regions AB and CD in such a way that the monocularly viewed region BC can fit in between. In situations where this is not true, binocular rivalry can ensue, as described in Grossberg (1987c).

The CC loop does not substantially change the boundary representation of Figure 29c except to attach endpoints of allelotropically shifted edges to near-zero disparity edges. Boundaries are not completed in the  $D_1$  representation because inhibition from  $D_2$  disparity cells propagates

into the CC loop via complex off-cells and hypercomplex off-cells (Grossberg, 1991).

Figure 29d indicates the regions of Figure 29c that can be successfully filled-in within the monocular syncytia, as in Figure 28e. Figure 29e describes the boundaries that survive the inhibition due to FB intercopies, as in Figure 28f. A similar inhibition of FCS signals for region AB occurs at disparity  $D_2$  due to binocular FF intercopies, as in Figure 28g. Figure 29f shows the effect of BF inter-

copies on the final connected boundary segmentations, as in Figure 28h, and the final filling-in of the binocular syn-  
 cytia, as in Figure 28i. Surface AB selectively fills-in at  
 disparity  $D_1$ , and surface BCD selectively fills-in at  
 disparity  $D_2$ . The ambiguous region BC hereby inherits the  
 depth of region CD.

**29. An Explanation of the Closer Appearance  
 of Higher Spatial Frequencies Than of  
 Lower Spatial Frequencies**

An explanation of the depthful spatial frequency per-  
 cepts that were described in Section 3 can also be derived

from these mechanisms. The explanation begins by not-  
 ing that a high spatial frequency sinusoid activates a large  
 receptive field more than does a low spatial frequency  
 sinusoid, other things being equal, if the receptive field  
 is no larger than one-quarter of the sinusoidal period. This  
 is true because the luminance of the high spatial frequency  
 sinusoid increases more quickly across space, and thus  
 causes a larger contrast change per unit area, than does  
 the low spatial frequency sinusoid (Figure 30a). As a re-  
 sult, just as at the ends of the cross in the Redies-Spillmann  
 display of Figure 9a, the vertically oriented complex cells  
 that are activated by the high spatial frequency sinusoid  
 inhibit the contiguous vertically oriented hypercomplex

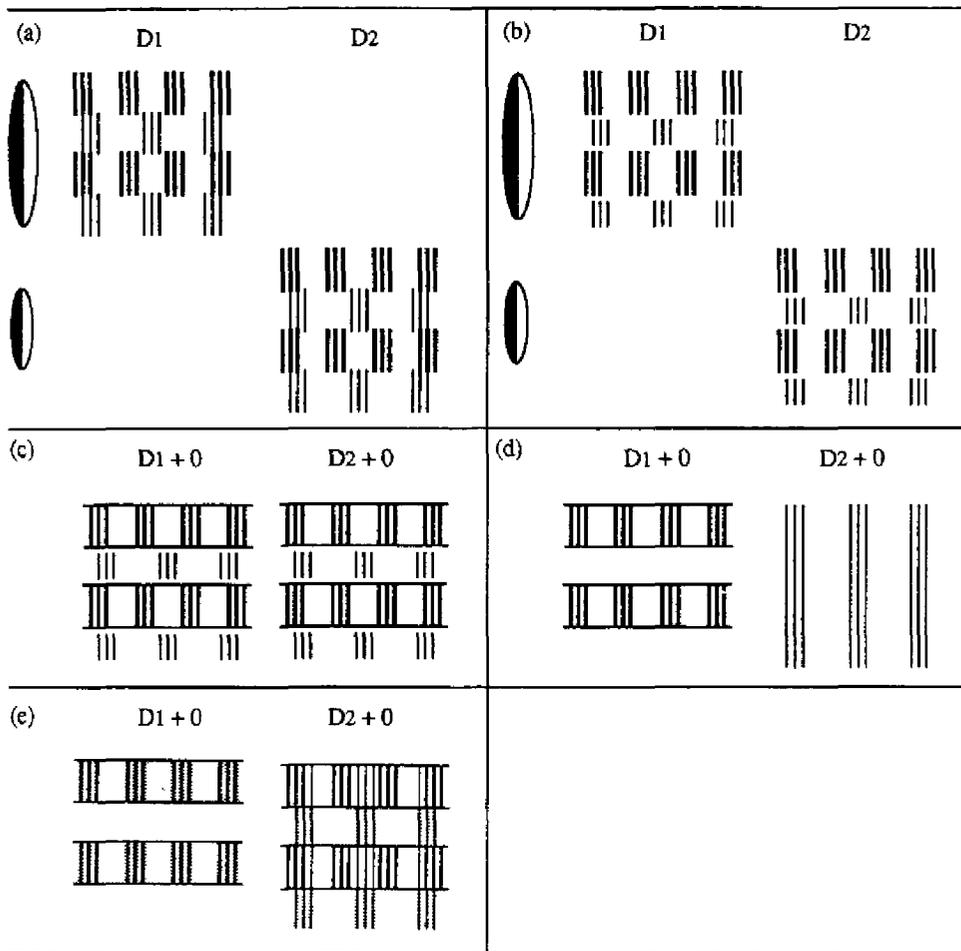


Figure 30. Why high spatial frequency inputs appear closer than low spatial frequency inputs: (a) Complex cells at the larger scale and disparity respond more strongly to the higher spatial frequency. The larger scale, smaller disparity cells do not respond due to inhibition from BB intrascales. The smaller scale and smaller disparity cells do respond because the smaller scale cannot fuse the larger disparity. (b) Effect of end gaps on the low spatial frequency activations of hypercomplex cells. (c) End cuts about the high spatial frequency activations of higher order hypercomplex cells. (d) FB intercopies from the filled-in large disparity feature contour system (FCS) copy to the smaller disparity boundary contour system (BCS) cells inhibit the high spatial frequency responses there, and enable the smaller disparity cooperative-competitive (CC) loop to complete vertical illusory contours among the low spatial frequency responses. (e) Filling-in takes place at the large disparity representation of the high spatial frequency input and at the smaller disparity representation of the low spatial frequency input. FF intercopies and BF intercopies prevent the latter representation from filling-in the high spatial frequency input.

cells that are activated by the low spatial frequency sinusoid, more than conversely. End gaps hereby begin to form at these locations (Figure 30b). These complex cells are activated by the continuously changing contrasts in the sinusoids. The activated cells generate a *boundary web* of form-sensitive boundary activations (see Figure 16 and Grossberg & Mingolla, 1987a, for computer simulations of boundary webs).

The asymmetric inhibition of hypercomplex cells at the first competitive stage enables the higher-order hypercomplex cells at the second competitive stage to form end cuts that bound the high frequency sinusoids (Figure 30b). The CC loop binds the stronger high spatial frequency activations and end cuts into an emergent boundary segmentation as it deepens the end gaps at the ends of the low spatial frequency sinusoids (Figure 30c). The CC loop hereby generates an emergent boundary segmentation that builds closed compartments out of horizontal boundaries and high spatial frequency vertical boundaries, but also opens end gaps between the horizontal boundaries and the vertical low spatial frequency boundaries.

FB intercopies from the larger disparity  $D_1$  inhibit the closed compartments at the smaller disparity  $D_2$ . The surviving lower spatial frequency vertical boundaries can hereupon use the CC loop at disparity  $D_2$  to colinearly complete vertical boundaries over the regions that were previously occluded by the high spatial frequency sinusoid (Figure 30d). These completed low spatial frequency boundaries can be recognized via the direct BCS → ORS pathway (Figure 7). Binocular FF intercopies and BF intercopies act next to complete surface properties of the high spatial frequency sinusoids at disparity  $D_1$  and of the low spatial frequency sinusoids at disparity  $D_2$  (Figure 30e). Hence, the high spatial frequency surface looks closer than the low spatial frequency surface.

This explanation also clarifies how the depth percept can reverse itself through time. This can be explained, without changing the theory, by invoking two additional theoretical mechanisms that are in the right place to do the job. These mechanisms control spatial frequency adaptation and attention shifts. Habituated transmitter gates exist in the pathways to the hypercomplex cells of the second competitive stage and in the bipole cell feedback pathways (Grossberg, 1987c). These are the same habituated transmitter gates that help to limit visual persistence and to trigger reset of a boundary segmentation when stimulus conditions change (Section 21). In the present example, if the habituation attenuates the initially more active high spatial frequency activations until they fall below the low spatial frequency activations, then the end gaps will switch to the high spatial frequency locations and the depth percept will flip. When the low frequency transmitter gates habituate, another depth flip can occur, and so on cyclically thereafter, with the advantage of the high frequency scale showing in its more persistent percept as a nearer figure. This is a preattentive mechanism for a bistable depth reversal.

A spatial attention mechanism can also operate via ORS → BCS feedback pathways (see Figure 7 and Section 49)

to influence such a bistable depth percept. A shift in attention can prime the CC loop of one part of the image more than another part. Such a top-down prime can amplify the attended CC loop activations. A sufficiently large amplification of the low spatial frequency boundaries could reverse the position of the end gaps, and hence the relative depth percept.

### 30. An Explanation of 3-D Neon Color Spreading and Transparency

The theory will now be used to explain the Nakayama, Shimojo, and Ramachandran (1990) demonstrations of 3-D neon color spreading and transparency (see Section 6), as well as 3-D percepts where opaque rather than transparent percepts emerge. These explanations clarify how the normal role of filling-in for surface synthesis can sometimes lead to paradoxical percepts of transparency.

Figure 31 depicts the network stages that account for a percept of a neon disk in front of an occluded cross lying on a more distant surface. Figure 31a shows the responses of the vertically oriented nonzero disparity complex cells at two scales and disparities that were chosen as in the previous examples. The darker lines signify the large complex cell activations due to white/black contrasts. Thinner lines represent weaker activations at the white/gray and gray/black contrasts where the gray cross meets the black background and the four Ehrenstein limbs. Note how sparse these cues are for purposes of continuous surface perception. Figure 31b adds signals from the near-zero disparity cells. As a result, connected boundaries form at the  $D_2 + 0$  disparity cell pool.

The hypercomplex and bipole cells of the CC loop respond to the larger scale activations by forming vertical end cuts at the horizontal ends of the cross. These vertical hypercomplex cells cooperate via bipole cells with the horizontal hypercomplex cells at the vertical ends of the cross (Figure 31b). Circular illusory boundaries are hereby formed. The vertical hypercomplex cells do not cooperate as well with the horizontal hypercomplex cells at the distal ends of the vertical Ehrenstein limbs for two reasons: (a) these Ehrenstein limbs are farther away, and (b) they form an angle with respect to the vertical end cuts that cannot colinearly cooperate across all four sets of inducers. End cuts also cannot form among the  $D_2$  disparity cells because of BB intrascale inhibition from the corresponding  $D_1$  disparity cells, and the propagation of this inhibition via complex off-cells and hypercomplex off-cells into the CC loop of the  $D_2$  cells (Grossberg, 1991).

Due to the nonoverlapping nature of the complete connected boundaries in Figure 31c, FB intercopies from the larger scale do not inhibit any boundaries at the smaller scale. Binocular FF intercopies inhibit FCS inducers of the gray cross at its horizontal and vertical ends, but leave intact FCS inducers of gray color along the edges of the cross. The BF intercopies in Figure 31d lead to a complete  $D_2 + 0$  boundary representation of the cross, the Ehrenstein limbs, the illusory circle, and the background. Filling-in at the  $D_1 + 0$  binocular syncytium creates a cir-

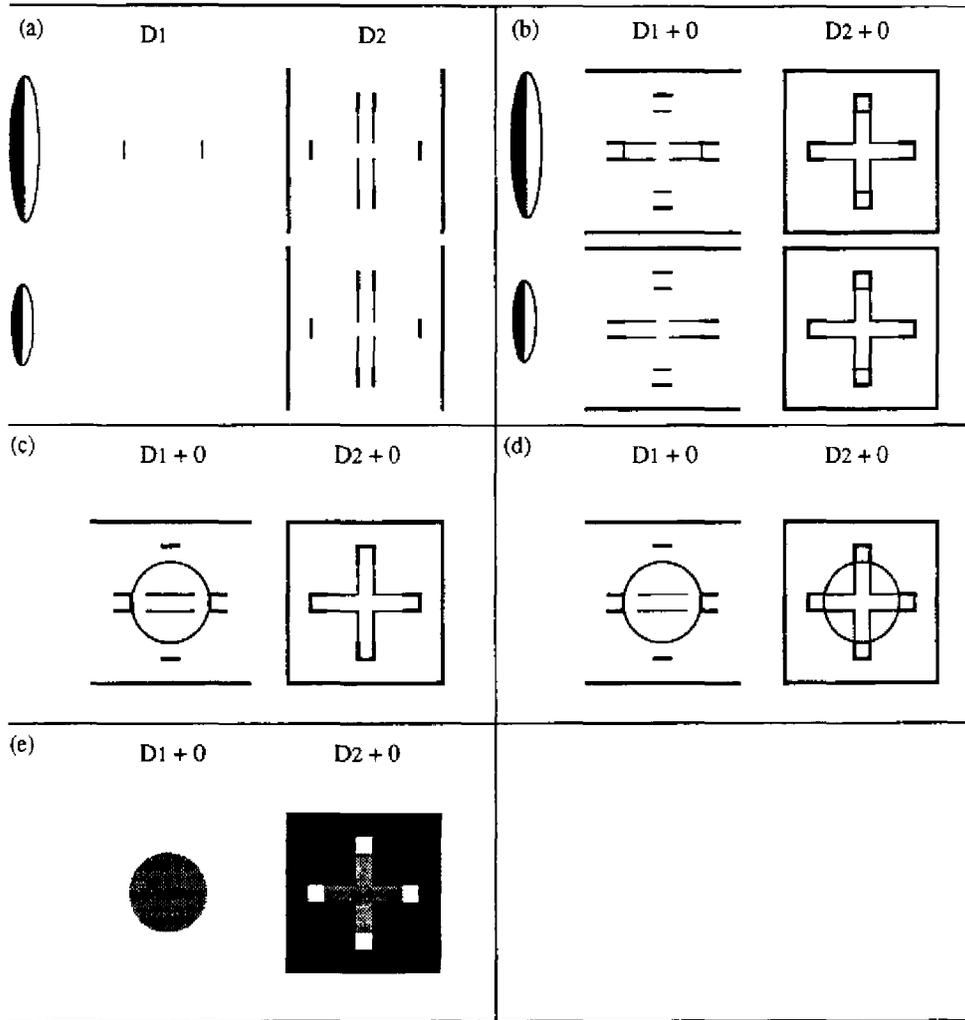


Figure 31. Explanation of 3-D transparency in response to the Redies-Spillmann stereogram: (a) Complex cell activations at two scales and disparities, (b) addition of near-zero disparity activations to those in (a), (c) cooperative-competitive (CC) loop activations, (d) complete boundary segmentations after BF intercopies act, (e) filled-in surface representations at the binocular Filling-In Domains (FIDOs) of the feature contour system (FCS). See Figure 11.

cular gray disk (Figure 31e). The black background, gray cross, and white Ehrenstein limbs are filled-in only at disparity  $D_2$ . A percept of a gray neon disk lying transparently in front of a gray cross on a black background is hereby generated.

### 31. An Explanation of 3-D Filling-In Without Transparency

Why is the neon disk no longer seen when its pair of vertical inducers at the horizontal ends of the cross code a farther away, rather than a nearer, disparity? Why is an observer nonetheless aware of a circular structure that is partially occluded by the cruciform black background?

Figure 32 illustrates the theory's explanation of this percept.

The combined activations of  $D_1$ ,  $D_2$ , and near-zero disparity complex cells are shown for two scales in Figure 32a. Note that a connected boundary can form only at the larger scale of the  $D_1 + 0$  cell pool. The CC loop at  $D_1$  cannot respond to the pair of horizontal Ehrenstein limbs to form vertical end cuts with which to generate a circular illusory contour (Figure 32b). This is due to BB intrascale inhibition among the larger scale complex cells. The  $D_1$ -tuned cells are inhibited by the  $D_2$ -tuned cells, which are maximally sensitive to disparity  $D_2$ . These larger scale  $D_2$  cells inhibit the responses of larger scale  $D_1$  cells at the corresponding positions throughout the

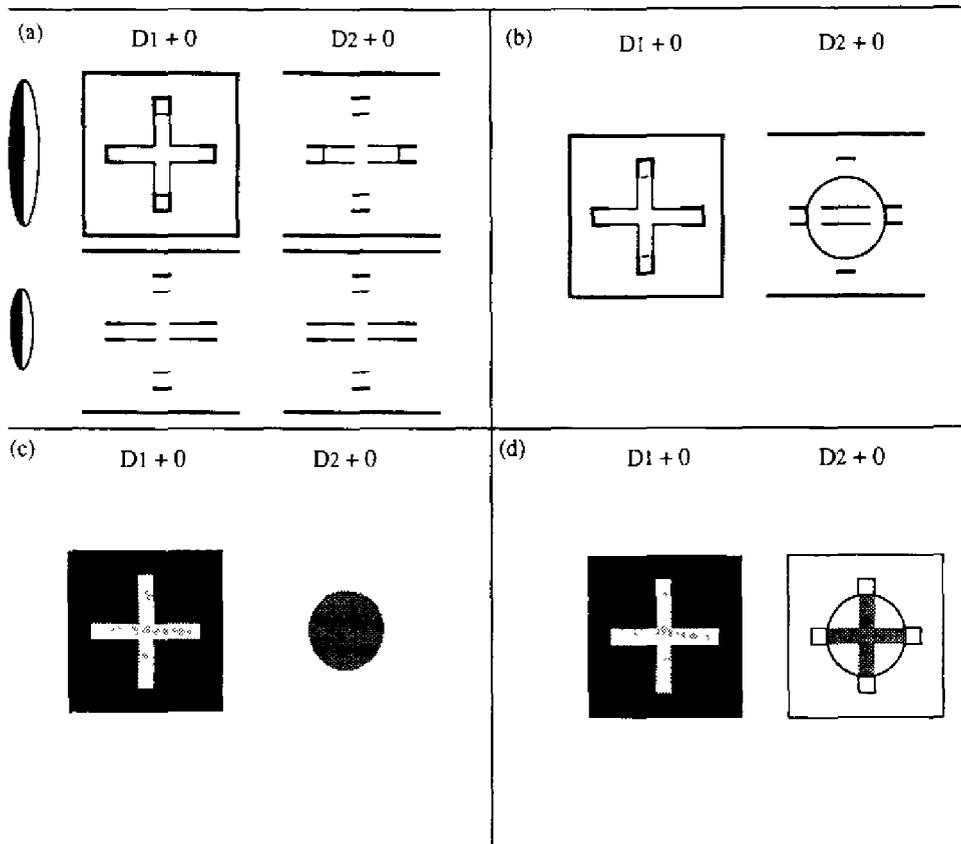


Figure 32. Reversal of relative depth eliminates percept of 3-D neon and transparency. Instead, an occluded circle is completed but only partly seen through an opaque occluding surface. See text for details.

complex-hypercomplex-bipole cell hierarchy. Such  $D_2$ -to- $D_1$  inhibition does not occur during 2-D percepts of the Redies-Spillmann display (Figure 9a), and thus does not prevent end gaps, end cuts, circular boundaries, and neon spreading from occurring in that situation.

In the 3-D layout of Figure 32, the absence of these vertical end cuts eliminates the source of cooperation with which the horizontal boundaries at the vertical ends of the cross formed an illusory circle in Figure 31. Instead, the vertical ends of the cross are subjected to conditions that typically form neon flanks (Grossberg, 1987b, Figure 4; Redies & Spillmann, 1981). Here, the strongly activated vertical hypercomplex cells at the ends of the vertical Ehrenstein limbs cooperate best with the vertical hypercomplex cells at the vertical ends of the cross, as they weaken the horizontal boundaries at the vertical cross ends (Figure 32b). A single connected boundary that surrounds both the cross and the Ehrenstein limbs is hereby created, with weak boundaries at the vertical limb-cross interfaces. In contrast, at the  $D_2 + 0$  cells, a circular illusory boundary can be formed through the four limb-cross interfaces (Figure 32b). Vertical illusory boundaries

are not formed perpendicular to the horizontal boundaries, because they are inhibited by BB intrascales from the  $D_1 + 0$  cells at the corresponding positions. The connected components that fill-in the monocular FIDOs are depicted in Figure 32c.

FB intercopies inhibit some of the Ehrenstein limb boundaries at the  $D_2$  cells, but not the vertical ones that support the circular illusory boundary after it forms. A circular boundary can thus be recognized via the BCS - ORS interaction at the depth corresponding to  $D_2$  (Figure 7). Binocular FF intercopies spare essentially all the FCS inducers of the gray cross color and some inducers of the white limb color. The  $D_1$  boundary at the binocular FIDO can support filling-in of a black background around the cross in its binocular syncytium (Figure 32d). A light gray color can also fill-in within the cross-plus-limb region. BF intercopies form a complete boundary representation at the  $D_2$  binocular FIDO around the cross, the Ehrenstein limbs, the circle, and the background (Figure 32d). This  $D_2$  boundary can support filling-in of the gray cross and the white Ehrenstein limbs. The final percept is one of a farther away gray cross on a white surface

that is partially occluded by a closer black region. This percept is augmented by direct BCS – ORS recognition of the complete circular boundary that joins the four ends of the cross.

Comparison of Figures 31 and 32 shows how changing the disparity of image edges that subtend infinitesimal areas of a scene can alter the percept of surfaces that subtend large scenic areas. These examples dramatize the radical break that students of perception need to make with classical mathematical notions of 3-D geometry in order to understand our percepts of the 3-D world.

### 32. An Explanation of 3-D Opaque Bars Instead of a Neon Disk

The next example further illustrates the context-sensitive nature of this perceptual geometry. As noted in Section 6, removal of the white Ehrenstein limbs abolishes the neon percept that was analyzed in Figure 31 and creates the percept of a horizontal bar in front of a vertical bar, with clear illusory horizontal boundaries across the vertical bar. An explanation of this percept is illustrated in Figure 33.

Figure 33a depicts the  $D_1+0$  and  $D_2+0$  boundaries that are induced by the cross and its background. The CC loop at  $D_1$  colinearly completes two horizontal boundaries and thereby forms a closed, connected, rectangular boundary (Figure 32b). The FB intercopies from this boundary to the  $D_2+0$  boundaries eliminate the corresponding horizontal boundaries at  $D_2+0$ . The remaining vertical boundaries are hereby freed to complete the two vertical boundaries and to thereby form another closed, connected, rectangular boundary. Binocular FF intercopies eliminate all FCS inducers at the  $D_2$  binocular FIDO of the horizontal bar that is subtended by the  $D_1$  connected boundary. BF intercopies add the two connected rectangular boundaries at the  $D_2$  binocular FIDO.

Figure 33c shows the result of filling-in at the binocular FIDOs. The  $D_1$  surface fills-in only a gray horizontal bar. The  $D_2$  surface fills-in the gray unoccluded region of the vertical bar, as well as the black unoccluded part of the background. The final percept is one of an opaque horizontal gray bar that partially occludes a more distant vertical gray bar on a black background. A comparison of Figures 31 and 33 clarifies the key role played by the

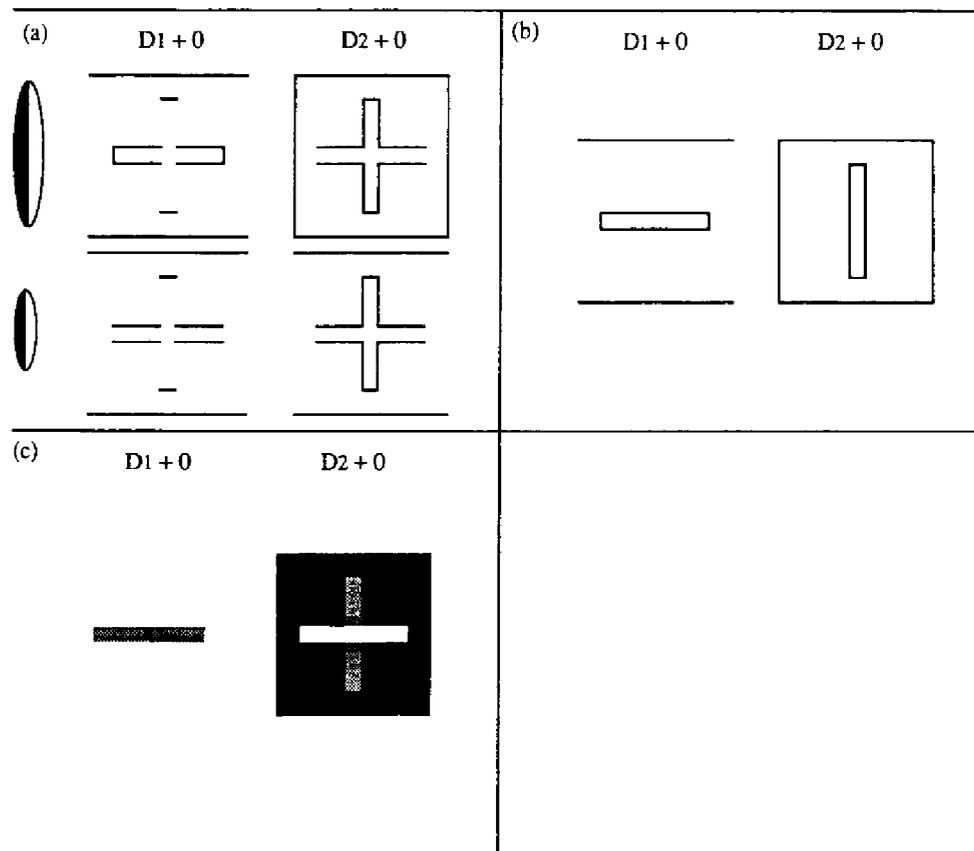


Figure 33. 3-D percept of a horizontal opaque bar occluding a more distant vertical bar, due to removal of the four Ehrenstein limbs. See text for details.

Ehrenstein limbs in creating the percept of a transparent disk, rather than that of an opaque bar.

### 33. An Explanation of Kanizsa-Varin Variations: Opaque or Transparent Square Occludes Four Disks on Background

The 3-D percepts using Kanizsa squares and Varin displays that were described in Section 7 will now be analyzed. These two types of images generate different complex cell activations, due to the absence in the Kanizsa square of gray wedges in the mouths of the pacman figures. Despite this difference, both images generate similar emergent boundary segmentations after the CC loops act. They do not, however, generate the same FCS signals for filling-in. As a result, the Kanizsa square does not induce a percept of transparency, but the Varin figure does.

Consider, first, a Kanizsa square with black pacman figures on a white background. Let the vertical boundaries of the pacman figures occur at the larger disparity,  $D_1$ . Figure 34a shows the  $D_1 + 0$  and  $D_2 + 0$  complex cell activations at two scales. Figure 34b shows how these activations induce emergent boundary segmentations via their respective CC loops. Figure 34c shows the connected boundaries after FB intercopies act. The square boundary at disparity  $D_1$  fills-in a white color at the binocular FIDO, as in Figure 34d. The FF intercopies induced by the square boundary eliminate the FCS inducers of the square at disparity  $D_2$ . The remaining FCS inducers at the  $D_2$  binocular FIDO fill-in the black wedges and the white background around the square. Hence, an opaque square is seen hovering above partially occluded black disks on a white background. The completed disk boundaries in Figure 34b also input directly from the BCS to the ORS (Figure 7).

The Varin display differs from the Kanizsa square in using white pacmen, gray wedges in the mouths of the pacmen, and a black background. As a result of the gray wedges, the circular disk boundaries do not have to be completed at disparity  $D_2$  using illusory contours of the CC loop. These contours are, instead, derived directly from the image by the simple cells. The boundaries induced by both images are similar at the CC loop stage, as in Figure 34b.

The filling-in events induced by Kanizsa and Varin images differ in a critical way, however. Figure 34e shows how filling-in at the binocular FIDOS generates a gray square surface at the disparity  $D_1$  binocular FIDO. At the disparity  $D_2$  binocular FIDO, the white pacmen fill-in, as does a black boundary outside the pacmen and square boundaries. These filling-in events are completely analogous to those in Figure 34d. In addition, the gray wedges also fill-in, due to the existence of uninhibited gray FCS inducers inside the round wedge boundaries. This filling-in event has no analog in Figure 34d. Another new filling-in event completes the black region that is surrounded by the square and the wedges. The black FCS inducers for this filling-in event exist outside the round gray wedge boundaries. In all, a surface is filled-in at disparity  $D_2$

that consists of a black background surrounding four white-and-gray disks. The final percept is of a transparent gray square in front of a black background in which four circular white disks are partially occluded by the gray square (Figure 34e).

This comparison between the Kanizsa and Varin percepts illustrates two important points: (1) Images that have different edges can have identical emergent boundary segmentations. (2) Images with identical emergent boundary segmentations can induce dramatically different percepts—even the difference between transparency and opacity—because they activate a different set of FCS inducers for filling-in surface properties.

### 34. An Explanation of Kanizsa-Varin Variations: Square Occluded by an Opaque Foreground With Four Circular Holes

A reversal of relative depth relationships in the Kanizsa square and Varin display creates a totally different percept (Section 7). In particular, no transparency occurs in response to the Varin display. Consider first the Kanizsa square. The  $D_1 + 0$  and  $D_2 + 0$  complex cell responses are shown in Figure 35a, and the CC loop boundary segmentations are shown in Figure 35b. The connected boundaries at the binocular FIDOS after FB intercopies and BF intercopies act are shown in Figure 35c.

The binocular FF intercopies inhibit only the FCS inducers that bound the circular disks and the outer square boundary of the background. As a result, the black pacmen and white wedges can fill-in the binocular FIDO at disparity  $D_2$ , as in Figure 35d. At disparity  $D_1$ , a white surface fills-in around the four circular apertures, and a light gray film—a mixture of black and white FCS inducers from around the inner circular boundaries—fills-in the four disks. The final percept is one of a near-white opaque surface with four circular apertures through which the four corners of a white square appear on a black background (Figure 35d). The completed square in Figure 35b is recognized via the BCS — ORS pathway (Figure 7).

The relative depth reversal of the image contrasts causes the Varin display to undergo essentially the same filling-in events as the Kanizsa square. The  $D_1 + 0$  boundary in Figure 35c does not include the square boundary, as it does in Figure 34c. Thus, the Varin wedges add nothing to the Kanizsa display but a different color for filling-in at disparity  $D_2$ . The binocular FIDO at disparity  $D_2$  fills-in white pacman figures and gray wedges by using the FCS inducers at the corners of the square (Figure 35e). Disparity  $D_1$  fills-in a black occluding surface with four circular apertures by using the black FCS inducers that abut the exterior background square and the exteriors of the circles. The circular apertures fill-in a light-gray film by mixing the white and gray FCS inducers at the inner circular boundaries. The final percept is one of a nearby black opaque surface with four circular apertures through which the four corners of a gray square appear on a more distant white background. The comparison between the transparent and opaque percepts summarized in Fig-

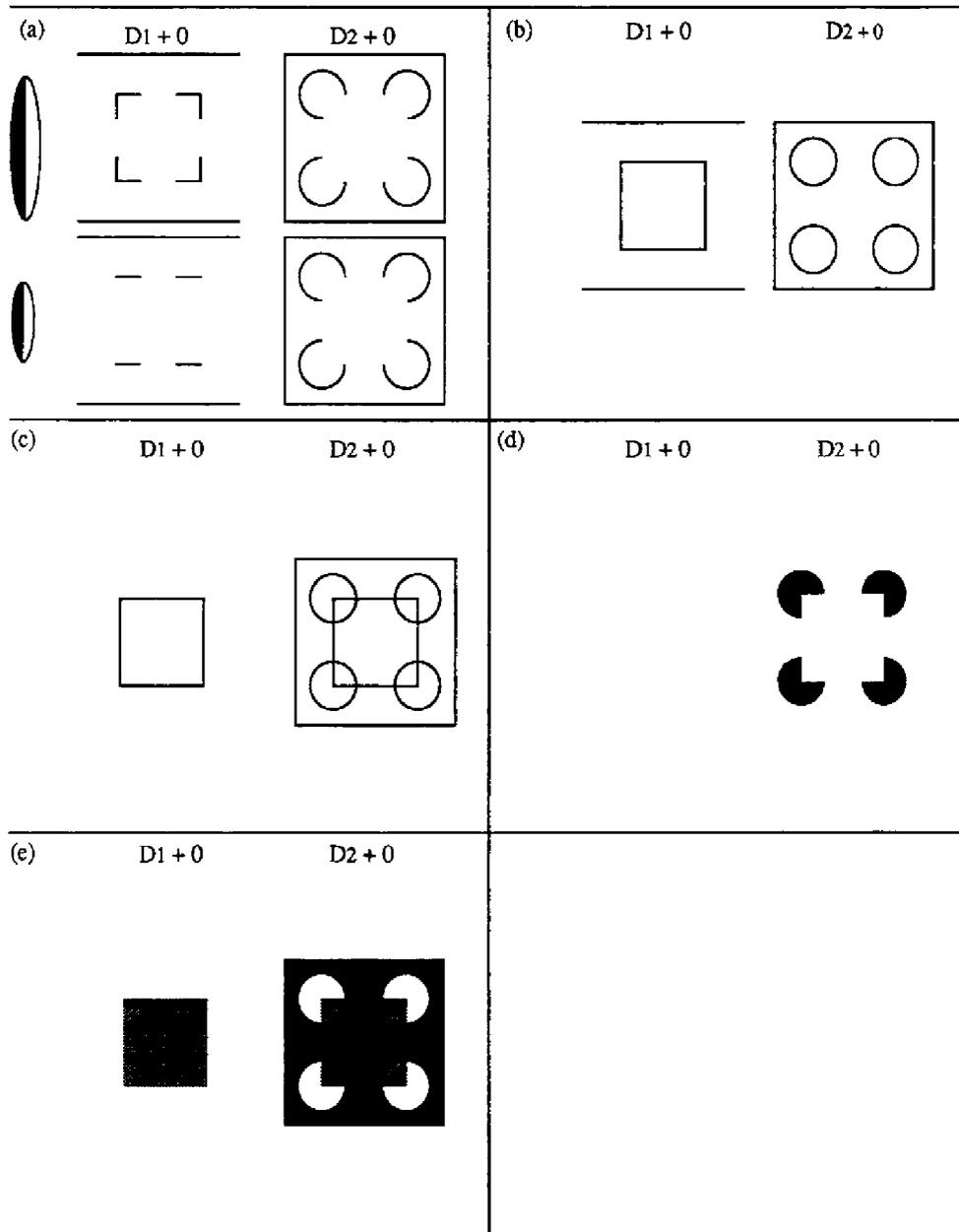


Figure 34. Perception of a 3-D Kanizsa square and transparent Varin display. See text for details.

ures 34 and 35 provides strong support for the theory's rules.

**35. Kanizsa-Varin Variations:  
Depth Reversal by Pacman Removal**

Nakayama et al. (1990) noted that removal of the pacman figures in a Varin display can cause a depth reversal in the final percept, even though "the disparity relations

of the various bounding contours ... are identical" (p. 506). With pacmen included, the display appears as a gray transparent square in front of four partially occluded circular disks (see Figure 34e). With pacmen removed, the display appears as a more distant square region upon which four wedges are seen through a closer square aperture.

Figure 36a shows the  $D_1 + 0$  and  $D_2 + 0$  complex cell activations after the pacmen are removed. Figure 36b

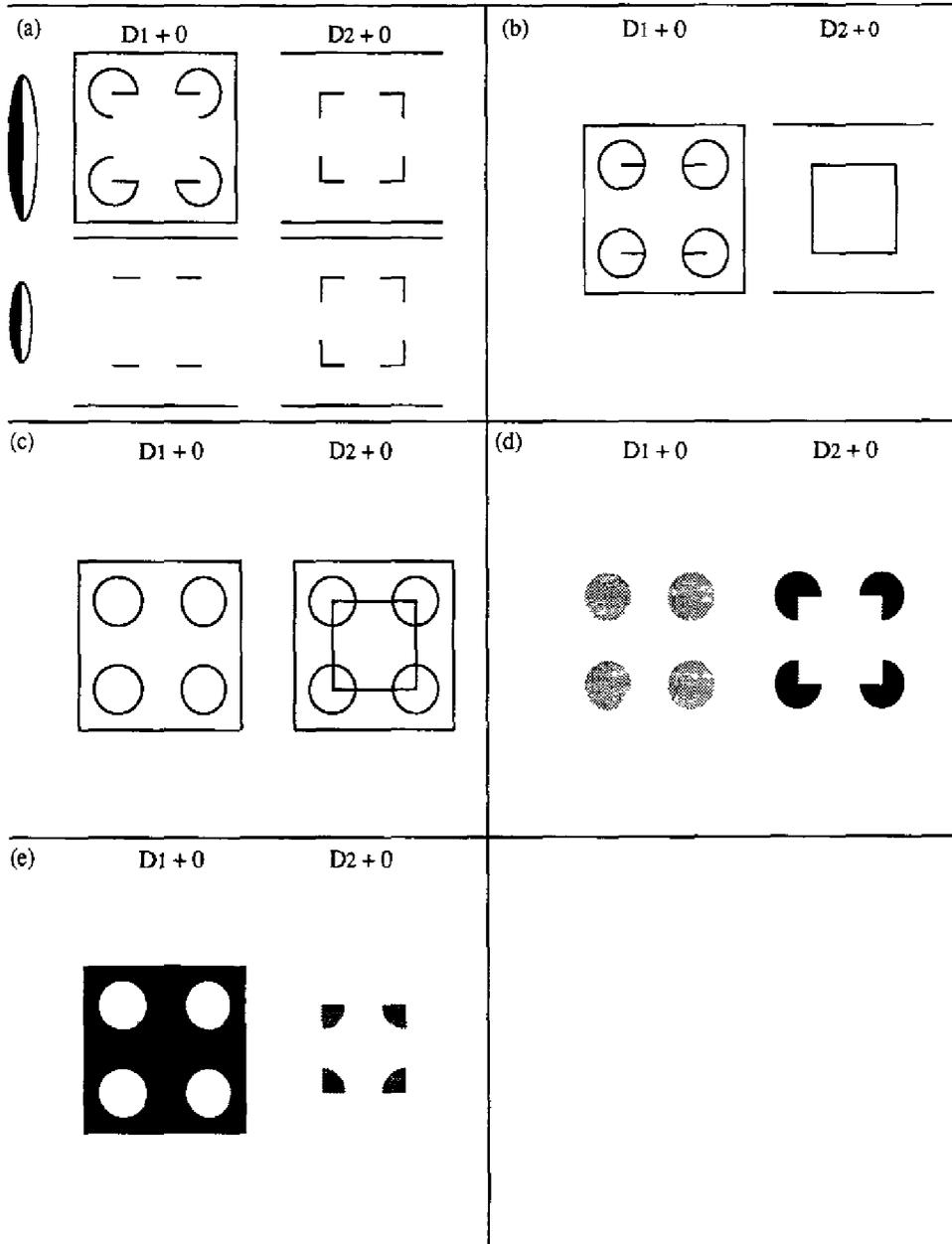


Figure 35. Perception of a 3-D Kanizsa square and opaque Varin display due to relative depth reversal. See text for details.

shows the CC loop boundary completions. For definiteness, it is assumed that the CC loop of the smaller scale, before interscale cooperation, cannot form vertical boundaries over such a large distance. This property does not affect the result, either way. Figure 36c shows the connected boundaries after FB intercopies and BF intercopies act. Figure 36d shows the filling-in events at the binocular FIDOS. At disparity  $D_1$ , the outer black frame fills-in, as does an inner rectangular film that mixes gray and black

FCS inducers. At disparity  $D_2$ , the FCS inducers on both sides of the curved wedge boundaries are spared by binocular FF intercopies. Hence, they can fill-in four gray wedges and a black region in between. The four wedges are thus seen through a square aperture in the black occluding surface. When humans perceive this image, the occluded wedges do not seem to exist beyond the square aperture. This property may be explained by the absence of boundaries in Figure 36c to complete the wedges into disks.

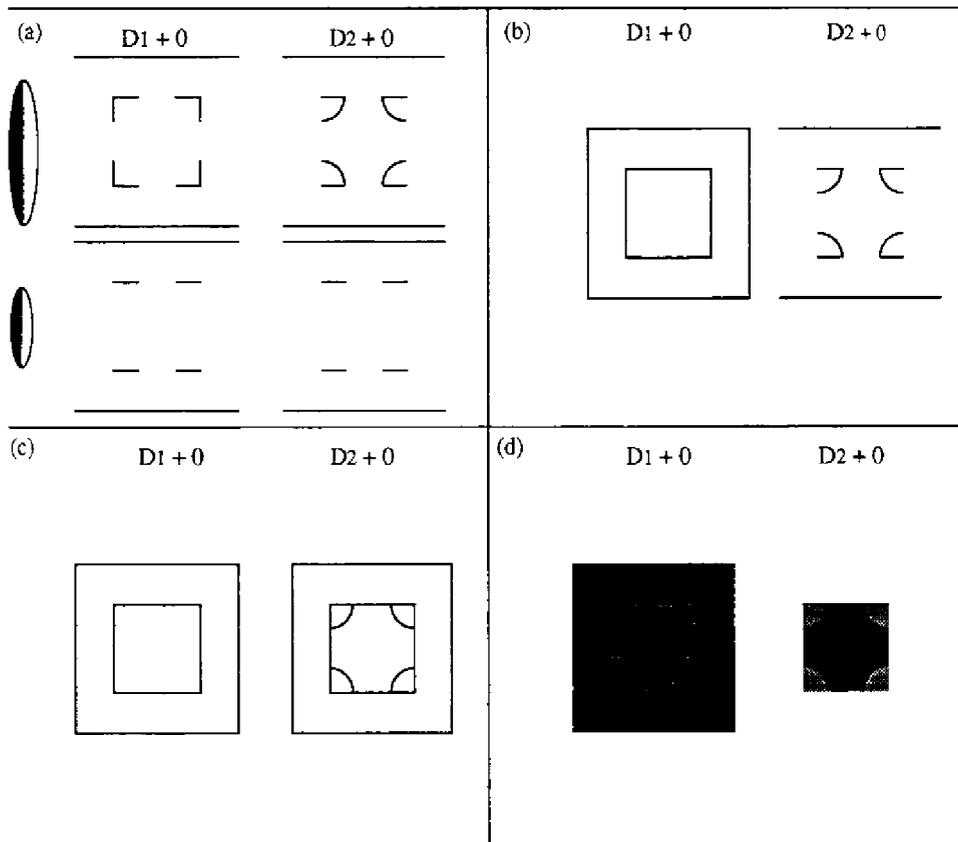


Figure 36. Effect of removing the Kanizsa pacman in Figure 34 on 3-D perception of the Varin wedges. See text for details.

### 36. Differences in BCS and FCS Processing of LGN Inputs at Isoluminance

The above data analyses have probed the properties and interactions of BCS boundary segmentation and FCS filling-in processes. These processes also clarify many other types of data. These processes will be described in greater computational detail in the remaining sections, while more data are explained to illustrate these details.

One type of data that has caused considerable discussion concerns the reduction or vanishing of a percept when chromatic inputs are adjusted to be isoluminant. The conclusion is then often drawn (Livingstone & Hubel, 1987, 1988) that the process supporting the percept receives inputs from luminance detectors but not color detectors. This conclusion is not warranted within FACADE theory because the SOC filter of the BCS groups inputs from the lateral geniculate nucleus (LGN) in a different way than the FCS does. In order to function well as a broad-band boundary detector, BCS complex cells combine inputs via simple cells from all types of LGN color opponent cells, albeit not necessarily with equal weights (Boynton, Eskew, & Olson, 1985; Tansley, Robertson, & Maughan,

1983; Thorell, DeValois, & Albrecht, 1984). In contrast, the FCS maintains the opponent organization of LGN cells and elaborates it into a double-opponent organization (Desimone et al., 1985; Livingstone & Hubel, 1984; Zeki, 1983a, 1983b) in order to carry out its functions of discounting the illuminant and filling-in surface properties such as brightness, color, and depth.

BCS boundary activation may be weakened in response to isoluminant stimuli even though its complex cells receive inputs from all types of LGN cells. This is because of the way model simple cell and complex cell receptive fields are built up from outputs of model LGN cells. For example, consider LGN cells that are sensitive to L cones and M cones in the L-M and L+M combinations (Mollon & Sharpe, 1983). The L-M LGN cells individually become less active at isoluminance. Hence, simple cell receptive fields that are built up from their output signals also become less active at isoluminance. The L+M LGN cells do not necessarily become less active at isoluminance. However, their simple cell targets do, because these simple cells estimate the contrast difference within their receptive fields, which is abolished or greatly reduced at isoluminance. A complex cell that

receives inputs from simple cells with these different types of receptive fields will also become less active at isoluminance. If the complex cell activations of the BCS become too weak near isoluminance to significantly activate FCS filling-in or ORS recognition processes, then the percept will also become weak as isoluminance is approached. Recent psychophysical (Cavanagh & Favreau, 1985) and neurophysiological (Logothetis, Schiller, Charles, & Hurlbert, 1990) experiments using isoluminant stimuli support the idea that attenuation of activation at isoluminance does not imply absence of chromatic inputs to BCS boundaries.

### 37. Sparse Blue Cones and Continuous Blue Surfaces

A related paradox about early visual processing may also be clarified in terms of BCS/FCS interactions. This paradox concerns the fact that blue surface properties may be vividly perceived even though blue cones are very sparsely distributed across the retina. How does such a sparse and discontinuous set of detectors generate a vivid and continuous surface percept? How are sharp borders of such a surface determined? The present remarks consider possible contributions to this process that supplement the local organization of receptive fields per se.

The BCS combines opponent inputs from all model LGN channels at individual complex cells in order to build up the strongest possible boundaries, whereas the FCS preserves the LGN opponent organization and elaborates it into a double-opponent organization in order to discount the illuminant and fill-in properties of surface brightness, color, and depth. A key stage in discounting the illuminant takes output signals derived from a field of similar photodetectors or other signal sources and inputs them into a shunting on-center off-surround network (Cohen & Grossberg, 1984; Grossberg, 1983; Grossberg & Todorović, 1988). Such a network tends to *normalize* the total activity of input patterns as it discounts the illuminant (see Figure 17b). This operation tends to make the network output independent of the total number of inputs per unit area. In other words, discounting the illuminant partially compensates for differences in receptor density. Filling-in then smooths out such a normalized, but still sparse, input pattern until it is obstructed by the best BCS boundaries that all the image data, working together, can generate. The result is a continuous, sharply bounded surface representation whose vividness is assured by shunting normalization.

Grossberg (1987b, Section 31) used these concepts to analyze experiments about the effect of blue cones on border distinctness (Boynton et al., 1985; Tansley & Boynton, 1976, 1978; Tansley et al., 1983). The hypothesis that opponent information from LGN is used in different ways to generate BCS boundary segmentations and FCS surface properties plays a central role in this analysis, as it did in Section 36, to conclude that attenuation of a percept at isoluminance does not imply that BCS boundaries receive no chromatic LGN inputs.

### 38. A Multiple-Scale Binocular Filter: Size-Disparity Correlations, Fusion, and Rivalry

The computational analysis of how multiple-scale 3-D boundary segmentations arise builds upon the theory of binocular vision that was introduced in Grossberg (1987c) and further developed through computer simulations in Grossberg and Marshall (1989). This theory shows how to convert the SOC filter whereby simple cell outputs combine to generate complex cell receptive fields (Figures 12 and 18) into a multiple-scale binocular filter that is capable of supporting binocular fusion and rivalry.

To accomplish this, inputs from the two eyes are organized into hypercolumns (Figure 37a). Small patches of the left (L) and right (R) eye retinas project to contiguous regions of visual cortex. The complete set of oriented L and R simple cells corresponding to that patch forms a hypercolumn (Hubel & Wiesel, 1977). Hypercolumns are joined together sequentially to form ocular dominance columns (Figure 37b). Traversal of an ocular dominance column leads to changes in the preferred orientation of receptive field responsiveness superimposed upon slower changes in retinal position (Blasdel, 1989; Hubel & Wiesel, 1977).

Scenic contrasts are converted into a pattern of activation that is distributed across the spatial map of hypercolumns. As a scenic edge is moved with respect to the two eyes, it generates distinct activation patterns across this spatially organized data structure (Figure 38). The binocular SOC filter from simple cells to complex cells converts such a distributed activation pattern into multiplexed tuning curves whereby individual complex cells code prescribed combinations of edge position, orientation, size, positional disparity, and orientational disparity.

The network that converts distributed activation patterns across the model simple cells (level  $F_1$ ) into multiplexed responses by model complex cells (level  $F_2$ ) is a variant of the competitive learning model. Competitive learning models—also called self-organizing feature maps and adaptive vector quantizers—were developed in Grossberg (1972, 1976a, 1976b) and von der Malsburg (1973), leading in Grossberg (1976a, 1978a, 1982b) and Willshaw and von der Malsburg (1976) to several versions of the model, of which many subsequent contributions are variations (Amari & Takeuchi, 1978; Bienenstock, Cooper, & Munro, 1982; Carpenter & Grossberg, 1987a, 1987b, 1988; Cohen & Grossberg, 1986, 1987; Grossberg & Kuperstein, 1989; Kohonen, 1982, 1984; Linsker, 1986a, 1986b, 1986c; Rumelhart & Zipser, 1985; Singer, 1983). A historical discussion of the development of these models is provided in Grossberg (1987a, 1988).

In a self-organizing feature map, normalized input patterns across level  $F_1$  pass through an adaptive filter to an  $F_2$  level, 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 and featurally more selec-

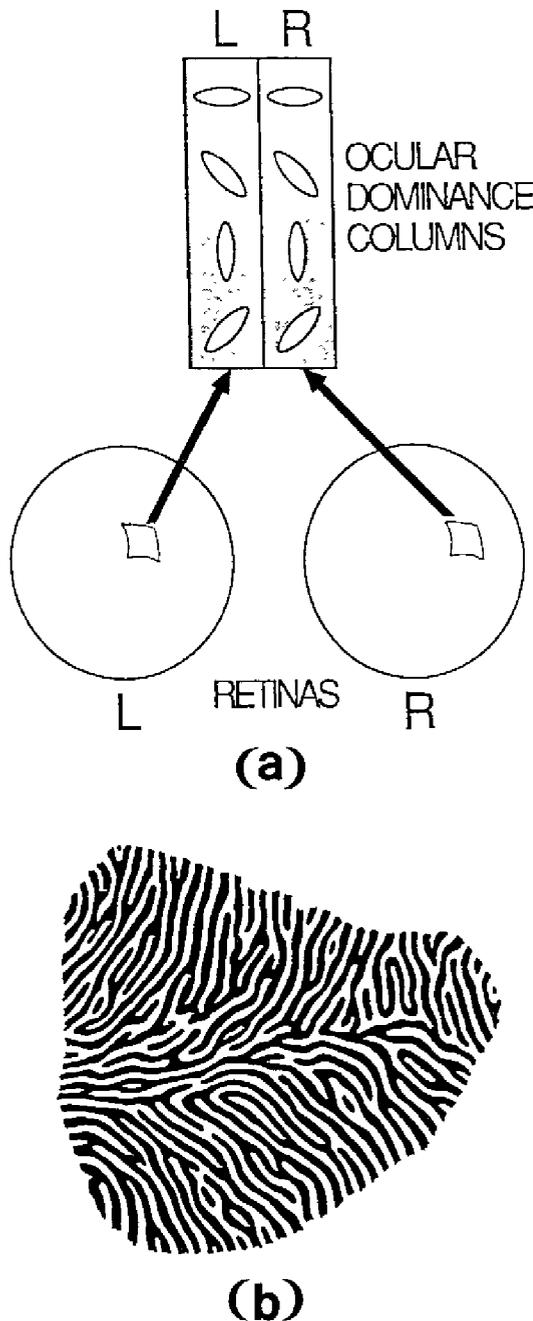


Figure 37. (a) Small patches of left (L) and right (R) eye retinas project to contiguous regions of visual cortex wherein a set of oriented receptive fields, or local contrast detectors, selectively react to oriented retinal contrasts. Such a complete set of oriented L and R detectors corresponding to a pair of retinal patches is called a *hypercolumn* (Hubel & Wiesel, 1977). (b) 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. From "Functional architecture of macaque monkey visual cortex" by D. H. Hubel and T. N. Wiesel, 1977, *Proceedings of the Royal Society of London (B)*, 198, 1-59. Copyright 1977 by The Royal Society. Reprinted by permission.

tive than its input pattern. The  $F_2$  cells that survive the competition with sufficiently positive activities can trigger learning within the adaptive weights of the filter. Learning occurs only in those adaptive weights whose pathways about the winning cells. Learning better tunes the  $F_2$  receptive fields to the statistics of the environment that activates them.

In the binocular SOC filter, level  $F_1$  models simple cells in the cortical hypercolumn map and level  $F_2$  models binocular complex cells (Figure 39). Learning in the adaptive filter from  $F_1$  to  $F_2$  refines the tuning curves of the binocular complex cells, notably their sensitivity to binocular disparity. Reciprocal, top-down signals also exist to stabilize the adaptive tuning process and to help select binocularly consistent LGN activations (Grossberg, 1980). Reciprocal, top-down connections are needed, more generally, to stabilize the learning in all adaptive filter circuits (Carpenter & Grossberg, 1987a, 1988, 1991), but will not be further discussed because they are not needed to explain the data that are analyzed herein.

Multiple copies of the  $F_1 - F_2$  binocular filter exist, corresponding to different simple cell receptive field sizes (Figure 40). Within each copy, the simple cell and complex cell receptive field sizes covary. This property of self-similarity across multiple scales enables the network to exhibit the size-disparity correlation that was used in Sections 27-35 to explain various 3-D percepts.

Figure 41 illustrates how self-similarity enables the larger scale complex cells to binocularly fuse the left eye and right eye monocular responses to an image edge over a broader range of disparities than can the smaller scale complex cells. Thus, there exist larger disparities which the smaller scales cannot fuse but which the larger scales can fuse. Large scale complex cells can fuse both low spatial frequency and high spatial frequency input patterns. Small scale complex cells can fuse only high spatial frequency patterns (Figure 41a). The model hereby provides an explanation of data showing that "there exists an association between large disparities and low spatial frequency, though the complementary association between small disparities and high spatial frequencies, while logically possible, lacks experimental support" (Julesz & Schumer, 1981, p. 609). The model thus does not posit separable pools of disparity-tuned and spatial frequency-tuned cells, consistent with experiments showing "the existence of a very limited number of discrete spatial frequency-tuned mechanisms in human stereopsis (Yang & Blake, 1991, p. 1187). On the other hand, the model clarifies why the maximum fusible disparity is larger for low spatial frequency patterns than for high spatial frequency patterns. The spatial frequency dependence of the disparity range for stereopsis is called the size-disparity correlation. This property has been reported in psychophysical studies (Julesz & Schumer, 1981; Richards & Kaye, 1974; Schor & Tyler, 1981; Schor & Wood, 1983; Schor et al., 1984; Tyler, 1975, 1983) and is observed when certain stereograms are viewed (Kulikowski, 1978). It suggests how the disparity limit may remain approxi-

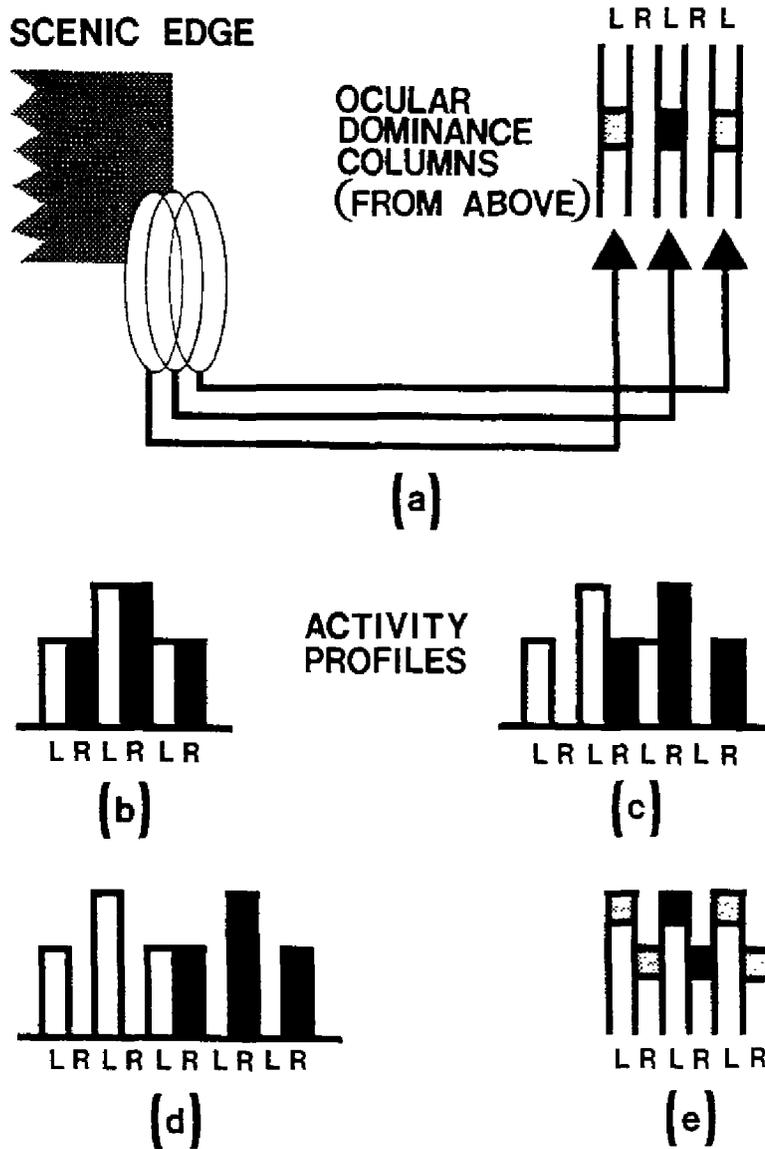


Figure 38. Translation of scenic contour information into spatial patterns of activity: (a) Overlapping like-oriented receptive fields generate a spatial pattern of activity at left eye monocular representations in response to a left eye monocularly viewed scenic edge. The figure portrays a view from above of cortical ocular dominance columns for the left (L) and right (R) eyes, and codes 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 from Grossberg, 1987c.)

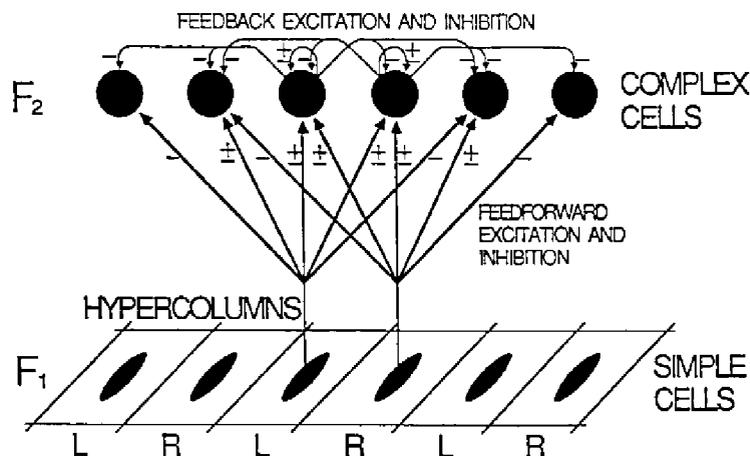


Figure 39. Visual signals broadly distributed across simple cells within the hypercolumn layer,  $F_1$ , input to a complex cell layer,  $F_2$ , via feedforward on-center off-surround shunting interactions. Feedback on-center off-surround shunting interactions transform broad activations at  $F_1$  into sharp, multiplexed activations at  $F_2$ . The Gaussian bandwidths 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. From "Stereo boundary fusion by cortical complex cells: A system of maps, filters, and feedback networks for multiplexing distributed data" by S. Grossberg and J. Marshall, 1989, *Neural Networks*, 2, 29-51. Copyright 1989 by Pergamon Press. Reprinted with kind permission from Pergamon Press Ltd., Headington Hill Hall, Oxford OX3 0BW, U.K.

mately invariant as the viewing distance is changed. The model also explains how a size-disparity correlation can coexist with data like those of the Weisstein effect (Section 3), which seem to contradict it.

BB intrascale competition converts the range of possible fusible disparities within a given scale into the choice of a disparity that best matches the image data. This competition occurs among complex cells that code the same position but different binocular disparities (Figure 41a). Consistent with neurophysiological recordings, it endows model complex cells with a tuning curve that is more sharply tuned to binocular disparity than are their activations without inhibition (Poggio, 1984, 1989; Poggio & Talbot, 1981; Sillito, 1974, 1975a, 1975b, 1977, 1979; Sillito, Salt, & Kemp, 1985).

After BB intrascales select the best matched complex cells, they input to the hypercomplex cells via the first competitive stage (Figure 18). This interaction clarifies how the size-disparity correlation regulates the onset of binocular fusion or rivalry. Within the fusion range, a single population of complex cells is activated which satisfies the size-disparity correlation. These complex cells, in turn, excite a single population of hypercomplex cells that is free from competition at the first competitive stage (Figure 41b). Within the rivalry range, spatially separate populations of complex cells are activated, because they fall outside the range of the size-disparity correlation. These populations mutually inhibit their target hypercomplex cells via the first competitive stage, thereby initiating binocular rivalry (Figure 41c).

The net effect of these interactions is to activate those hypercomplex cells which code the largest disparity, or disparities, that can fuse the monocular inputs from the left and right eyes. All other disparities are suppressed by competitive interactions between complex cells (fusion range) or from complex cells to hypercomplex cells (rivalry range). In particular, if the complex cells that would otherwise win the fusion competition do not exist due to the size-disparity correlation, then *the complex cells that code the next largest disparity will be activated*. Such a disparity shift was used, for example, to explain Bregman-Kanizsa figure-ground separation (Figure 27) and the pop-out of higher spatial frequencies (Figure 30).

The complex cells and hypercomplex cells hereby select the BCS signals that will be used to generate fused or rivalrous BCS boundaries. The hypercomplex cells at the second competitive stage respond to these signals by selecting the most favored orientations, using competition *within* each position and scale but *across* disparities, generating end cuts along the way. Bipole stage cooperation responds to and modifies the signals it receives from the hypercomplex cells via an interaction *across* positions and scales but *within* a given disparity (or relative depth) range (see Figure 41a and Section 42). Multiple boundary segmentations hereby begin to form within different relative depth ranges, and are modified by the exchange of BF intracopy and FB intercopy feedback signals.

Parametric properties of binocular rivalry may occur among these boundary segmentations when habituating transmitter gates are incorporated into the network's

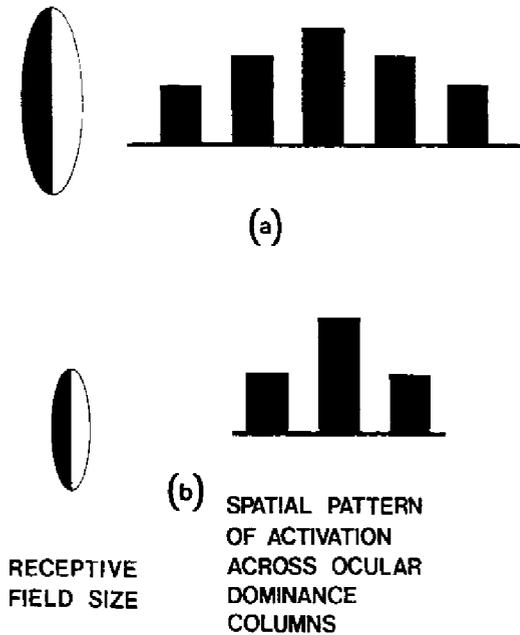
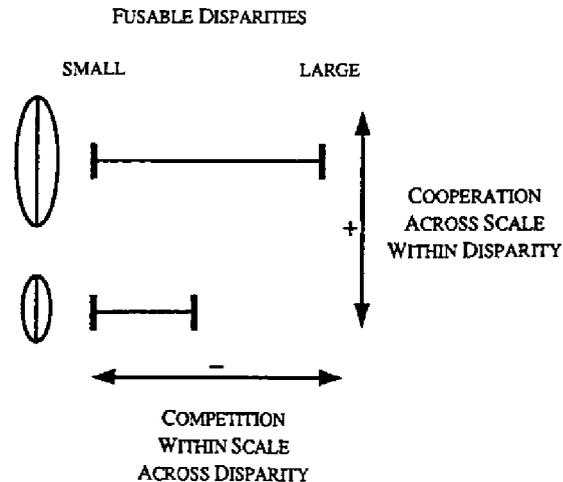


Figure 40. Early stages of spatial frequency sensitivity are represented by covariation of receptive field size with the spatial pattern of activity generated by cells of this size across  $P_1$ : Small receptive fields generate more localized patterns (a) than large receptive fields (b). Bar height represents activity at each cell position. (Reprinted from Grossberg, 1987b.)

cooperative-competitive interactions (Grossberg, 1987c). These are the same habituated interactions that have been used to explain visual persistence data (Section 21) and figure-ground reversal data (Section 29). Arrington (1992) has advanced this analysis with quantitative computer simulations of binocular rivalry data, notably the data of Mueller and Blake (1989) on interactions between stimulus contrast and eye dominance. Multiple types of data hereby support the model's processing stages.

The Marr and Poggio (1979) stereopsis model also embodies a size-disparity correlation. It performs a multiple-scale Gaussian average of the image before computing disparities at zero crossings of the averaged left- and right-eye images. This model has a number of properties that are inconsistent with data about stereopsis (Julesz & Schumer, 1981), and prevent it from explaining data about figure-ground separation. Some of these properties are summarized below to emphasize the new features of the present stereopsis model. The Marr-Poggio model does not create a binocularly fused boundary from a pair of spatially disparate monocular boundaries. It does not allow smaller disparities than the preferred disparity to be simultaneously activated, but suppressed due to complex cell inhibition. It does not distinguish between the range of spatial frequencies that can be detected by simple cells of a given size and the range of disparities that their target complex cells can binocularly fuse. It thus does not allow the complex cells of smaller scales to be activated

by disparities that are larger than those to which they are maximally tuned, a property that was used to discuss the Bregman-Kanizsa effect in Section 27. More generally, the Marr-Poggio model omits the key properties that are



(a)

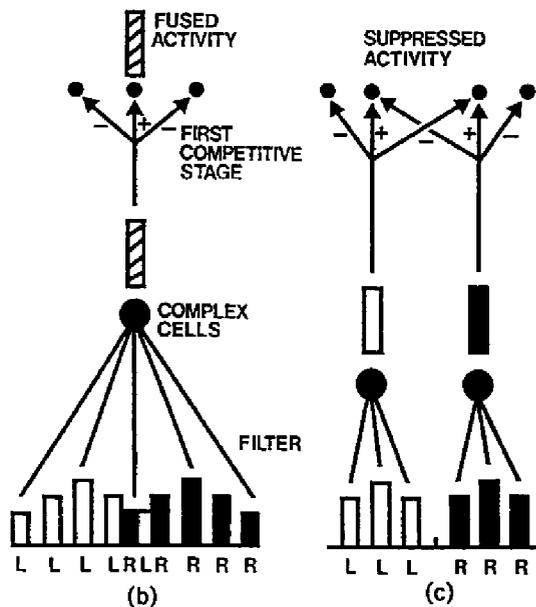


Figure 41. (a) Larger simple cell receptive fields can generate a spatially broader response across the ocular dominance columns than can smaller receptive fields. They can also send convergent signals to complex cells over a broader spatial region. Larger scales can therefore fuse a broader range of disparities than can smaller scales. BB intrascale competition selects the complex cells within each scale and position whose adaptive filter best matches the spatial distribution of simple cell inputs. Outputs from larger scale complex cells to hypercomplex cells (b) can binocularly fuse images at disparities for which the response in smaller scales (c) may be rivalrous, thereby relating the size-disparity correlation to a property of interscale self-similarity.

needed to understand the role of stereopsis in forming binocular boundary segmentations.

### 39. Better Discrimination of Figure Than of Ground

The better discrimination of features when they are perceived as figure than when they are perceived as ground, as summarized in Section 10, is clarified by the competitive disparity interactions within multiple spatial scales that were described in Section 38. These mechanisms show how a part of a planar image that is perceived as a figure can be preferentially processed by the largest scale that can optimally fuse its disparities. The analysis of Figure 30 in Section 10 used this property along with the property that, when another image part is perceived as ground, it often activates smaller scales that are not maximally tuned to this disparity. The stereopsis model summarized in Section 38 shows how this can happen using cells within each scale that can respond to a broad range of disparities. Because a poorer combination of left and right eye inputs survives the binocular competition in the smaller scales than in the larger scales, a weaker and spatially fuzzier reaction occurs in the BCS scales that build up the "ground" of the percept, as was reported in the Brown and Weisstein (1988a) experiments.

This explanation also clarifies the finding of Wong and Weisstein (1983) that sharp targets are detected better against a figure and blurred targets are detected better against a background. Sharp targets are detected better against a figure because they are processed by scales that can sharply fuse their features. Blurred targets are detected better against a ground because they can span the spatially fuzzy reactions of the ground scales. This explanation overcomes the objection, noted in Section 10, that a spatial frequency difference per se is insufficient to explain these figure-ground differences because the same image part can serve as figure or ground in a multistable percept.

This analysis also is consistent with the discussion in Section 29 of how figure-ground switching can occur between high and low spatial frequency patterns. When the high spatial frequency pattern is processed as a figure, it can complete a connected surface using the larger scale and disparity, as in Figure 30. The low spatial frequency pattern uses a smaller scale and disparity to build up its boundaries and fill-in its surface. When switching occurs, the low spatial frequency pattern can create end gaps in the boundaries of the high spatial frequency pattern. Then the low spatial frequency pattern can complete a connected surface using the larger scale and disparity, whereas the high spatial frequency pattern is relegated to the smaller scale and disparity. The greater discriminability of figure than of ground is hereby accounted for by contextually sensitive mechanisms that provide a better match between image data and the percept precisely when these data are perceived as "figure." Two separate systems for figure and ground are not needed and, as noted in Section 10, could not explain these data.

### 40. BB Intrascales: Ocular Dominance, Allelotropia, and Disparity Competition

Figure 42 illustrates a simple and testable self-organizing feature map within which the desired computational properties of multiple-scale competitive disparity computations are attained. This figure depicts an idealized response to an edge by the ocular dominance columns that correspond to a large scale. Because the scale's receptive field size is large, several left eye columns (in white) and right eye columns (in black) are activated. The adaptive filter from the simple cells to the complex cells groups the activity pattern across the simple cells in multiple ways. In particular, left eye and right eye inputs to the binocular cells add within the adaptive filter, before nonlinear competitive interactions select the winning combinations, consistent with the data of Freeman and Ohzawa (1990).

Figure 42 illustrates how properties of allelotropia, size-disparity correlation, and separation of disparity-tuned cell "pools" could be spatially organized across cortical complex cells. The spatial location of a complex cell varies with the total number and pattern of simple cell inputs that project to it. The complex cells that win the BB intra-scale competition in response to a monocular L or R input lie above the L and R markers, respectively, as do near-zero binocular inputs. The complex cells that get the largest input, and thus win, in response to binocular inputs to L and R typically lie symmetrically between L and R. Allelotropia (Section 2A) is hereby initiated via a coordinated shift in map location and disparity tuning to binocularly fused cells at a location midway between the monocular positions L and R. These symmetric binocular complex cells code a binocular disparity that is scaled to

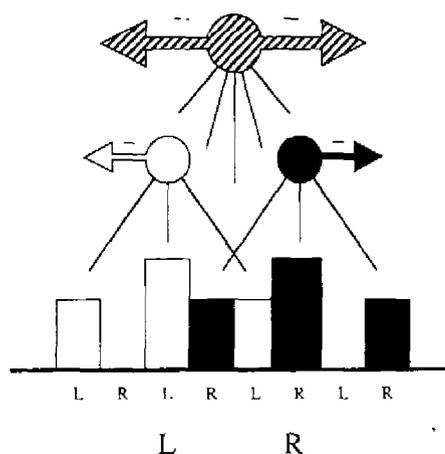


Figure 42. Simple cell ocular dominance columns tend to generate a spatial correlation between disparity-sensitivity, ocular dominance, total input strength, and, possibly, cortical layer at the complex cells. Monocular responses are marked in white or black, binocular responses are marked in white-and-black stripes. This arrangement gives rise to the requisite computational properties of BB intrascales.

the cortical distance  $\|L-R\|$  between L and R. Not all spatial scales can span this cortical distance. A size-disparity correlation derives from the fact that L and R simple cells with larger receptive fields can span a larger cortical distance via their self-similar output pathways, and thereby cause a unimodal activity profile at their target symmetric complex cells (cf. Figure 41b). Complex cells whose receptive fields span a narrower cortical distance, within a given scale, tend to code smaller disparities. Complex cells with larger receptive fields thus tend to code larger disparities and to be in different network layers. Cells tuned to different disparity ranges are hereby spatially segregated, thereby facilitating separation of their output signals to separate BCS copies.

Freeman and Ohzawa (1990) have suggested that complex cell disparity tuning might be set up by pairs of simple cells with the same receptive field center, rather than by the shifted centers used above. In this proposal, the disparity phase-shift between the cell pairs is achieved by using even-symmetric and odd-symmetric receptive fields. It is not yet clear how such a mechanism could explain allelotropia, the spatial separation of disparity-tuned "pools," and the fact that larger scales can fuse a broader range of disparities. Any cortical scheme that supports these properties is consistent with the rest of the theory.

#### 41. Cortical Magnification Factor, Scale Self-Similarity, and Hyperacuity

A BCS scale is not the same thing as a *receptive field size*. This is true because an object contour at a fixed depth from an observer generates a larger binocular disparity as it activates each retina at an increasing distance from the fovea. All of these disparities need to activate the same scale, so that they can fill-in a surface at a fixed depth, as illustrated in Sections 27-35. The explanation of the size-disparity correlation in Section 38 used the self-similarity of receptive field sizes in simple cells and their complex cell targets. This explanation implies that the largest disparity that a complex cell can binocularly fuse covaries with the corresponding simple cell receptive field size. How are receptive field sizes of a BCS scale organized to enable larger disparities to be fused at larger retinal eccentricities? The cortical magnification factor has properties that are consistent with this computational requirement (Daniel & Whitteridge, 1961; Fischer, 1973; Schwartz, 1984; Tootell, Silverman, Switkes, & DeValois, 1982; van Essen, Newsome, & Maunsell, 1984). Larger receptive fields with more receptive field scatter occur at more peripheral locations in the retina and LGN. This arrangement enables higher spatial resolution at the cortex for the registration of foveal retinal inputs than for the registration of peripheral retinal inputs. In particular, less eccentric, notably foveal, retinal locations are represented across more cortical area than are more eccentric retinal locations. A larger retinal disparity at the periphery is hereby compensated for by a smaller spatial separation on the cortex. Were this compensation perfect, larger peripheral disparities and smaller foveal disparities that corresponded to the same egocentric distance would activate

left and right eye cortical representations that were equally displaced on the cortex (Figure 38). In this limiting case, a single cortical receptive field size could be used in each BCS scale to fuse these larger peripheral disparities and smaller foveal disparities. In general, this compensation needs larger peripheral receptive fields at some stage of processing in order to fuse larger peripheral disparities. Figure 43 represents this self-similarity property with respect to the retina, but it needs to be kept in mind that the cortical magnification factor substantially, if not completely, carries out the necessary transformation before the simple cell stage. As a result, the range of disparities capable of binocular fusion (Figure 41) is matched to the degree of receptive field scatter.

The explanation of hyperacuity data in Grossberg (1987b, Section 30) uses spatial properties of this self-similar (complex cell)-to-(hypercomplex cell) filter. The cortical magnification of filter breadth with retinal eccentricity enables the theory to interpret data wherein vernier acuity varies with the cortical magnification factor (Levi, Klein, & Aitsehomo, 1985). These considerations clarify why the term *BCS scale* is used, rather than *receptive field size*. Each scale includes a range of receptive field sizes in order to begin the transformation from the size-disparity correlation to perceived surface depth.

#### 42. Multiscale Interactions of the CC Loop: Coarse-to-Fine Boundary Completion, Curvature Detection, and Depth Repulsion

With these computational interpretations of the cortical magnification factor (Section 41), the size-disparity correlation (Section 38), and the topography of ocular dominance columns (Section 40) in mind, we can now computationally realize the property that the bipole cells of each CC loop generate boundary segmentations for controlling filling-in at a prescribed relative depth from the observer. In particular, disparity is scaled to compensate for the cortical magnification factor, as in Figure 43. It is also assumed that the (hypercomplex cell)-to-(bipole cell) filter is self-similar (Grossberg, 1987c, Section 20). This assumption means that the receptive field sizes of the bipole cells increase with the maximal receptive field

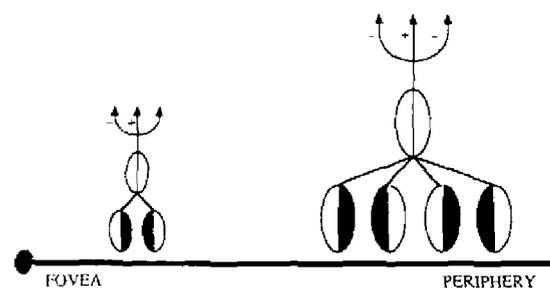


Figure 43. The receptive field sizes and maximum fusible disparities increase with distance from the fovea within each boundary contour system (BCS) copy as part of the transformation from receptive field size and stereo disparity into surface depth.

sizes of the (simple cell) → (complex cell) → (hypercomplex cell) filter that feeds them. As a result, a large bipole cell needs more scenic input to fire than does a small bipole cell. Finally, it is assumed that each CC loop receives inputs from cells that are tuned to its disparity range from *all* the BCS scales to compute a BCS copy (Figure 22b). These converging inputs enable the CC loop to use all the available scenic evidence to build boundaries whose positions and orientations are as accurate as possible. The several BCS scales can selectively cooperate at each position and disparity in the model by using the spatially regular organization of disparities with respect to the ocular dominance columns (Figure 42).

The explanations of 3-D data in Sections 27–35 illustrate how the larger scale activations may stabilize before the smaller scale activations that they modulate. Larger scales, however, may provide a poorer measure of a scenic edge's location, especially at locations of high curvature. The CC loop cooperation across scales within each disparity range enables the large-scale, spatially coarse boundary groupings to be positionally adjusted as small-scale, finer boundary groupings are activated. This process continues until all the scales get a chance to cooperate and compete to complete the boundary web across all positions at each depth. Computer simulations of positional adjustment using multiple-scale cooperation are described in Carpenter et al. (1989).

A number of studies have provided evidence for large-scale spatial channels being activated before smaller scale spatial channels (Arend, 1976; Ferraro & Foster, 1986; Watt, 1987). Such studies have led to the hypothesis that low spatial frequency, fast channels generate a percept's coarse background, whereas high spatial frequency, slower channels elaborate the percept's finer figural representations (Breitmeyer & Ganz, 1976; Ginsburg, 1982; Julesz, 1978; Wong & Weisstein, 1983). The present theory shows how this can happen without contradicting such basic properties as the size-disparity correlation and the fact that an object's image on the retina increases in size as it approaches an observer. In particular, the largest scale that can binocularly fuse a nearby object may be used to generate the representation of that object as a figure in front of a background.

CC loop cooperation across scales helps to explain Griffiths and Chubb's (1993) data showing that subjects can integrate information across different spatial frequency channels to determine contrast boundaries. Also clarified are Wilson and Richards's (1989) data indicating that only (relatively) small, high spatial frequency mechanisms are involved in curvature discrimination, except when a curve is low-pass filtered. To see this, suppose that, at a given disparity, the largest fusible scales are too coarse to respond with positional accuracy to a high curvature edge. Then intrascale competition across positions at the hypercomplex cells of these scales' first competitive stages (Figure 18) attenuates the output from them to the CC loop. This suppression is like the lateral inhibitory interactions that have been reported across positions at the same scale and disparity (Klein & Stromeyer, 1980; Quinn, 1985;

Sagi & Hochstein, 1984). Finer scales that respond more selectively to the high curvature region will not be attenuated. Thus, as in the explanation of greater acuity when features are perceived as figure than when they are perceived as background (Section 39), in the present explanation, the scales that provide a better match to the image data use competitive interactions to reduce or eliminate the influence of the weaker scales. In particular, when the image data exceed the bounds of the size-disparity correlation, as in Figure 41c, or one eye's data mismatch the other eye's data as a result of "ghosts" or other false binocular matches (Julesz, 1971; Kaufman, 1974), then binocular suppression or rivalry can occur.

These mechanisms are also relevant to the results of Westheimer and Levi (1987) on depth attraction and repulsion of disparate foveal stimuli. In their experiments, depth attraction occurs if targets are separated by less than 2'–8' of arc. Depth repulsion occurs at larger separations, between targets with the same or opposite contrasts relative to the background. The attractive effect at small separations may be analyzed in terms of the cooperative pooling across ocular dominance columns by the binocular filter in Figure 42 to select a best disparity at the complex cells and, to a lesser degree, the longer range cooperative linking of nearly colinear activations within a prescribed disparity range. The repulsive effect at larger separations may be analyzed in terms of the composite effects of the first and second competitive stages and FB intercopies, as in the explanations of binocular rivalry (Figure 41c) and Bregman-Kanizsa figure-ground separation (Figure 27). The repulsive competitive interactions work for either direction-of-contrast, because they occur after the complex cell stage (Figure 18).

Westheimer and Levi (1987) noted that these depth attraction and repulsion effects are not the same as the positional attraction and repulsion effects that their lab had earlier reported (Badcock & Westheimer, 1985a, 1985b). These positional displacements were analyzed in Grossberg (1987b, Section 30) in terms of pooling by simple cell receptive fields (short-range attraction) and competition at the first competitive stage (long-range repulsion). The properties of these two types of effects may be used in future computational studies to help select optimal parameters for the filters and competitive interactions.

For bipole cells to achieve the type of multiscale disparity processing suggested above, the connections between hypercomplex cells and bipole cells, no less than those between simple cells and complex cells, need to be adaptively tuned by experience. Marshall (1990) has shown how long-range cooperative connections can self-organize in his model of motion perception. The learning mechanisms that he used are drawn from models of competitive learning and adaptive resonance (Carpenter & Grossberg, 1987a, 1987b; Grossberg, 1982b, 1987c; Kohonen, 1984). Although the Marshall model does not use bipole cells, it provides an example of how selective long-range connections can self-organize by sensing statistically reliable spatiotemporal correlations among spatially distributed inputs.

### 43. Near-Zero Disparity Boundaries, Disparity Pools, and Interpool Cooperation

To prevent brightness and color signals from flowing out of a region, the region boundary must be closed. When we view even a 2-D picture, however, we see it at a distance, so that our eyes receive disparate images. The disparities of nonfoveated vertical and oblique edges are thus nonzero. On the other hand, the disparities of horizontal edges are zero, or approximately so. These facts suggested the hypothesis, used extensively in Sections 27-35, that near-zero disparity boundaries interact with nonzero disparity boundaries in order to close region boundaries even when we view a 2-D picture, and certainly when we view 3-D scenes. Near-zero disparity boundaries must be able to group with the several different nonzero disparity boundaries generated by the scene, or else many region boundaries could not close. It is therefore assumed in the theory that near-zero disparity boundaries are processed in a separate pool from nonzero disparity boundaries. We assumed above that these near-zero boundaries can add multiple copies of themselves to all the pools of nonzero disparity boundaries to enable certain boundaries to close. The theory hereby makes the prediction that near-zero disparity cells and multiple pools of nonzero disparity cells can cooperate via multiple pools of bipole cells in area V2 (and possibly V4) of the prestriate cortex. With these results about multiple-scale BCS boundary segmentation in view, we can now turn to a more detailed discussion of how the BCS controls FCS filling-in.

### 44. Binocular BCS Modulation of Monocular FCS Filling-In Domains

The FCS is decomposed into multiple copies, each with its own Filling-In Domains, or FIDOS. The theory needs to describe (Figure 11): (a) how a binocular BCS segmentation, or BCS copy, is defined; (b) how a binocular BCS segmentation interacts with each monocular FIDO, so that only those monocular feature contour (FC) inputs to the FIDO that are consistent with the binocular segmentation generate outputs, while binocularly inconsistent FC inputs are suppressed; (c) how the output signals from a pair of left eye and right eye monocular FIDOS are binocularly matched at a binocular FIDO; and (d) how these binocularly matched FC signals interact with signals from the binocular BCS segmentation to trigger filling-in of a FACADE representation at each binocular FIDO.

An answer to (a) was developed in Sections 40 and 41. Each BCS copy binds boundary activations that correspond to a prescribed range of nonzero disparities, as well as signals from the near-zero disparity pool which interact with these nonzero activations via each copy's CC loop. Within each BCS copy, the nonzero disparity range varies with distance from the fovea according to the cortical magnification factor in order to define a prescribed relative depth from the observer. Absolute depth measures also require information about the vergence of the eyes (Blank, 1978; Foley, 1980; Greve, Grossberg, Guenther, & Bullock, 1993; Rock, 1984).

We next develop an answer to (b). Each BCS copy sends topographically organized signals to a corresponding monocular FCS copy via BF intracopies. Because each BCS copy is maximally sensitive to a *range* of disparities at each position, this separation is not absolute. Thus, each FCS copy receives inputs from its preferred BCS copy and, possibly, lesser inputs from BCS copies with similar disparity selectivities. We now analyze how a monocular FCS filling-in network, followed by a double-opponent network, reacts to binocular BCS signals in the desired way.

### 45. The FACADE Filter: Combining Filling-In, Double-Opponent, and Boundary-Gating Operations

Each FCS copy is organized into opponent pairs of FIDOS, where each pair corresponds to opponent colors: (red, green), (blue, yellow), (white, black). Figure 44a illustrates that each FC input is delivered to all the FCS copies. Such an FC input selectively activates all, and only, the FIDOS corresponding to its color, across multiple FCS copies. A mixed color can activate more than one FIDO in each FCS copy, much as a mixture of red and yellow codes orange. Each FCS copy also receives topographically organized BF intracopy signals from its BCS copy. Figure 44a shows that each such BCS → FCS input is delivered to all the FIDOS of its FCS copy. Figures 44b and 44c show that the FC input to a FIDO generates an output from the FIDO only if the BCS input occurs in the proper spatial relationship to the FC input. Figure 45 explains how this happens.

Figure 45 illustrates the simplest cellular geometry wherein a shunting on-center off-surround network can receive inputs from a filling-in syncytium. Illuminant-discounted FC signals input to a filling-in syncytium. Filling-in occurs across the syncytium until it hits a BCS boundary or is attenuated by its spatial spread. Each shunting on-center off-surround network responds to spatial nonuniformities in the filled-in activity pattern across its syncytium. Its outputs code the ratio-contrasts of activity levels, and thereby tend to preserve the brightness or color code of the FC inputs that activated the syncytium (Grossberg & Todorović, 1988). In Figure 44b, an FC input to a syncytium fills-in freely around the input position because no BCS input occurs to impede its spread. The shunting network therefore does not register a detectable ratio-contrast, so the FIDO generates no output signal. Thus, the FC input to the lower FIDO in Figure 44a generates no output signal. In Figure 44c, the FC input to the syncytium occurs spatially adjacent to a BCS input to the syncytium. Filling-in occurs only to one side of the BCS input. The shunting network detects the ratio-contrast of the filled-in activity levels, and generates an output signal that increases with this ratio-contrast. Together, Figures 44b and 44c show that a monocular FC input to a FIDO that receives a BF intracopy input can generate an output signal, whereas the same FC input to a FIDO that does not receive a BF intracopy input cannot generate an output signal.

This argument can now be extended to an understanding of how only binocularly consistent FC signals generate

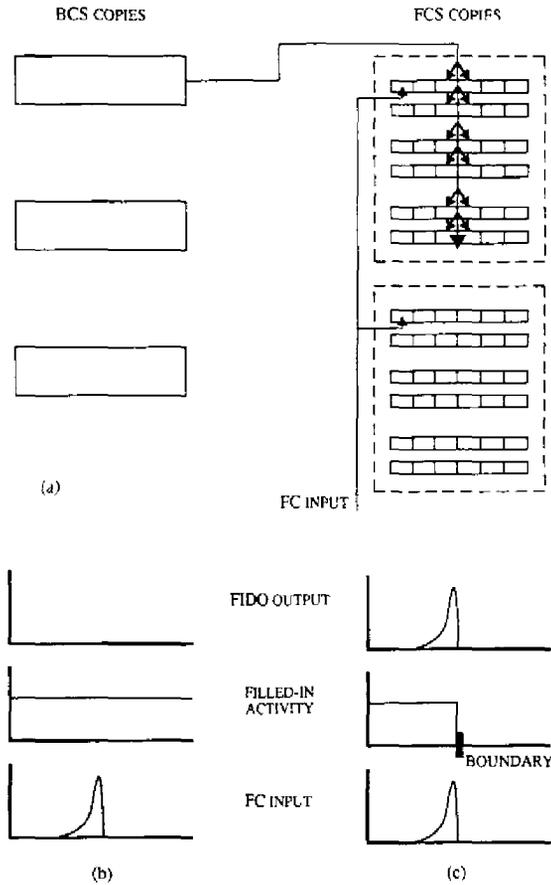


Figure 44. How BF intracopies control monocular Filling-In-Domain (FIDO) output signals: (a) Each illuminant-discounted feature contour (FC) input is broadcast, within its position, to all the filling-in syncytia that code its color across all feature contour system (FCS) scales. Each BF intracopy is broadcast, within its position, across all syncytia corresponding to its targeted FCS copy. (b) If no BF intracopy impedes the filling-in of the FC input, then no output is generated by that FIDO. (c) If a BF intracopy causes a spatial discontinuity to occur in the filling-in of the FC input, then an output is generated by that FIDO.

outputs from a FIDO. Each FIDO consists of a pair of opponent filling-in syncytia that activates a double-opponent output network. Such a double-opponent network consists of three parts: (a) a shunting on-center off-surround network fed by one syncytium, (b) a shunting on-center off-surround network fed by the opponent syncytium, and (c) subtractive opponent interactions between the opponent syncytia at each position (Figure 46). The double-opponent network responds to the filled-in activation patterns of the opponent syncytia by selecting the monocular FC signals that are consistent with the binocular BCS segmentation. Because a FIDO combined with binocular BCS gating signals can select appropriate Form-And-Color-And-Depth combinations, we call it a *FACADE* filter. The following sections explain how this selection process works.

### 46. Binocular Fusion and Rivalry: The Kaufman Stereogram

Binocular rivalry phenomena illustrate how a *FACADE* filter suppresses the monocular featural data that are not consistent with the binocular boundary segmentation. The rivalrous percept generated by a Kaufman stereogram (Figure 47) provides a particularly vivid example. During this percept, the mutually perpendicular diagonal lines in the left eye image and right eye image cannot simultaneously be perceived. When one orientation is perceived in a given region, the perpendicular orientation is suppressed by binocular rivalry. Binocular rivalry between perpendicular orientations is initiated within the CC loop when they compete at the higher order hypercomplex cells of the second competitive stage (Figure 18). Suppose that one of the orientations (say, the right eye orientation) has won the competition in a given region during a given time interval. We need to show how this binocular BCS segmentation interacts with the monocular FIDOS to select the monocular FC signals from the right eye for further processing at the binocular FIDOS, while it suppresses the monocular FC signals from the left eye (Figure 48).

We consider first the role of opponent processing by the FIDOS, and then the role of double-opponent processing. Given opponent processing, a scenic input generates an on-input (e.g., red or white) to its syncytium and an off-input (e.g., green or black) to the opponent syncytium

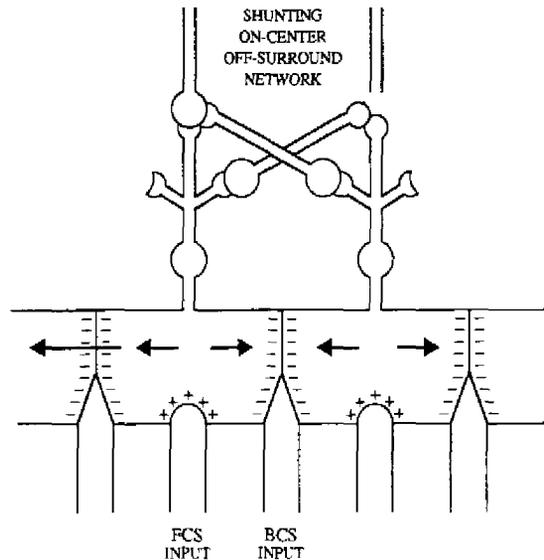


Figure 45. BF intracopy inputs and illuminant-discounted feature contour (FC) inputs converge on a filling-in syncytium. FC inputs activate a syncytium, which carries activation electrotonically to neighboring syncytial cells, except at cell membranes that receive BF intracopy inputs. Syncytial cells activate a shunting on-center off-surround network whose output signals encode the ratio-contrast of spatial discontinuities in the filled-in syncytial activation pattern. Thus, if no BF intracopy prevents lateral spread of the FC input, then no output signals occur, as in Figure 44b.

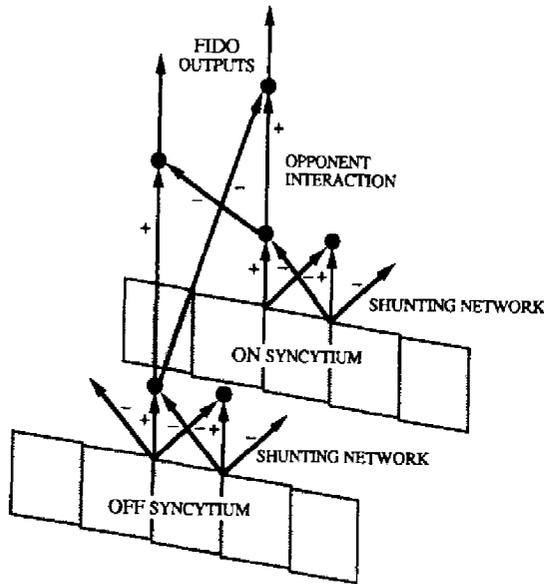


Figure 46. A Filling-In-Domain, or FIDO: The filled-in activity patterns of the on-syncytium and the off-syncytium are filtered by contrast-sensitive on-center off-surround shunting networks. In addition, the output signals from the shunting nets compete at each position to compute the on and off outputs from their respective FIDOs.

(Figure 49a). The off-input has a shape complementary to that of the on-input. Its maximum is thus spatially displaced relative to that of the on-input. Suppose that a BCS boundary signal is received by both syncytia at a location between these maximal opponent responses (Figure 49b). Then the on-syncytium fills-in on one side of the boundary and the off-syncytium fills-in on the other side of the boundary. As a result, the shunting network that is attached to the on-syncytium generates an on-output, while the shunting network that is attached to the off-syncytium generates a spatially displaced off-output.

Consider how to explain the percept corresponding to Figure 48b using the network properties summarized in Figure 49b. The explanation summarized in Figure 49b shows how on-inputs (+ signs) fill-in the on-syncytium on one side of the boundary, and off-inputs (- signs) fill-in the off-syncytium on the other side of the boundary. After opponent processing between the syncytia, the double-opponent output network generates a row of on-outputs and a spatially displaced row of off-outputs for further processing at the binocular FIDOs. This would also occur in the absence of double-opponent competition between the on-cells and off-cells.

Consider how to explain the percept corresponding to Figure 48a. Here the BCS boundary does not occur between the on-inputs and off-inputs. Consequently, filling-in can occur on both sides of the on-inputs and the off-inputs. At positions away from the BCS boundary, the

shunting networks attached to the on-syncytium and off-syncytium can, by themselves, attenuate FIDO outputs, as in Figure 49c.

47. Tissue Contrast and Binocular Rivalry: Lines Are Thick

Why is the double-opponent competition between pairs of opponent syncytium outputs needed, as in Figure 46? It is needed to prevent outputs from occurring along the row of FCS positions that is spatially contiguous to the BCS boundary. To clarify why this is so, I explain the percept of tissue contrast (Helmholtz, 1962). To demonstrate tissue contrast, place a gray circular disk on top of a red background. Cover the whole figure with a white piece of tissue paper. The tissue paper lets the colors be seen and creates an edge without lines, shadows, or other edge-thickening contrasts. Then the gray area looks green. Now draw a circular line with a black pen on the tissue to divide the disk from the background. The gray area looks gray again.

The red-green tissue percept may be explained as follows. As in Figure 49a, the red surround creates red FC inputs just outside the disk boundary and green FC inputs just inside the disk boundary. The red-gray edge creates a narrow boundary signal between these FC inputs. As in Figure 49b, green fills-in the disk interior and red fills-in the disk exterior. Gray also fills-in the disk interior and desaturates the green percept.

When the black line is drawn on the tissue paper, it causes changes in both BCS and FCS processing. This is because the black line has a determinate thickness. It generates a band, or boundary web, of BCS boundaries (see Figure 16 and Grossberg, 1987b; Grossberg & Mingolla, 1987a; Grossberg & Todorović, 1988). The red FC inputs that are registered outside the black line again induce spatially displaced opponent green FC inputs, as in Figure 49a. However, the green inputs now occur within the region that is covered by the boundary web. Filling-in of green therefore occurs in the same region where filling-in of black, from the black line, occurs. Green is therefore not seen. The interior of the gray region looks

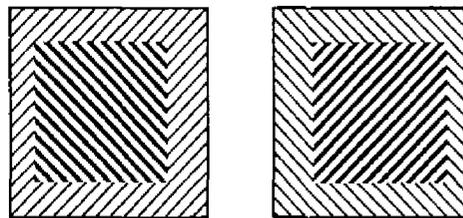


Figure 47. The Kaufman stereogram induces a percept of a square-in-depth as the perpendicular line patterns undergo binocular rivalry. From *Sight and mind: An introduction to visual perception*, by L. Kaufman, 1974, New York: Oxford University Press. Copyright 1974 by Oxford University Press. Reprinted by permission.

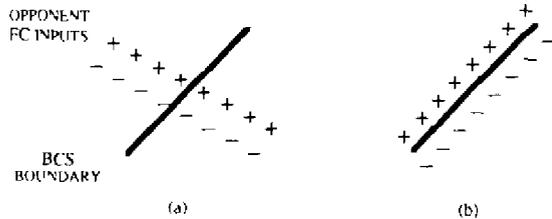


Figure 48. Two-dimensional overlap of a boundary contour system (BCS) boundary contour and a line of correlated on and off inputs to opponent on and off syncytia determine whether output signals will be emitted by their filling-In DOmains (FIDOs): (a) If on (+) and off (-) inputs are not interpolated by a boundary (solid line), the filled-in activities within the on and off syncytia can cancel each other's output signals. (b) If a boundary (solid line) interpolates on (+) and off (-) inputs, output signals from the on FIDO and the off FIDO, respectively, can be generated from opposite sides of the boundary.

gray, not green, because the opponent FC interaction that occurs around the gray/black edge is an achromatic black/gray interaction, rather than the red/gray interaction that occurred before the black line was drawn. In summary, the black line converts the tissue image from two filling-in regions to three filling-in regions, and traps the opponent green induction within a region where it is overwhelmed by black.

The main insight from the tissue contrast percept that we need to explain binocular rivalry is that lines are thick. In particular the lines in the Kaufman stereogram (Figure 47) are thick. Consider what this implies for processing the pattern in Figure 48a, where spatially displaced opponent FC inputs are not separated by a BCS boundary. Within the on-syncytium, filling-in occurs up to the BCS boundary. However, lines are thick! Thus, there exists a sharp discontinuity in the filled-in activity pattern on both sides of the line-induced boundary web. A sharp discontinuity also occurs in the filled-in activity pattern within the off-syncytium. Without additional processing, the shunting network attached to each syncytium would generate large on-outputs and off-outputs to the binocular FIDOs on either side of every line, and the suppressed image would be seen. This additional processing is provided by the competition that occurs at each position between the output signals from each opponent pair of on-syncytium and off-syncytium (Figure 46). The opponent competition works together with the filling-in process to suppress outputs near the line boundaries, as follows.

The FC on-inputs and off-inputs to their respective filling-in syncytia are spatially displaced, as in Figure 49a. On the other hand, they are of approximately equal size. Filling-in obliterates this difference in spatial phase, as in Figure 49c. Thus, the filled-in activity levels at each position along the BCS boundary are essentially identical in the on-syncytium and the off-syncytium. Opponent competition between the syncytium outputs at each position therefore annihilates the outputs that would otherwise have occurred at each line, and suppresses the percept

from the nondominant eye. Experiments that combine tissue contrast and binocular rivalry operations might be useful to further test this explanation.

In summary, interactions of an unoriented filling-in process and an unoriented double-opponent network with an oriented boundary segmentation can selectively filter oriented image properties, including properties that are selective for appropriate combinations of FACADE. This result suggests that color representations are already formed in area V2 before they are binocularly matched to form the final figure-ground FACADE representation in area V4. Whether this predicted V2 representation also contributes to recognition of color remains an open question.

#### 48. The FACADE Representation in Prestriate Area V4

The output signals from the FF binocular intercopies of the FCS interact with the BF intercopies of the BCS to fill-in multiple surface representations at the binocular

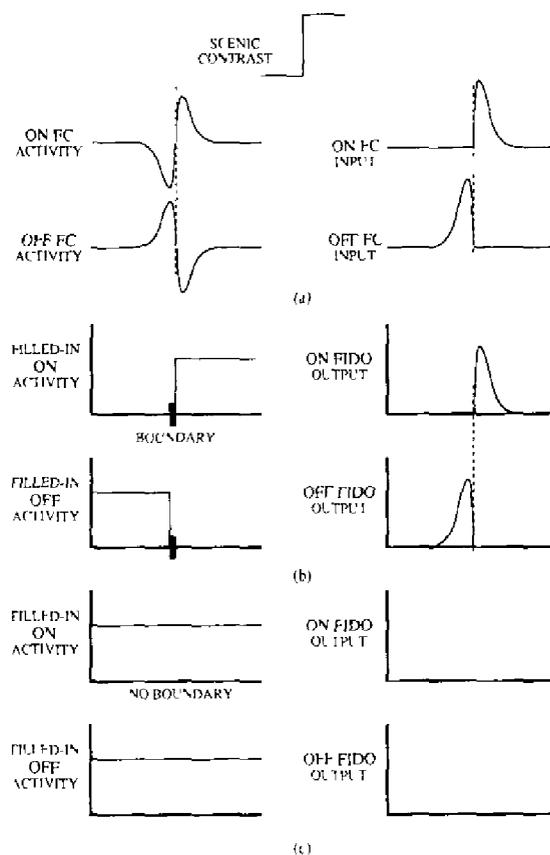


Figure 49. (a) A scenic contrast generates a feature contour (FC) on-input to its syncytium and a spatially displaced FC off-input to the opponent syncytium. (b) If a BF intercopic positionally interpolates the FC on-input and off-input, then output signals are generated by the shunting output networks of the target syncytia. (c) If no BC intercopic is processed, then filling-in occurs in both syncytia and their shunting networks compute zero output signals.

FIDOS (Figures 11 and 28). As in the monocular FIDOS, several FCS copies exist at the binocular FIDOS, corresponding to the several BCS boundary segmentations that process different disparity ranges. The final FACADE representation multiplexes perceptual properties of FACADE in a manner that is qualitatively in accord with the multiplexed receptive field properties that have been reported in prestriate cortical area V4 (Desimone et al., 1985; Zeki, 1983a, 1983b). The theory also predicts that area V4 may play an important role in figure-ground perception. Some data consistent with this prediction are reported in Schiller and Lee (1991).

Such multiplexed surface representations begin to clarify how a visual percept combines several different properties, such as orientation, size, color, and depth, at a single perceptual location. In contrast, other binocular theories have typically failed to explain the appearance and figure-ground properties of surface percepts. Instead, they show how disparity planes or other geometric surface properties may be computed (Marr & Poggio, 1979). Even these computations collapse, however, when a surface percept is induced by very sparse image contrasts, or does not obey continuity constraints, as is typically the case in many stereograms (Julesz, 1971) and in most of the percepts explained above.

#### 49. Attentive Object Recognition, Spatial Orientation, and Visual Search

The organization of the BCS boundary and FCS surface representations suggests how preattentive visual processes in the prestriate cortical areas may interact reciprocally with attentive cortical processes. These attentive processes form part of the What cortical processing stream for object learning and recognition and the Where processing stream for spatial localization and orientation (Goodale & Milner, 1992; Mishkin et al., 1983; Ungerleider & Mishkin, 1982). The concepts of object-based attention (Duncan, 1984) and spatial-based attention (Posner, 1980) have also been used to interpret psychological experiments that probe these processes.

The What processing stream includes the parvocellular area V4 and the inferotemporal (IT) cortex. The Where processing stream passes through the magnocellular area MT and parietal cortex (Figure 1). When the distinction between the What and Where streams was first made, the functional separation of these streams was emphasized. However, as Zeki (1990) has noted, "all the specialized areas communicate, either directly or indirectly, with parietal and temporal areas" (p. 658). FACADE theory provides some indications of how and why this may occur. In particular, the theory offers a framework for analyzing how both the FCS and the static BCS contribute to attentive object recognition, how both the FCS and the motion BCS contribute to attentive spatial orientation, and how both processing streams may interact during visual search (Figure 7b). In neural terms, this analysis suggests why both parvocellular cortical streams interact

with IT cortex, and why both the parvocellular and magnocellular streams interact with parietal cortex.

Some of these processes have been used in previous sections to clarify how a preattentively completed BCS segmentation can directly activate the visual ORS, whether or not the boundary segmentation supports visible brightness or color differences within the FCS (Figure 7). This distinction between recognition and seeing was used, for example, to explain how partially occluded objects can be recognized so well even if their occluded regions are not seen. The ORS can, in turn, read out learned top-down expectations, or priming signals, that focus attention upon those prototypical boundary configurations that have been learned through prior experience (Carpenter & Grossberg, 1987b, 1991, 1993; Carpenter, Grossberg, & Reynolds, 1991; Grossberg, 1980, 1982b). Thus, in response to familiar objects in a scene, the final 3-D segmentation may be *doubly* completed, first by automatic preattentive segmentation processes and then by attentive learned expectation processes. Unfamiliar boundary segmentations are completed only by automatic BCS processes, as a prelude to learning reciprocal visual object-recognition codes in the BCS  $\leftrightarrow$  ORS feedback loop. These interactions are thus part of the object-based attention system.

Growing neurophysiological evidence suggests that such a boundary segmentation process takes place in prestriate area V2 and possibly beyond in V4, and exhibits properties like those of the CC loops in the static BCS (Peterhans & von der Heydt, 1989; von der Heydt et al., 1984; Zeki & Shipp, 1988). Much neurophysiological evidence suggests that the visual ORS includes IT cortex (Desimone, 1991, 1992; Desimone et al., 1985; Desimone & Ungerleider, 1989; Gochin, 1990; Gochin, Miller, Gross, & Gerstein, 1991; Harries & Perrett, 1991; Miller, Li, & Desimone, 1991; Mishkin, 1982; Mishkin & Appenzeller, 1987; Perrett, Mistlin, & Chitty, 1987; Riches, Wilson, & Brown, 1991; Schwartz et al., 1983; Spitzer, Desimone, & Moran, 1988). Accumulating neural evidence suggests, moreover, that the operations of IT cortex are similar in some key respects to those of ART. See Carpenter and Grossberg (1993) for a review.

ART networks model how learned top-down expectations prime, focus attention upon, and bind prototypical feature combinations as part of the process whereby new recognition categories are continually learned throughout life without new learning forcing unselective forgetting of previously learned, but still useful, memories. Within ART, top-down signals, by themselves, cannot fully activate their target cells. Their subliminal priming function is said to obey the "2/3 rule." Such a priming effect may be compared with data showing that cortical top-down or "backward connections seem not to excite cells in lower areas, but instead influence the way they respond to stimuli" (Zeki & Shipp, 1988, p. 316). In addition, although preattentively completed boundary segmentations do not require attentive feedback in order to form, they may be weakened or suppressed if they are not confirmed by a top-down expectation after attention is engaged. Thus

object-based attention (Duncan, 1984) is analyzed within ART as a property of the process whereby object recognition categories are stably learned and remembered.

These model BCS  $\leftrightarrow$  ART interactions make predictions about how areas V2, V4, and IT may interact. In particular, there are three main types of organization that are consistent with the theory: (1) The complete BCS boundaries, as in Figure 27f, are formed in area V2, and project BF intercopies, as in Figure 27h, directly to the FCS FIDOS of area V4. (2) The complete BCS boundaries are formed in area V2, and project to a BCS region of area V4 where the BF intercopies are formed, after which these composite boundary signals project to the FCS FIDOS of area V4. (3) The complete BCS boundaries, whether formed in area V2 or area V4, are also represented in area V4, after which BF intercopies are formed as in (1) or (2). In alternative (1) or (2), direct V2  $\rightarrow$  IT interactions are predicted to subserve the boundary recognition of the completed letters B in Figure 27f. In alternative (3), direct V4  $\rightarrow$  IT interactions could perform this function. Thus, the theory requires that complete BCS boundaries project to the ORS before the stage at which BF intercopies reorganize these boundaries to control surface filling-in within the FCS binocular FIDOS. If V2  $\rightarrow$  IT signals do not perform the boundary recognition function, area V4 should contain both the complete BCS boundaries and the BF intercopy representation.

These reciprocal interactions between the BCS and the ORS are predicted to bind boundary segmentations that are distributed across multiple BCS copies into object recognition categories. These interactions are consistent with the following types of data: Object superiority effects occur using outline stimuli with little surface detail (Davidoff & Donnelly, 1990; Homa, Haver, & Schwartz, 1976). The number of errors in tachistoscopic recognition and the speed of identification are often comparable using appropriately and inappropriately colored objects (Mial, Smith, Doherty, & Smith, 1974; Ostergaard & Davidoff, 1985). There is no difference in the speed with which recognition occurs using black-and-white photographs or line drawings that are carefully derived from them (Biederman & Ju, 1988). Davidoff (1991) has summarized a wealth of additional evidence concerning the ability of normals and clinical patients to recognize objects using only boundary segmentations.

On the other hand, FCS surface representations are also predicted to interact reciprocally with the ORS to generate more complete object recognition codes than the BCS can support on its own (Figure 7). Categorical representations of BCS boundary segmentations within the ORS are hereby joined with categorical representations of the corresponding FCS surface representations. These attentive BCS  $\leftrightarrow$  ORS  $\leftrightarrow$  FCS binding interactions supplement the preattentive BCS  $\leftrightarrow$  FCS interactions that form the separate boundary and surface representations. A theoretical analysis of how this binding and categorization process takes place goes beyond the scope of this article (although see Asfour, Carpenter, Grossberg, & Leshner,

1993). Several types of experimental evidence nonetheless support the hypothesis that distinct boundary and surface representations are formed before being actively bound together. These include the following: A failure to attentively bind colored surfaces to the correct forms occurs during illusory conjunctions (McLean, Broadbent, & Broadbent, 1983; Stefurak & Boynton, 1986; Treisman & Schmidt, 1982). Color can facilitate object naming if the objects to be named are structurally similar or degraded (Christ, 1975; Price & Humphreys, 1989). Colors are coded categorically prior to the processing stage at which they are named (Davidoff, 1991; Rosch, 1975).

The types of surface categories that can form depend upon the organization of the FCS binocular FIDOS. In each FIDO, as illustrated in Figures 3b and 44, unique combinations of depth and color are *structurally* separated from other conjunctions of these features. These FCS color-depth slabs interact with ORS recognition categories via adaptive filters. Different combinations of FIDOS may input to each filter, and thus converge on different ORS cells. In this way, different combinations of filters may be biased to be sensitive to different combinations of brightness, color, depth, and surface form properties. Spatially distributed competitive interactions among the ORS recognition categories choose the categories whose adaptive weights best match the filtered visual data. All the while, nonspecific gain control, possibly mediated by the pulvinar, modulates the balance between bottom-up and top-down (FCS, BCS)  $\leftrightarrow$  ORS interactions (Carpenter & Grossberg, 1993; Desimone, 1992; Robinson & Petersen, 1992).

These categorical recognition codes of the ORS are assumed to occur in a coordinate frame wherein retinotopic variations in position, orientation, and size—but not necessarily viewpoint (Perrett et al., 1987)—are partially overcome by appropriate preprocessing. Such a transformation towards spatially invariant recognition categories greatly reduces the computational load on learning and memory resources, since otherwise a combinatorial explosion of recognition categories coding the same object at all possible retinotopic positions, orientations, and sizes would be needed. Within the model, the Where processing stream helps to organize this invariance transformation. Both the FCS and the motion BCS input to the multiplexed spatial map that carries out the model's Where processing operations (Figure 7b). Thus the FCS is predicted to play a dual role: in categorical recognition of colored 3-D surface forms via the What processing stream, and in organizing spatial orientation and visual search via the Where processing stream. Data supportive of this hypothesis are linked below to the way color-depth feature conjunctions are structurally separated from one another on their own binocular FIDOS.

As in the case of FCS  $\leftrightarrow$  ORS interactions, the reciprocal FCS interactions with the multiplexed spatial map also operate through adaptive filters, albeit adaptive filters that preserve a record of the spatial locations of their inputs. Carpenter, Grossberg, and Leshner (1992, 1993) have de-

veloped a model called a What-and-Where filter to illustrate how such a spatial map may form. The filter transforms a filled-in FCS figure that has been separated from its background into activation of a cell population within the multiplexed spatial map. This visuospatial transformation uses the same types of competitive and cooperative mechanisms that are used for boundary segmentation by the BCS. Activation of individual cells in the map represents a figure's position, as well as its overall orientation and size. When a FIDO inputs such a What-and-Where filter, it generates a spatial map whose individual cells represent a figure's position, orientation, size, depth, and color. When more than one figure occurs in a scene, the spatial map represents, in parallel, all their positions, as well as their orientations, sizes, depths, and colors. The color-depth FIDOS hereby organize the overall spatial map into separate color-depth maps of position, orientation, and size. Competitive interactions across these spatial maps work to select the map locations that are most active at any time.

This spatial organization provides a framework for making sense out of many data about visual search that no single previous theory has been able to handle. Many of these data were discovered subsequent to the seminal finding that fast parallel search seems to occur if a target is distinguished from distractors along a single stimulus dimension, whereas serial search seems to occur if a target is defined by the conjunction of two or more stimulus dimensions (Treisman & Gelade, 1980; Treisman & Souter, 1985). A typical exception to this rule was discovered by Nakayama and Silverman (1986), who showed that targets that differ from distractors by a combination of depth and color, or of depth and motion, can be rapidly separated without serial search. Why can some feature conjunctions be rapidly searched while others cannot? One contributing factor is that a unique conjunction of depth and color is *structurally* separated from other conjunctions of these features on its own FIDO. This observation suggests that feature conjunctions that can be structurally separated from other conjunctions on their own FIDO are among those that can be rapidly searched.

The problem of which feature combinations can be rapidly searched is hereby related to the problem of which feature combinations are multiplexed to lie on structurally separate parts of the total 3-D FACADE representation and its multiplexed spatial map. This hypothesis is consistent with the data of Enns and Rensink (1990) showing that the attributes that control visual search form part of the 3-D representation of a scene, and the data of He and Nakayama (1992), which emphasized the importance of 3-D surface representations in the search process. Moreover, any operation that primes, and thereby amplifies the activities of, a particular color-depth slab in Figure 44 can amplify all representations with that color throughout the FCS - ORS - BCS feedback loop in Figure 7, thereby enabling the amplified targets to be searched first. This mechanism helps to explain the data of Egeth et al.

(1984), who demonstrated that subjects could restrict their visual search to just the red items when the targets were red Os and the distractors were red Ns and black Os.

These observations indicate how the multiplexed spatial map may be organized, but more structure is needed to understand how attention is dynamically distributed among spatial and object representations, and how attention shifts control eye and arm movements. Inputs from the motion BCS to the multiplexed spatial map are also needed to understand how this happens. The main new properties of this interaction are clarified by comparison with properties of the BCS - ORS interaction. The BCS - ORS object attention and recognition process results from a combination of bottom-up activation of a learned category, top-down readout of a learned expectation, and their mutual resonant binding through a *feedback* process that is relatively slow to develop. In contrast, the motion BCS input to the multiplexed spatial map can begin to operate quickly enough to start tracking a rapidly flashing or moving target even before it is recognized. This process uses a *feedforward* filter mechanism that is called the MOC filter (Grossberg & Mingolla, 1993; Grossberg & Rudd, 1989, 1992).

The MOC filter also activates a motion segmentation mechanism that generates representations of form-from-motion. These motion segmentations require slower feedback interactions to develop using a network, called a MOCC loop, that is homologous to the CC loop of the static BCS. The fast feedforward signals from the MOC filter are capable of generating motion direction signals from unambiguous visual data; the slower feedback signals of the MOCC loop supplement these signals with motion direction signals from ambiguous visual data, thereby overcoming the aperture problem (Grossberg & Mingolla, 1993).

The fast inputs to the multiplexed spatial map from the motion BCS have properties that are useful to draw attention to the spatial location of rapidly moving targets. For example, if a targeted predator or prey is moving rapidly across a scene, perhaps darting behind protective cover, then an animal may be able to see the target only intermittently. Grossberg (1991, in press) has suggested how the Where processing stream may use the motion system to interpolate these temporally discrete views with a continuous motion signal that adapts its speed to the varying speed of the target, much as apparent motion signals do (Kolers, 1972). These continuous motion signals can speed up or slow down to cover variable distances in equal time. The results of Remington and Pierce (1984) and Kwak, Dagenbach, and Egeth (1991) are of particular interest in this regard, since they report a speedup of spatial attention to cover variable distances in equal time.

In the MOC filter, a spatially continuous motion signal is generated only under certain spatiotemporal conditions, the speed of the motion signal between two successive target locations is nonuniform in time, and spatially discrete jumps in activation may occur in cases where continuous motion is not observed (Grossberg & Rudd, 1992).

Hence, both spatially discrete and continuous shifts of attention can be reconciled by such a model (Eriksen & Murphy, 1987; LaBerge & Brown, 1989; Remington & Pierce, 1984). In addition, the front end of the MOC filter is sensitive to stimulus on-transients, which can therefore preferentially activate and draw attention to their respective locations. Yantis and Johnson (1990) have reported that abrupt cue onsets capture attention in visual search experiments.

Such a continuous motion signal may be used to predict the location and speed of a moving target and, accordingly, to command such motor responses as eye movements. Experimental evidence consistent with attentive eye movement control by the parietal cortex has been reported by Fischer (1986), Fischer and Breitmeyer (1987), Maylor and Hockey (1985), Mountcastle, Anderson, and Motter (1981), Rizzolatti, Riggio, Dascola, and Umiltá (1987), and Wurtz et al. (1982). Grossberg and Kuperstein (1986, 1989) have modeled how attention is competitively allocated to potential saccadic eye movement targets. In their model, the control of visually reactive saccades, saccades directed to an intended target, and planned sequences of saccades are self-organized through interactions that are interpreted to occur across superior colliculus, parietal cortex, and frontal cortex, among other structures. These various input sources compete for attention to determine the next saccadic target. In the model, multiple target locations can be stored in a spatially organized working memory for the control of sequential eye movements. These top-down planned targets compete with bottom-up transient visual cues for priority. The model applies the working memory model of Grossberg (1978b, Sections 29 and 42), which has a "transient memory span" for ordered readout of stored locations that equals (appropriately) four. Consistent experimental evidence for such a model has been reported by Yantis and Johnson (1990) and Yantis and Jones (1991), who have shown that up to four simultaneous abrupt-onset stimuli can sequentially attract attention in a visual search task, and by Yantis and Jonides (1990), who have reported that a highly valid top-down attentional cue can override attentional capture by an abrupt bottom-up stimulus onset.

Attention can be allocated to several positions of the multiplexed spatial map at the same time. That this can happen in vivo is shown by experiments wherein subjects track multiple visual targets moving about a display screen among identical nontarget elements (Pylyshyn & Storm, 1988; Yantis, 1992). Any initial selection cue, such as flashing the targets before they begin to move, can set up an invisible BCS boundary segmentation that groups the targets together while it also activates feedback between the BCS, the FCS, and the multiplexed spatial map that attentively maintains the enhanced salience (and, perhaps, the enhanced brightness) of this grouping via positive feedback.

In summary, the multiplexed spatial map may be activated either by FCS surface representations or by motion BCS signals. During visual search, these spatial repre-

sentations interact reciprocally with primed ORS categories to recursively search for and identify desired targets (Figure 7b). Grossberg, Mingolla, and Ross (1993a, 1993b; also see Ross, Grossberg, & Mingolla, 1993) have simulated these interactions to derive quantitative fits of visual search data, including those of Cohen and Ivry (1991), Mordkoff, Yantis, and Egeth (1990), Treisman and Sato (1990), and Wolfe, Cave, and Franzel (1989). The model thus suggests an explanation of the types of data that have led to variations of feature integration theory (Treisman, 1982; Treisman & Gelade, 1980; Treisman & Sato, 1990) and the guided search model (Wolfe, 1992; Wolfe et al., 1989). Its mechanisms can be neurally interpreted in terms of prestriate, inferotemporal, parietal, and frontal interactions and can be linked to other types of perceptual phenomena via the models of vision, spatial attention, object recognition, and eye movements that were outlined above.

## 50. Concluding Remarks

The present article suggests a solution to the classical 3-D figure-ground problem of biological vision. In so doing, it provides a unified explanation of a large data base from visual psychophysics and neurobiology that has not been explained by alternative theories. It also clarifies why many of these models work where they do. Along the way, the theory makes testable predictions that may lead to novel experimental designs because the theory integrates many types of data that have not previously been linked. The theory thus presents an opportunity for large-scale cooperation between theory and experiment, guided by a body of perceptual principles and model neural mechanisms that are capable of informing each new fact with theoretical significance.

## REFERENCES

- AMARI, S., & TAKEUCHI, A. (1978). Mathematical theory on formation of category detecting nerve cells. *Biological Cybernetics*, **29**, 127-136.
- ANDREOU, A. G., & BOAHKEN, K. A. (1991). Modeling inner and outer plexiform retinal processing using nonlinear, coupled resistive networks. In *Human vision, visual processing, and digital display, II*. Bellingham, WA: Society of Photo-Optical Instrumentation Engineers, **1453**, 270-281.
- AREND, L. E. (1976). Response of the human eye to spatially sinusoidal gratings at various exposure durations. *Vision Research*, **16**, 1311-1315.
- ARRINGTON, K. F. (1992). *Neural network models for color and brightness perception and binocular rivalry*. Unpublished doctoral dissertation, Boston University.
- ARRINGTON, K. F. (1993). *The temporal dynamics of brightness filling-in*. Manuscript submitted for publication.
- ASFOUR, Y. R., CARPENTER, G. A., GROSSBERG, S., & LESHER, G. W. (1993). *Fusion ARTMAP: A neural network architecture for multi-channel data fusion and classification* (Tech. Rep. No. CAS/CNS-TR-93-006). Boston: Boston University. (*Proceedings of the World Congress on Neural Networks* [Vol. II, pp. 210-215]. Hillsdale, NJ: Erlbaum.)
- BADCOCK, D. R., & WESTHEIMER, G. (1985a). Spatial location and hyperacuity: The centre/surround localization contribution function has two substrates. *Vision Research*, **25**, 1259-1267.

- BADCOCK, D. R., & WESTHEIMER, G. (1985b). Spatial location and hyperacuity. Flank position within the centre and surround zones. *Spatial Vision*, **1**, 3-11.
- BECK, J., GRAHAM, N., & SUTTER, A. (1991). Lightness differences and the perceived segregation of regions and populations. *Perception & Psychophysics*, **49**, 257-269.
- BECK, J., PRAZDNY, K., & IVRY, R. (1984). The perception of transparency with achromatic colors. *Perception & Psychophysics*, **35**, 407-422.
- BECK, J., PRAZDNY, K., & ROSENFELD, A. (1983). A theory of textural segmentation. In J. Beck, B. Hope, & A. Rosenfeld (Eds.), *Human and machine vision*. New York: Academic Press.
- BECK, J., ROSENFELD, A., & IVRY, R. (1990). Line segregation. *Spatial Vision*, **4**, 75-101.
- BECK, J., SUTTER, A., & IVRY, R. (1987). Spatial frequency channels and perceptual grouping in texture segregation. *Computer Vision, Graphics, & Image Processing*, **37**, 299-325.
- BERGEN, J. R. (1991). Theories of visual texture perception. In D. M. Regan (Ed.), *Spatial vision* (pp. 114-134). New York: Macmillan.
- BIEDERMAN, I. (1985). Human image understanding: Recent research and a theory. *Computer Vision, Graphics, & Image Processing*, **32**, 29-73.
- BIEDERMAN, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, **94**, 115-147.
- BIEDERMAN, I., & JU, G. (1988). Surface versus edge-based determinants of visual recognition. *Cognitive Psychology*, **20**, 38-64.
- 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.
- BLANK, A. A. (1978). Metric geometry in human binocular perception: Theory and fact. In E. L. J. Leeuwenberg & H. F. J. M. Buffart (Eds.), *Formal theories of visual perception* (pp. 83-102). New York: Wiley.
- BLASDEL, G. G. (1989). Topography of visual function as shown with voltage-sensitive dyes. In J. J. Lund (Ed.), *Sensory processing in the mammalian brain* (pp. 242-268). New York: Oxford University Press.
- BOWEN, R., POLA, J., & MATIN, L. (1974). Visual persistence: Effects of flash luminance, duration, and energy. *Vision Research*, **14**, 295-303.
- BOYNTON, R. M., ESKEW, R. T., JR., & OLSON, C. X. (1985). Research note: Blue cones contribute to border distinctness. *Vision Research*, **25**, 1349-1352.
- BREGMAN, A. S. (1981). Asking the "what for" question in auditory perception. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 99-118). Hillsdale, NJ: Erlbaum.
- BREGMAN, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.
- BREITMEYER, B., & GANZ, L. (1976). Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing. *Psychological Review*, **83**, 1-36.
- BROWN, J. M., & WEISSTEIN, N. (1988a). A phantom context effect. Visual phantoms enhance target visibility. *Perception & Psychophysics*, **43**, 53-56.
- BROWN, J. M., & WEISSTEIN, N. (1988b). A spatial frequency effect on perceived depth. *Perception & Psychophysics*, **44**, 157-166.
- BUCKLEY, D., FRISBY, J. P., & MAYHEW, J. E. W. (1989). Integration of stereo and texture cues in the formation of discontinuities during three-dimensional surface interpolation. *Perception*, **18**, 563-588.
- CARPENTER, G. A., & GROSSBERG, S. (1987a). ART 2: Stable self-organization of pattern recognition codes for analog input patterns. *Applied Optics*, **26**, 4919-4930.
- CARPENTER, G. A., & GROSSBERG, S. (1987b). A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics, & Image Processing*, **37**, 54-115.
- CARPENTER, G. A., & GROSSBERG, S. (1988). The ART of adaptive pattern recognition by a self-organizing neural network. *Computer*, **21**, 77-88.
- CARPENTER, G. A., & GROSSBERG, S. (Eds.) (1991). *Pattern recognition by self-organizing neural networks*. Cambridge, MA: MIT Press.
- CARPENTER, G. A., & GROSSBERG, S. (1993). Normal and amnesic learning, recognition, and memory by a neural model of cortico-hippocampal interactions. *Trends in Neurosciences*, **16**, 131-137.
- CARPENTER, G. A., GROSSBERG, S., & LESHER, G. W. (1992). *A what-and-where neural network for invariant image processing* (Tech. Rep. No. CAS/CNS-TR-92-006). Boston: Boston University [Proceedings of the International Joint Conference on Neural Networks (Vol. 3, pp. 303-308). Piscataway, NJ: IEEE Service Center].
- CARPENTER, G. A., GROSSBERG, S., & LESHER, G. W. (1993). *The what-and-where filter: A neural network for spatial mapping, object recognition, and image understanding*. Manuscript submitted for publication.
- CARPENTER, G. A., GROSSBERG, S., & MEHANIAN, C. (1989). Invariant recognition of cluttered scenes by a self-organizing ART architecture: CORT-X boundary segmentation. *Neural Networks*, **2**, 169-181.
- CARPENTER, G. A., GROSSBERG, S., & REYNOLDS, J. H. (1991). ART-MAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network. *Neural Networks*, **4**, 565-588. Reprinted in G. A. Carpenter & S. Grossberg (Eds.), *Pattern recognition by self-organizing neural networks* (pp. 503-544). Cambridge, MA: MIT Press, 1991.
- CAVANAGH, P., & FAVREAU, O. E. (1985). Color and luminance share a common motion pathway. *Vision Research*, **25**, 1595-1601.
- CHRIST, R. E. (1975). Review and analysis of color coding research for visual displays. *Human Factors*, **17**, 562-570.
- COHEN, A., & IVRY, R. B. (1991). Density effects in conjunctive search: Evidence for a coarse location mechanism of feature integration. *Journal of Experimental Psychology: Human Perception & Performance*, **17**, 891-901.
- COHEN, M. A., & GROSSBERG, S. (1984). Neural dynamics of brightness perception: Features, boundaries, diffusion, and resonance. *Perception & 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.
- COLLETT, T. S. (1985). Extrapolating and interpolating surfaces in depth. *Proceedings of the Royal Society of London (B)*, **244**, 43-56.
- COLTHEART, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics*, **27**, 183-228.
- CRICK, F., & KOCH, C. (1990). Some reflections on visual awareness. *Cold Spring Harbor Symposium on Quantitative Biology*, **55**, 953-962.
- CRUTHIRDS, D. R., GOVE, A., GROSSBERG, S., & MINGOLLA, E. (1991). Preattentive texture segmentation and grouping by the boundary contour system. In *Proceedings of the International Joint Conference on Neural Networks* (Vol. 1, pp. 655-660). Piscataway, NJ: IEEE Service Center.
- CRUTHIRDS, D. R., GOVE, A., GROSSBERG, S., MINGOLLA, E., NOWAK, N., & WILLIAMSON, J. (1992). Processing of synthetic aperture radar images by the boundary contour system and feature contour system. In *Proceedings of the International Joint Conference on Neural Networks* (Vol. 4, pp. 414-419). Piscataway, NJ: IEEE Service Center.
- CRUTHIRDS, D. R., GROSSBERG, S., & MINGOLLA, E. (1993). Emergent groupings and texture segregation. *Investigative Ophthalmology & Visual Science*, **34**, 1237.
- DANIEL, P. M., & WHITTERIDGE, D. (1961). The representation of the visual field on the cerebral cortex in monkeys. *Journal of Physiology*, **159**, 203-221.
- DAVIDOFF, J. B. (1991). *Cognition through color*. Cambridge, MA: MIT Press.
- DAVIDOFF, J. B., & DONNELLY, N. (1990). Object superiority: A comparison of complete and part probes. *Acta Psychologica*, **73**, 225-243.
- DESIMONE, R. (1991). Face-selective cells in the temporal cortex of monkeys. *Journal of Cognitive Neuroscience*, **3**, 1-8.
- DESIMONE, R. (1992). Neural circuits for visual attention in the primate brain. In G. A. Carpenter & S. Grossberg (Eds.), *Neural networks for vision and image processing* (pp. 343-364). Cambridge, MA: MIT Press.
- DESIMONE, R., SCHEIN, S. J., MORAN, J., & UNGERLEIDER, L. G. (1985). Contour, color, and shape analysis beyond the striate cortex. *Vision Research*, **25**, 441-452.
- DESIMONE, R., & UNGERLEIDER, L. G. (1989). Neural mechanisms of

- visual processing in monkeys. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 2, pp. 267-299). Amsterdam: Elsevier Publishing.
- DEVALOIS, R. L., ALBRECHT, D. G., & THORELL, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, **22**, 545-559.
- DEYOE, E. A., & VAN ESSEN, D. C. (1988). Concurrent processing streams in monkey visual cortex. *Trends in Neurosciences*, **11**, 219-226.
- DRESP, B., LORENCEAU, J., & BONNET, C. (1990). Apparent brightness enhancement in the Kanizsa square with and without illusory contour formation. *Perception*, **19**, 483-489.
- DUNCAN, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, **113**, 501-517.
- ECKHORN, R., BAUER, R., JORDAN, W., BROSCHE, M., KRUSE, W., MUNK, M., & REITBOECK, H. J. (1988). Coherent oscillations: A mechanism of feature linking in the visual cortex? *Biological Cybernetics*, **60**, 121-130.
- ECKHORN, R., & SCHANZE, T. (1991). Possible neural mechanisms of feature linking in the visual system: Stimulus-locked and stimulus-induced synchronizations. In A. Babloyantz (Ed.), *Self-organization, emerging properties, and learning* (pp. 63-80). New York: Plenum Press.
- EGETH, H. E., VIRZI, R. A., & GARBART, H. (1984). Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception & Psychophysics*, **25**, 319-327.
- ELLIAS, S., & GROSSBERG, S. (1975). Pattern formation, contrast control, and oscillations in the short term memory of shunting on-center off-surround networks. *Biological Cybernetics*, **20**, 69-98.
- EMMERT, E. (1881). Grössenverhältnisse der Nachbilder. *Klinische Monatsblätter für Augenheilkunde*, **19**, 443-450.
- ENNS, J. T., & RENSINK, R. A. (1990). Influence of scene-based properties on visual search. *Science*, **247**, 721-723.
- ERIKSEN, C. W., & MURPHY, T. D. (1987). Movement of attentional focus across the visual field: A critical look at the evidence. *Perception & Psychophysics*, **42**, 299-305.
- ESKEW, R. T., JR. (1989). The gap effect revisited: Slow changes in chromatic sensitivity as affected by luminance and chromatic borders. *Vision Research*, **29**, 717-729.
- ESKEW, R. T., JR., STROMEYER, C. F., III, PICOTTE, C. J., & KRONAUER, R. E. (1991). Detection uncertainty and the facilitation of chromatic detection by luminance contours. *Journal of the Optical Society of America A*, **8**, 394-403.
- FERRARO, M., & FOSTER, D. H. (1986). Discrete and continuous modes of curved-line discrimination controlled by effective stimulus duration. *Spatial Vision*, **1**, 219-230.
- FIELD, D. J., HAYES, A., & HESS, R. F. (1993). Contour integration by the human visual system: Evidence for a local "association field." *Vision Research*, **33**, 173-193.
- FISCHER, B. (1973). Overlap of receptive field centers and representation of the visual field in the cat's optic tract. *Vision Research*, **13**, 2113-2120.
- FISCHER, B. (1986). The role of attention in the preparation of visually guided eye movements in monkey and man. *Psychological Research*, **48**, 251-257.
- FISCHER, B., & BREITMEYER, B. (1987). Mechanisms of visual attention revealed by saccadic eye movement. *Neuropsychologia*, **25**, 73-83.
- FOLEY, J. M. (1980). Binocular distance perception. *Psychological Review*, **87**, 411-434.
- FOSTER, K. H., GASKA, J. P., NAGLER, M., & POLLEN, D. A. (1985). Spatial and temporal frequency selectivity of neurons in visual cortical areas V1 and V2 of the macaque monkey. *Journal of Physiology*, **365**, 331-363.
- FRANCIS, G., GROSSBERG, S., & MINGOLLA, E. (in press). Cortical dynamics of feature binding and reset: Control of visual persistence. *Vision Research*.
- FREEMAN, R. D., & OHZAWA, I. (1990). On the neurophysiological organization of biological vision. *Vision Research*, **30**, 1661-1676.
- GIBSON, J. J. (1950). *Perception of the visual world*. Boston: Houghton Mifflin.
- GILLAM, B., & BORSTING, E. (1988). The role of monocular regions in stereoscopic displays. *Perception*, **17**, 603-608.
- GINSBURG, A. P. (1982). On the filter approach to understanding the perception of visual form. In D. G. Albrecht (Ed.), *Recognition of pattern and form* (pp. 215-262). Hillsdale, NJ: Erlbaum.
- GLASS, L., & SWITKES, E. (1976). Pattern recognition in humans: Correlations which cannot be perceived. *Perception*, **5**, 67-72.
- GOCHIN, P. (1990). Pattern recognition in primate temporal cortex. But is it ART? In *Proceedings of the International Joint Conference on Neural Networks* (Vol. 1, pp. 77-80). Hillsdale, NJ: Erlbaum.
- GOCHIN, P. M., MÜLLER, E. K., GROSS, C. G., & GERSTEIN, G. L. (1991). Functional interactions among neurons in inferior temporal cortex of the awake macaque. *Experimental Brain Research*, **84**, 505-516.
- GOGEL, W. C. (1956). The tendency to see objects as equidistant and its reverse relations to lateral separation. *Psychological Monograph*, **70**(Whole No. 411).
- GOGEL, W. C. (1965). Equidistance tendency and its consequences. *Psychological Bulletin*, **64**, 153-163.
- GOGEL, W. C. (1970). The adjacency principle and three-dimensional visual illusions. *Psychonomic Monograph Supplement*, **3**(Whole No. 45), 153-169.
- GOODALE, M. A., & MILNER, D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, **15**, 20-25.
- GRAHAM, N., BECK, J., & SUTTER, A. (1992). Nonlinear processes in spatial-frequency channel models of perceived texture segregation. Effects of sign and amount of contrast. *Vision Research*, **32**, 719-743.
- GRAY, C. M., KONIG, P., ENGEL, A. K., & SINGER, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, **338**, 334-337.
- GRAY, C. M., & SINGER, W. (1989). Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proceedings of the National Academy of Sciences*, **86**, 1698-1702.
- GREVE, D., GROSSBERG, S., GUENTHER, F., & BULLOCK, D. (1993). Neural representations for sensory-motor control I. Head-centered 3-D target positions from opponent eye commands. *Acta Psychologica*, **82**, 115-138.
- GRIFFITHS, A. F., & CHUBB, C. (1993). Integration of information across spatial frequency channels in detection of contrast boundaries. *Investigative Ophthalmology & Visual Science*, **34**, 1289.
- GROSSBERG, S. (1972). Neural expectation: Cerebellar and retinal analogs of cells fired by learnable or unlearned pattern classes. *Kybernetik*, **10**, 49-57.
- GROSSBERG, S. (1973). Contour enhancement, short-term memory and constancies in reverberating neural networks. *Studies in Applied Mathematics*, **52**, 217-257.
- GROSSBERG, S. (1976a). Adaptive pattern classification and universal recoding. I: Parallel development and coding of neural feature detectors. *Biological Cybernetics*, **23**, 121-134.
- GROSSBERG, S. (1976b). On the development of feature detectors in the visual cortex with applications to learning and reaction-diffusion systems. *Biological Cybernetics*, **21**, 145-159.
- GROSSBERG, S. (1978a). A theory of human memory: Self-organization and performance of sensory-motor codes, maps, and plans. In R. Rosen & F. Snell (Eds.), *Progress in theoretical biology* (Vol. 5, pp. 233-374). New York: Academic Press.
- GROSSBERG, S. (1978b). A theory of visual coding, memory, and development. In E. Leeuwenberg & H. Buffart (Eds.), *Formal theories of visual perception* (pp. 7-26). New York: Wiley.
- GROSSBERG, S. (1980). How does a brain build a cognitive code? *Psychological Review*, **87**, 1-51.
- GROSSBERG, S. (1982a). Processing of expected and unexpected events during conditioning and attention: A psychophysiological theory. *Psychological Review*, **89**, 529-572.
- GROSSBERG, S. (1982b). *Studies of mind and brain: Neural principles of learning, perception, development, cognition, and motor control*. Boston: Reidel Press.
- GROSSBERG, S. (1983). The quantized geometry of visual space: The coherent computation of depth, form, and lightness. *Behavioral & Brain Sciences*, **6**, 625-657.

- GROSSBERG, S. (1984). Outline of a theory of brightness, color, and form perception. In E. Degreef & J. van Buggenhout (Eds.), *Trends in mathematical psychology* (pp. 5559-5586). Amsterdam: Elsevier/North-Holland.
- GROSSBERG, S. (1987a). Competitive learning: From interactive activation to adaptive resonance. *Cognitive Science*, **11**, 23-63.
- GROSSBERG, S. (1987b). Cortical dynamics of three-dimensional form, color, and brightness perception: I. Monocular theory. *Perception & Psychophysics*, **41**, 87-116.
- GROSSBERG, S. (1987c). Cortical dynamics of three-dimensional form, color, and brightness perception: II. Binocular theory. *Perception & Psychophysics*, **41**, 117-158.
- GROSSBERG, S. (Ed.) (1988). *Neural networks and natural intelligence*. Cambridge, MA: MIT Press.
- GROSSBERG, S. (1990). A model cortical architecture for the preattentive perception of 3-D form. In E. L. Schwartz (Ed.), *Computational neuroscience* (pp. 117-138). Cambridge, MA: MIT Press.
- GROSSBERG, S. (1991). Why do parallel cortical systems exist for the perception of static form and moving form? *Perception & Psychophysics*, **49**, 117-141.
- GROSSBERG, S. (1993). Boundary, brightness, and depth interactions during preattentive representation and attentive recognition of figure and ground (Tech. Rep. No. CAS/CNS-TR-93-003). Boston: Boston University. (*Italian Journal of Psychology*, in press.)
- GROSSBERG, S. (in press). Neural dynamics of motion perception, recognition learning, and spatial attention. In R. F. Port & T. van Gelder (Eds.), *Mind as motion: Dynamics, behavior, and cognition*. Cambridge, MA: MIT Press.
- GROSSBERG, S., & KUPERSTEIN, M. (1986). *Neural dynamics of adaptive sensory-motor control: Ballistic eye movements*. Amsterdam: Elsevier/North-Holland.
- GROSSBERG, S., & KUPERSTEIN, M. (1989). *Neural dynamics of sensory-motor control, Expanded edition*. Elmsford, NY: Pergamon Press.
- GROSSBERG, S., & MARSHALL, J. (1989). Stereo boundary fusion by cortical complex cells: A system of maps, filters, and feedback networks for multiplexing distributed data. *Neural Networks*, **2**, 29-51.
- GROSSBERG, S., & MINGOLLA, E. (1985a). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review*, **92**, 173-211.
- GROSSBERG, S., & MINGOLLA, E. (1985b). Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations. *Perception & Psychophysics*, **38**, 141-171.
- GROSSBERG, S., & MINGOLLA, E. (1987a). Neural dynamics of surface perception: Boundary webs, illuminants, and shape-from-shading. *Computer Vision, Graphics, & Image Processing*, **37**, 116-165.
- GROSSBERG, S., & MINGOLLA, E. (1987b). A neural network architecture for preattentive vision: Multiple scale segmentation and regularization. In M. Caudill & C. Butler (Eds.), *Proceedings of the First International Conference on Neural Networks* (Vol. 4, pp. 177-184). Piscataway, NJ: IEEE Service Center.
- GROSSBERG, S., & MINGOLLA, E. (1992). Neural dynamics of visual motion perception: Local detection and global grouping. In G. A. Carpenter & S. Grossberg (Eds.), *Neural networks for vision and image processing* (pp. 293-342). Cambridge, MA: MIT Press.
- GROSSBERG, S., & MINGOLLA, E. (1993). Neural dynamics of motion perception: Direction fields, apertures, and resonant grouping. *Perception & Psychophysics*, **53**, 243-278.
- GROSSBERG, S., MINGOLLA, E., & ROSS, W. D. (1993a). *A neural theory of attentive visual search: Interactions of visual, spatial, and object representations* (Tech. Rep. No. CAS/CNS-TR-93-038). Boston: Boston University.
- GROSSBERG, S., MINGOLLA, E., & ROSS, W. D. (1993b). A neural theory of visual search: Recursive attention to segmentations and surfaces (Tech. Rep. No. CAS/CNS-TR-93-023). Boston: Boston University. (In *Proceedings of the World Congress on Neural Networks*. Hillsdale, NJ: Erlbaum, in press.)
- GROSSBERG, S., MINGOLLA, E., & TODOROVIĆ, D. (1989). A neural network architecture for preattentive vision. *IEEE Transactions on Biomedical Engineering*, **36**, 65-84.
- GROSSBERG, S., & RUDD, M. (1989). A neural architecture for visual motion perception: Group and element apparent motion. *Neural Networks*, **2**, 421-450.
- GROSSBERG, S., & RUDD, M. (1992). Cortical dynamics of visual motion perception: Short-range and long-range apparent motion. *Psychological Review*, **99**, 78-121.
- GROSSBERG, S., & SOMERS, D. (1991). Synchronized oscillations during cooperative feature linking in a cortical model of visual perception. *Neural Networks*, **4**, 453-466.
- GROSSBERG, S., & SOMERS, D. (1992). Synchronized oscillations for binding spatially distributed feature codes into coherent spatial patterns. In G. A. Carpenter & S. Grossberg (Eds.), *Neural networks for vision and image processing* (pp. 385-405). Cambridge, MA: MIT Press.
- GROSSBERG, S., & STONE, G. (1986). Neural dynamics of word recognition and recall: Attentional priming, learning, and resonance. *Psychological Review*, **93**, 46-74.
- GROSSBERG, S., & TODOROVIĆ, D. (1988). Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena. *Perception & Psychophysics*, **43**, 241-277.
- GROSSBERG, S., & WYSE, L. (1991). Invariant recognition of cluttered scenes by a self-organizing ART architecture: Figure-ground separation. *Neural Networks*, **4**, 723-742.
- GROSSBERG, S., & WYSE, L. (1992). Figure-ground separation of connected scenic figures: Boundaries, filling-in, and opponent processing. In G. A. Carpenter & S. Grossberg, *Neural networks for vision and image processing* (pp. 161-194). Cambridge, MA: MIT Press.
- HARRIES, M. H., & PERRETT, D. I. (1991). Visual processing of faces in temporal cortex: Physiological evidence for a modular organization and possible anatomical correlates. *Journal of Cognitive Neuroscience*, **3**, 9-24.
- HE, Z. J., & NAKAYAMA, K. (1992). Surface features in visual search. *Nature*, **359**, 231-233.
- HELMHOLTZ, H. L. F. VON (1962). *Treatise on physiological optics* (J. P. C. Southall, Trans.). New York: Dover.
- HOMA, D., HAVER, B., & SCHWARTZ, T. (1976). Perceptibility of schematic face stimuli: Evidence for a perceptual gestalt. *Memory & Cognition*, **4**, 176-185.
- HUBEL, D. H., & WIESEL, T. N. (1965). Receptive fields and functional architecture 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.
- HUMPHREYS, G. W., QUINLAN, P. T., & RIDDICH, N. J. (1989). Grouping processes in visual search: Effects with single- and combined-feature targets. *Journal of Experimental Psychology: General*, **118**, 258-279.
- INTRAUB, H. (1985). Visual dissociation: An illusory conjunction of pictures and forms. *Journal of Experimental Psychology: Human Perception & Performance*, **11**, 431-442.
- JULESZ, B. (1971). *Foundations of cyclopean perception*. Chicago: University of Chicago Press.
- JULESZ, B. (1978). Perceptual limits of texture discrimination and their implications for figure-ground separation. In E. Leeuwenberg (Ed.), *Formal theories of visual perception* (pp. 205-216). New York: Wiley.
- JULESZ, B., & SCHUMER, R. A. (1981). Early visual perception. *Annual Review of Psychology*, **32**, 572-627.
- KANZISA, G. (1979). *Organization in vision: Essays in Gestalt perception*. New York: Praeger Press.
- KAUFMAN, L. (1974). *Sight and mind: An introduction to visual perception*. New York: Oxford University Press.
- KAWABATA, N. (1984). Perception at the blind spot and similarity grouping. *Perception & Psychophysics*, **36**, 151-158.
- KAYE, M. (1978). Stereopsis without binocular correlation. *Vision Research*, **18**, 1013-1022.
- KELLMAN, P. J., & SHIPLEY, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology*, **23**, 141-221.
- KLEIN, S., & STROMEYER, C. F., III (1980). On inhibition between spatial

- frequency channels. Adaptation to complex gratings. *Vision Research*, **20**, 459-466
- KLYMENKO, V., & WEISSSTEIN, N. (1986). Spatial frequency differences can determine figure-ground organization. *Journal of Experimental Psychology: Human Perception & Performance*, **12**, 324-330
- 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* (pp. 248-266). New York: Springer-Verlag.
- KOHONEN, T. (1984). *Self-organization and associative memory*. New York: Springer-Verlag.
- KOLERS, P. A. (1972). *Aspects of motion perception*. Oxford, U.K.: Pergamon Press.
- KULIKOWSKI, J. J. (1978). Limit of single vision in stereopsis depends on contour sharpness. *Nature*, **275**, 126-127.
- KWAK, H.-W., DAGENBACH, D., & EGETH, H. (1991). Further evidence for a time-independent shift of the focus of attention. *Perception & Psychophysics*, **49**, 473-480.
- LABERGE, D., & BROWN, V. (1989). Theory of attentional operations in shape identification. *Psychological Review*, **96**, 101-124
- LAWSON, R. B., & GULICK, W. L. (1967). Stereopsis and anomalous contour. *Vision Research*, **1**, 271-297.
- LEHAR, S., HOWELLS, T., & SMOTROFF, I. (1990). Application of Grossberg and Mingolla neural vision model to satellite weather imagery. In *Proceedings of the International Neural Network Conference* (Vol. 2, pp. 805-808). Boston: Kluwer Press
- LEHAR, S., WORTH, A. J., & KENNEDY, D. N. (1990). Application of the boundary contour/feature contour system to magnetic resonance brain scan imagery. In *Proceedings of the International Joint Conference on Neural Networks* (Vol. 1, pp. 435-440). Piscataway, NJ: IEEE Service Center.
- LENNEBERG, E. H. (1967). *Biological foundations of language*. New York: Wiley
- LESHER, G. W., & MINGOLLA, E. (1993). The role of edges and line-ends in illusory contour formation. *Vision Research*, **33**, 2253-2270.
- LEVI, D. M., KLEIN, S. A., & AITSEBAOMO, A. P. (1985). Vernier acuity, crowding, and cortical magnification. *Vision Research*, **25**, 963-977.
- 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 basic 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.
- LIVINGSTONE, M. S., & HUBEL, D. H. (1984). Anatomy and physiology of a color system in the primate visual cortex. *Journal of Neuroscience*, **4**, 309-356.
- LIVINGSTONE, M. S., & HUBEL, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *Journal of Neuroscience*, **7**, 3416-3648.
- LIVINGSTONE, M. S., & HUBEL, D. H. (1988). Segregation of form, color, movement, and depth: Anatomy, physiology, and perception. *Science*, **240**, 740-749.
- LOGOTHETIS, N. K., SCHILLER, P. H., CHARLES, E. R., & HURLBERT, A. C. (1990). Perceptual deficits and the activity of the color-opponent and broad-band pathways at isoluminance. *Science*, **247**, 214-217
- MARR, D., & POGGIO, T. (1979). A computational theory of human stereo vision. *Proceedings of the Royal Society of London (B)*, **204**, 301-328.
- MARSHALL, J. (1990). Self-organizing neural networks for perception of visual motion. *Neural Networks*, **3**, 45-74.
- MAYLOR, E. A., & HOCKEY, R. (1985). Inhibitory components of externally controlled covert orienting in visual space. *Journal of Experimental Psychology: Human Perception & Performance*, **11**, 777-787.
- MCLEAN, J. P., BROADBENT, D. E., & BROADBENT, M. H. P. (1983). Combining attributes in rapid serial visual presentation tasks. *Quarterly Journal of Experimental Psychology*, **35**, 171-186
- METELLI, F. (1974a). Achromatic color conditions in the perception of transparency. In R. B. MacLeod & H. L. Pick (Eds.), *Perception: Essays in honor of J. J. Gibson*. Ithaca, NY: Cornell University Press.
- METELLI, F. (1974b). The perception of transparency. *Scientific American*, **230**, 90-98
- METELLI, F., D'APPOS, O., & CAVEDON, A. (1985). Balanced and unbalanced, complete and partial transparency. *Perception & Psychophysics*, **38**, 354-366
- MEYER, G. E., & DOUGHERTY, T. (1987). Effects of flicker-induced depth on chromatic subjective contours. *Journal of Experimental Psychology: Human Perception & Performance*, **13**, 353-360
- MEYER, G. E., & MING, C. (1988). The visible persistence of illusory contours. *Canadian Journal of Psychology*, **42**, 479-488.
- MEYER, G. E., & SENECA, M. (1983). The illusion of transparency and chromatic subjective contours. *Perception & Psychophysics*, **34**, 58-64.
- MIAL, R. P., SMITH, P. C., DOHERTY, M. E., & SMITH, D. W. (1974). The effect of memory color on form identification. *Perception & Psychophysics*, **16**, 1-3.
- MIKAELIAN, H. H., LINTON, M. J., & PHILLIPS, M. (1990). Orientation-specific luminance aftereffects. *Perception & Psychophysics*, **47**, 575-582.
- MILLER, E. K., LI, L., & DESIMONE, R. (1991). A neural mechanism for working and recognition memory in inferior temporal cortex. *Science*, **254**, 1377-1379.
- MISHKIN, M. (1982). A memory system in the monkey. *Philosophical Transactions of the Royal Society of London (B)*, **298**, 85-95.
- MISHKIN, M., & APPENZELLER, T. (1987). The anatomy of memory. *Scientific American*, **256**, 80-89
- MISHKIN, M., UNGERLEIDER, L. G., & MACKO, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences*, **6**, 414-417.
- MOLLON, J. D., & SHARPE, L. T. (Eds.) (1983). *Colour vision*. New York: Academic Press.
- MORDKOFF, J. T., YANTIS, S., & EGETH, H. (1990). Detecting conjunctions of color and form in parallel. *Perception & Psychophysics*, **5**, 157-168.
- MOUNTCASTLE, V. B. (1978). Brain mechanisms for directed attention. *Journal of the Royal Society of Medicine*, **71**, 14-28
- MOUNTCASTLE, V. B., ANDERSON, R. A., & MOTTER, B. C. (1981). The influence of attentive fixation upon the excitability of the light-sensitive neurons of the posterior parietal cortex. *Journal of Neuroscience*, **1**, 1218-1235.
- MUELLER, T. J., & BLAKE, R. (1989). A fresh look at the temporal dynamics of binocular rivalry. *Biological Cybernetics*, **61**, 223-232.
- NAKAYAMA, K., & SHIMOJO, S. (1988). Depth, rivalry and subjective contours from unpaired monocular points. *Investigative Ophthalmology & Visual Science*, **29**(Suppl.), 21.
- NAKAYAMA, K., & SHIMOJO, S. (1990). Da Vinci stereopsis: Depth and subjective occluding contours from unpaired image points. *Vision Research*, **30**, 1811-1825.
- NAKAYAMA, K., SHIMOJO, S., & RAMACHANDRAN, V. S. (1990). Transparency: Relation to depth, subjective contours, luminance, and neon color spreading. *Perception*, **19**, 497-513
- NAKAYAMA, K., SHIMOJO, S., & SILVERMAN, G. H. (1989). Stereoscopic depth: Its relation to image segmentation, grouping, and the recognition of occluded objects. *Perception*, **18**, 55-68.
- NAKAYAMA, K., & SILVERMAN, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature*, **320**, 264-265.
- NISLY, S. L., & WASSERMAN, G. S. (1989). Intensity dependence of perceived duration: Data, theories, and neural integration. *Psychological Bulletin*, **106**, 481-496.
- 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.
- OSTERGAARD, A. L., & DAVIDOFF, J. B. (1985). Some effects of color on naming and recognition of objects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **11**, 579-587.
- PARADISO, M. A., & NAKAYAMA, K. (1991). Brightness perception and filling-in. *Vision Research*, **31**, 1221-1236.

- PENTLAND, A. P. (1985) The focal gradient: Optics ecologically salient. *Investigative Ophthalmology & Visual Science*, **26**, 243
- PERRETT, D. I., MISTLIN, A. J., & CHITTY, A. J. (1987) Visual cells responsive to faces. *Trends in Neurosciences*, **10**, 358-364
- PETERHANS, E., & VON DER HEYDT, R. (1989). Mechanisms of contour perception in monkey visual cortex, II. Contours bridging gaps. *Journal of Neuroscience*, **9**, 1749-1763
- PICCOLINO, M., NEYTON, J., & GERSCHENFELD, H. M. (1984). Decrease of gap junction permeability induced by dopamine and cyclic adenosine 3' 5'-monophosphate in horizontal cells of turtle retina. *Journal of Neuroscience*, **4**, 2477-2488.
- POGGIO, G. F. (1984) Processing of stereoscopic information in monkey visual cortex. In G. M. Edelman, W. E. Gall, & W. M. Cowan (Eds.), *Dynamic aspects of neocortical function* (pp. 613-635). New York: Wiley
- POGGIO, G. F. (1989) Neural responses serving stereopsis in the visual cortex of the alert macaque monkey: Position-disparity and image-correlation. In J. J. Lund (Ed.), *Sensory processing in the mammalian brain* (pp. 226-241). New York: Oxford University Press
- POGGIO, G. F., & TALBOT, W. H. (1981). Mechanism of static and dynamic stereopsis in foveal cortex of the rhesus monkey. *Journal of Physiology*, **315**, 469-492
- POLLEN, D. A., FOSTER, K. H., & GASKA, J. P. (1985). Phase-dependent response characteristics of visual cortical neurons. In D. Rose & V. G. Dobson (Eds.), *Models of the visual cortex* (pp. 281-291). New York: Wiley.
- POLLEN, D. A., & RONNER, S. E. (1981). Phase relationships between adjacent simple cells in the visual cortex. *Science*, **212**, 1409-1411.
- POLLEN, D. A., & RONNER, S. E. (1982) Spatial computation performed by simple and complex cells in the visual cortex of the cat. *Vision Research*, **22**, 101-118.
- POSNER, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, **32**, 2-25.
- PRICE, C. J., & HUMPHREYS, G. W. (1989). The effects of surface detail on object categorization and naming. *Quarterly Journal of Experimental Psychology*, **41A**, 797-827.
- PRINZMETAL, W. (1990) Neon colors illuminate reading units. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 584-597.
- PRINZMETAL, W., & BOAZ, K. (1989) Functional theory of illusory conjunctions and neon colors. *Journal of Experimental Psychology General*, **118**, 165-190.
- PYLYSHYN, Z. W., & STORM, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, **3**, 179-197.
- QUINN, P. C. (1985) Suprathreshold contrast perception as a function of spatial frequency. *Perception & Psychophysics*, **38**, 408-414.
- RAMACHANDRAN, V. S. (1992). Perception: A biological perspective. In G. A. Carpenter & S. Grossberg (Eds.), *Neural networks for vision and image processing* (pp. 45-92). Cambridge, MA: MIT Press.
- RAMACHANDRAN, V. S., & NELSON, J. I. (1976). Global grouping overrides point-to-point disparities. *Perception*, **5**, 125-128.
- REDIES, C., & SPILLMANN, L. (1981) The neon color effect in the Ehrenstein illusion. *Perception*, **10**, 667-681.
- REGAN, D., ERKELENS, C. J., & COLLEWIN, H. (1986). Visual field defects for vergence eye movements and for stereomotion perception. *Investigative Ophthalmology & Visual Science*, **27**, 806-819.
- REMINGTON, R., & PIERCE, L. (1984) Moving attention: Evidence for time-invariant shifts of visual selective attention. *Perception & Psychophysics*, **35**, 393-399.
- RICHARDS, W. A., & KAYE, M. G. (1974). Local versus global stereopsis: Two mechanisms. *Vision Research*, **14**, 1345-1347
- RICHARDS, W. A., & REGAN, D. (1973). A stereofield map with implications for disparity processing. *Investigative Ophthalmology*, **12**, 904-909.
- RICHES, I. P., WILSON, F. A. W., & BROWN, M. W. (1991) The effects of visual stimulation and memory on neurons of the hippocampal formation and the neighboring parahippocampal gyrus and inferior temporal cortex of the primate. *Journal of Neuroscience*, **11**, 1763-1779
- RIZZOLATTI, G., RIGGIO, L., DASCOLA, I., & UMILTÀ, C. (1987) Reorienting attention across the horizontal and vertical meridians. *Neuropsychologia*, **25**, 31-40.
- ROBINSON, D. L., & PETERSEN, S. E. (1992). The pulvinar and visual salience. *Trends in Neurosciences*, **15**, 127-132.
- ROCK, I. (1984) *Perception*. New York: W. H. Freeman
- ROSCH, E. (1975). The nature of mental codes for color categories. *Journal of Experimental Psychology: Human Perception & Performance*, **1**, 303-322
- ROSS, W. D., GROSSBERG, S., & MINGOLLA, E. (1993) A neural model of visual search. *Investigative Ophthalmology & Visual Science*, **34**, 1235
- RUBIN, E. (1921) *Visuell wahrgenommene Figuren*. Copenhagen: Gyldendalske
- RUBIN, E. (1958) Figure and ground. In D. C. Beardslee & M. Wertheimer (Eds.), *Readings in perception* (pp. 194-203). New York: Van Nostrand.
- RUMELHART, D. E., & ZIPSER, D. (1985). Feature discovery by competitive learning. *Cognitive Science*, **9**, 75-112
- SAGI, D., & HOCHSTEIN, S. (1984). The contrast dependence of spatial frequency channel interactions. *Vision Research*, **24**, 1357-1365
- SCHILLER, P. H., & LEE, K. (1991) The role of the primate extrastriate area V4 in vision. *Science*, **251**, 1251-1253.
- SCHOR, C. M., & TYLER, C. W. (1981). Spatio-temporal properties of Panum's fusional area. *Vision Research*, **21**, 683-692.
- SCHOR, C. M., & WOOD, I. (1983) Disparity range for local stereopsis as a function of luminance spatial frequency. *Vision Research*, **23**, 1649-1654
- SCHOR, C. M., WOOD, I., & OGAWA, J. (1984). Binocular sensory fusion is limited by spatial resolution. *Vision Research*, **24**, 661-665
- SCHWARTZ, E. L. (1984) Spatial mapping and spatial vision in primate striate and inferotemporal cortex. In L. Spillmann & B. R. Wooten (Eds.), *Sensory experience, adaptation and perception* (pp. 73-104). Hillsdale, NJ: Erlbaum.
- SCHWARTZ, E. L., DESIMONE, R., ALBRIGHT, T., & GROSS, C. G. (1983). Shape recognition and inferior temporal neurons. *Proceedings of the National Academy of Sciences*, **80**, 576-578.
- SHIPLEY, T. F., & KELLMAN, P. J. (1992). Perception of partly occluded objects and illusory figures: Evidence for an identity hypothesis. *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 106-120
- SILLITO, A. M. (1974). Modification of the receptive field properties of neurons in the visual cortex by bicuculline, a GABA antagonist. *Journal of Physiology*, **239**, 36P-37P.
- SILLITO, A. M. (1975a). The contribution of inhibitory mechanisms to the receptive field properties of neurons in the striate cortex of the cat. *Journal of Physiology*, **250**, 305-329
- SILLITO, A. M. (1975b) The effectiveness of bicuculline as an antagonist of GABA and visually evoked inhibition in the cat's striate cortex. *Journal of Physiology*, **250**, 287-304.
- SILLITO, A. M. (1977). Inhibitory processes underlying the directional specificity of simple, complex, and hypercomplex cells in the cat's visual cortex. *Journal of Physiology*, **271**, 699-720.
- SILLITO, A. M. (1979). Inhibitory mechanisms influencing complex cell orientation selectivity and their modification at high resting discharge levels. *Journal of Physiology*, **289**, 33-53.
- SILLITO, A. M., SALT, T. E., & KEMP, J. A. (1985) Modulatory and inhibitory processes in the visual cortex. *Vision Research*, **25**, 375-381.
- 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* (pp. 89-101). New York: Springer-Verlag.
- SPITZER, H., DESIMONE, R., & MORAN, J. (1988). Increased attention enhances both behavioral and neuronal performance. *Science*, **240**, 338-340
- SPITZER, H., & HOCHSTEIN, S. (1985) A complex-cell receptive field model. *Journal of Neurophysiology*, **53**, 1266-1286
- STEFURAK, D. L., & BOYNTON, R. M. (1986). Independence of memory for categorically different colors and shapes. *Perception & Psychophysics*, **39**, 164-174

- SLUTTER, Å., BECK, J., & GRAHAM, N. (1989). Contrast and spatial variables in texture segregation: Testing a simple spatial-frequency channel model. *Perception & Psychophysics*, **46**, 312-332.
- TAKEICHI, H., SHIMOJO, S., & WATANABE, T. (in press). Processing levels of neon flank, illusory contours, and neon color spreading. *Perception*.
- TAKEICHI, H., WATANABE, T., & SHIMOJO, S. (1992). Illusory occluding contours and surface formation by depth propagation. *Perception*, **21**, 177-184.
- TANSLEY, B. W., & BOYNTON, R. N. (1976). A line, not a space, represents visual distinctness of borders formed by different colors. *Science*, **191**, 954-957.
- TANSLEY, B. W., & BOYNTON, R. N. (1978). Chromatic border perception: The role of the red- and green-sensitivity cones. *Vision Research*, **18**, 683-697.
- TANSLEY, B. W., ROBERTSON, A. W., & MAUGHAN, K. E. (1983). Chromatic and achromatic border perception: A two-cone model accounts for suprathreshold border distinctness judgments and cortical pattern-evoked response amplitudes to the same stimuli. In J. D. Mollon & L. T. Sharpe (Eds.), *Colour vision* (pp. 445-454). New York: Academic Press.
- TAUSCH, R. (1953). Die beidäugige Raumwahrnehmung—ein Prozess auf Grund der Korrespondenz und Disparation von Gestalten anstelle der Korrespondenz oder Disparation einzelner Netzhautelemente. *Zeitschrift für experimentelle und angewandte Psychologie*, **1**, 394-421.
- THORPE, 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. A. (1987). Perception of three-dimensional form from patterns of optical texture. *Journal of Experimental Psychology: Human Perception & Performance*, **13**, 242-255.
- TOLHURST, D. J. (1972). Adaptation to square-wave gratings: Inhibition between spatial frequency channels in the human visual system. *Journal of Physiology*, **226**, 231-248.
- TOOTELL, R. B. H., SILVERMAN, M. S., SWITKES, E., & DEVALOIS, R. L. (1982). Deoxyglucose analysis of retinotopic organization in primate striate cortex. *Science*, **218**, 902-904.
- TREISMAN, A. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology: Human Perception & Performance*, **8**, 194-214.
- TREISMAN, A., & GELADE, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, **12**, 97-136.
- TREISMAN, A., & SATO, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 459-478.
- TREISMAN, A., & SCHMIDT, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, **14**, 107-141.
- TREISMAN, A., & SOUTHER, J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General*, **114**, 285-310.
- TS'O, D. Y. (1989). The functional organization and connectivity of color processing. In D. M.-K. Lam & C. D. Gilbert (Eds.), *Neural mechanisms of visual perception: Proceedings of the Retina Research Foundation Symposia* (Vol. 2, pp. 87-115). The Woodlands, TX: Portofolio Publishing.
- TYLER, C. W. (1975). Spatial organization of binocular disparity sensitivity. *Vision Research*, **15**, 583-590.
- TYLER, C. W. (1983). Sensory processing of binocular disparity. In C. M. Schor & K. J. Cuffreda (Eds.), *Vergence eye movements* (pp. 199-295). Boston: Butterworths.
- TYNAN, P., & SEKULER, R. (1975). Moving visual phantom: A new contour completion effect. *Science*, **188**, 951-952.
- UNGERLEIDER, L. G., & MISHKIN, M. (1982). Two cortical visual systems: Separation of appearance and location of objects. In D. L. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549-586). Cambridge, MA: MIT Press.
- VAN ALLEN, E. J., & KOLODZY, P. J. (1987). Application of a boundary contour neural network to illusions and infrared sensor imagery. In M. Caudill & C. Butler (Eds.), *Proceedings of the IEEE First International Conference on Neural Networks* (Vol. 4, pp. 193-197). Piscataway, NJ: IEEE Service Center.
- VAN ESSEN, D. C., NEWSOME, W. T., & MAUNSELL, J. H. R. (1984). The visual representation in striate cortex of macaque monkey: Asymmetries, anisotropies, and individual variability. *Vision Research*, **24**, 429-448.
- VAN TUJL, H. F. J. M., & DE WEERT, C. M. M. (1979). Sensory conditions for the occurrence of the neon spreading illusion. *Perception*, **8**, 211-215.
- VARIN, D. (1971). Fenomini di contrasto e diffusione cromatica nell'organizzazione spaziale del campo percettivo. *Rivista di Psicologia*, **65**, 101-128.
- VON DER HEYDT, R., HEITGER, F., ROSENTHALER, L., PETERHANS, E., & KLÜBLER, O. (1992). Simulation of cortical contour mechanisms: Illusory figures, depth and distortions, superior image segmentation of natural scenes. *Investigative Ophthalmology & Visual Science Supplement*, **33**, 1343.
- VON DER HEYDT, R., PETERHANS, E., & BAUMGARTNER, G. (1984). Illusory contours and cortical neuron responses. *Science*, **224**, 1260-1262.
- VON DER MALSBERG, 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: Thomas.
- WALLACH, H., & BACON, J. (1976). Two forms of retinal disparity. *Perception & Psychophysics*, **19**, 375-382.
- WALLACH, H., & BARTON, W. (1975). Adaptation to optically produced curvature of frontal planes. *Perception & Psychophysics*, **18**, 21-25.
- WALLACH, H., & LINDAUER, J. (1962). On the definition of retinal disparity. *Psychologische Beiträge*, **6**, 521-530.
- WATANABE, T., & CAVANAGH, P. (1992). Depth capture and transparency of regions bounded by illusory and chromatic contours. *Vision Research*, **32**, 527-532.
- WATANABE, T., & SATO, T. (1989). Effects of luminance contrast on color spreading and illusory contour in the neon color spreading effect. *Perception & Psychophysics*, **45**, 427-430.
- WATANABE, T., & TAKEICHI, H. (1990). The relation between color spreading and illusory contours. *Perception & Psychophysics*, **47**, 457-467.
- WATT, R. J. (1987). Scanning from coarse to fine spatial scales in the human visual system after the onset of a stimulus. *Journal of the Optical Society of America*, **4**, 2006-2021.
- WATT, R. J., & CAMPBELL, F. W. (1985). Vernier acuity: Interactions between length effects and gaps when orientation cues are eliminated. *Spatial Vision*, **1**, 31-38.
- WERNER, H. (1937). Dynamics in binocular depth perception. *Psychological Monograph* (Whole No. 218).
- WESTHEIMER, G., & LEVI, D. M. (1987). Depth attraction and repulsion of disparate foveal stimuli. *Vision Research*, **8**, 1361-1368.
- WHEATSTONE, C. (1838). On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society of London*, **128**, 371-394.
- WILDE, K. (1950). Der Punktreiheneffekt und die Rolle der binocularen Querdissparation beim Tiefenshen. *Psychologische Forschung*, **23**, 223-262.
- WILLSHAW, D. J., & VON DER MALSBERG, C. (1976). How patterned neural connections can be set up by self-organization. *Proceedings of the Royal Society of London (B)*, **194**, 431-445.
- WILSON, H. R., BLAKE, R., & HALPERN, D. L. (1991). Coarse spatial scales constrain the range of binocular fusion on fine scales. *Journal of the Optical Society of America*, **8**, 229-236.
- WILSON, H. R., & RICHARDS, W. A. (1989). Mechanisms of contour curvature discrimination. *Journal of the Optical Society of America*, **6**, 106-115.
- WOLFE, J. (1992). Effortless texture segmentation and "parallel" visual search are not the same thing. *Vision Research*, **32**, 757-763.
- WOLFE, J., CAVE, K. R., & FRANZEL, S. L. (1989). Guided search: An alternative to the feature integration model of visual search. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 419-433.

- WONG, E., & WEISSTEIN, N. (1982) A new perceptual context-superiority effect. Line segments are more visible against a figure than against a ground. *Science*, **218**, 587-589
- WONG, E., & WEISSTEIN, N. (1983). Sharp targets are detected better against a figure and blurred targets are detected better against a background. *Journal of Experimental Psychology: Human Perception & Performance*, **9**, 194-202
- WORTH, A. J., LEHAR, S., & KENNEDY, D. N. (1992) A recurrent cooperative/competitive field for segmentation of magnetic resonance brain slices. *IEEE Transactions on Knowledge & Data Engineering*, **4**, 156-161
- WURTZ, R. H., GOLDBERG, M. E., & ROBINSON, D. L. (1982, June). Brain mechanisms of visual attention. *Scientific American*, pp 124-135
- YANG, Y., & BLAKE, R. (1991) Spatial frequency tuning of human stereopsis. *Vision Research*, **31**, 1177-1189
- YANTIS, S. (1992) Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology*, **24**, 295-340.
- YANTIS, S., & JOHNSON, D. N. (1990) Mechanisms of attentional priority. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 812-825
- YANTIS, S., & JONES, E. (1991). Mechanisms of attentional selection: Temporally modulated priority tags. *Perception & Psychophysics*, **50**, 166-178.
- YANTIS, S., & JONIDES, J. (1990) Abrupt visual onsets and selective attention: Voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 121-134
- ZEKI, S. (1983a). Colour coding in the cerebral cortex. The reaction of cells in monkey visual cortex to wavelengths and colours. *Neuroscience*, **9**, 741-766.
- ZEKI, S. (1983b). Colour coding in the cerebral cortex: The responses of wavelength-selective and colour coded cells in monkey visual cortex to changes in wavelength composition. *Neuroscience*, **9**, 767-791
- ZEKI, S. (1990). Parallelism and functional specialization in human visual cortex. *Cold Spring Harbor Symposia on Quantitative Biology*, **LV**, 651-661
- ZEKI, S., & SHIPP, S. (1988) The functional logic of cortical connections. *Nature*, **335**, 311-317

## APPENDIX

The following is a list of abbreviations from this article.

ART	Adaptive Resonance Theory
BC	Boundary Contour
BCS	Boundary Contour System
BB	Boundary-Boundary
BF	Boundary-Feature
CC	Cooperative-Competitive
FACADE	Form-And-Color-And-DEpth
FB	Feature-Boundary
FC	Feature Contour
FCS	Feature Contour System
FF	Feature-Feature
FIDO	Filling-In-Domain
IT	InferoTemporal
LGN	Lateral Geniculate Nucleus
MOC	Motion Oriented Contrast-sensitive
MOCC	Motion Oriented Cooperative-Competitive
MP	Monocular Preprocessing
ORS	Object Recognition System
SOC	Static Oriented Contrast-sensitive

(Manuscript received October 19, 1992;  
revision accepted for publication June 15, 1993.)