Neural Networks, Vol. 6, pp. 463–483, 1993
Printed in the USA. All rights reserved.

0893-6080/93 $6.00 + .00

ORIGINAL CONTRIBUTION

A Solution of the Figure-Ground Problem for Biological Vision

STEPHEN GROSSBERG

Boston University

(*Received* 10 *January* 1992; *revised and accepted* 14 *October* 1992)

Abstract—*A neural network model of 3-D visual perception and figure-ground separation by visual cortex is introduced. The theory provides a unified explanation of 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 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 detector 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 3-D neon color spreading, visual phantoms, and tissue contrast percepts are generated; how conjunctions of color-and-depth may rapidly pop-out during visual search. These explanations are derived from an ecological analysis of how monocularly viewed parts of an image fill-in the appropriate depth from contiguous binocularly viewed parts, as during DaVinci stereopsis. The model predicts the functional role and ordering of multiple interactions within and between the two parvocellular processing streams that join lateral geniculate nuclei (LGN) to prestriate area V4. Interactions from cells representing larger scales and disparities to cells representing smaller scales and disparities are of particular importance. An application to painterly color technique is noted.*

Keywords—Vision, Neural networks, Visual cortex, Figure-ground separation, Segmentation, Stereopsis, Surface perception, Filling-in, Spatial frequency analysis, Color perception.

1. INTRODUCTION

This article introduces a neural network theory of biological vision that suggests solutions to some long-standing problems concerning how we perceive a 3-D world, notably the classical figure-ground problem of biological vision. Illustrative explanations concern how a 2-D image may generate a 3-D percept; how figures pop-out from cluttered backgrounds; how binocular fusion of objects at different depths can deform perceptual space by different amounts without destroying its seamless properties; how local properties such as multiple spatial scales and stereo disparities are transformed into global properties such as surface depth; 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 both color and depth can fill-in surfaces defined by sparse image contrasts; and how conjunctions of color-and-depth or other 3-D object properties may pop-out as single attributes during visual search. 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.

The theory suggests how key processing stages in the two parvocellular processing streams from the lateral geniculate nuclei (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 number of processing stages that are needed in the theory well matches the number of stages in the LGN Parvo → Blob → Thin Stripe → V4 processing stream

Supported in part by the Air Force Office of Scientific Research (AFOSR 90-0175), DARPA (AFOSR 90-0083 and ONR N00014-92-J-4015), and the Office of Naval Research (ONR N00014-91-J-4100).

Acknowledgments: Thanks to Cynthia E. Bradford and Diana J. Meyers for their valuable assistance in the preparation of the manuscript, and to Harald Ruda for his assistance with the figures.

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

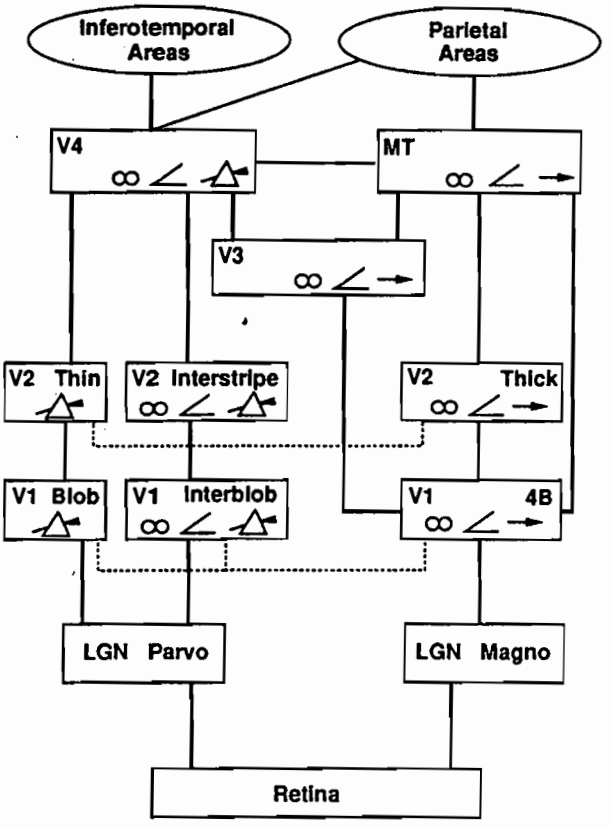


FIGURE 1. Schematic diagram of anatomical connections and neuronal selectivities of early visual areas in the macaque monkey. LGN = lateral geniculate nucleus (parvocellular and magnocellular divisions). Divisions of V1 and V2: blob = cytochrome oxidase blob regions; interblob = cytochrome oxidase-poor 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; MT = middle temporal area. Areas V2, V3, V4, 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%); pointing arrow = direction of motion selectivity (incidence at least 20%). (Adapted with permission from DeYoe and van Essen (1988).)

and the LGN Parvo → Interblob → Interstripe → V4 processing stream.

This paper describes some of the paradoxical data and conceptual problems about 3-D vision that the theory treats. It also outlines how the theory explains these data and resolves the problems. A more detailed and extensive analysis of these and other data about cortical mechanisms of 3-D vision and figure-ground separation is provided in Grossberg (1992).

2. DAVINCI 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 DaVinci, and is often called DaVinci stereopsis (Gillam & Borsting, 1988; Kaye, 1978; Lawson & Gulick, 1967; 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.

2.1. 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 that object region which 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

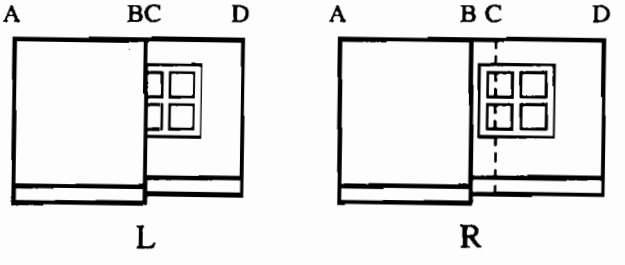


FIGURE 2. When a scene is viewed by both eyes, most of it may be binocularly detected, such as regions AB and CD, but part of it may be detected by only one eye, such as region BC. An appropriate depth of the monocularly viewed region is often filled-in using information from retinally contiguous, binocularly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

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, since images of closer objects are more disparate than images of further 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. In particular, during DaVinci stereopsis, the vertical boundaries of regions AB and CD in the left eye and right eye images of Figure 2 need to be deformed by different amounts in order to be binocularly fused. Given all this deformation of monocular boundaries to form fused binocular boundaries with different amounts of deformation required to fuse objects at different distances from the observer, we need to analyse why no "holes" in binocular perceptual space are created.

2.2. 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 further from the fixation point generate larger binocular disparities. Why does a plane not recede towards optical infinity at the fixation point and curve towards 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? 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). In order to explain how this occurs, the theory utilizes the following types of processes.

2.3. Binocular and Monocular Boundary Representation

The filling-in process is contained by internal representations of scenic boundaries. Some boundaries are binocularly viewed, others monocularly viewed. 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 horizontal 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 DaVinci 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.

2.4. 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 those scales and disparities 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, if the following problem can be solved.

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

2.6. 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 generated, including their brightnesses, colors, and surface depths? 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. 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.

An application to image processing technology of the idea that boundaries that enclose connected regions can use filling-in of their enclosed region for purposes of figure-ground separation was developed in Grossberg and Wyse (1991, 1992).

2.7. Near Boundaries Obstruct Filling-In of Occluded Regions

How does the monocularly viewed surface BC *only* get filled-in at the depth of CD? 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 obstruct filling-in of region AB within the filling-in domain whose depth corresponds to region CD. 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, then why do not *all* such surfaces look transparent?

The theory suggests that this does not happen because the boundaries corresponding to closer objects are added to the boundaries corresponding to further objects in the filling-in domains. 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 may be recognized via their completed boundaries, even if visible surface properties are not filled-in behind the occluding object.

These properties of DaVinci stereopsis illustrate 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 inter-

actions is also indicated by 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. An illustration of this effect is shown in Figure 3, where the low spatial frequency region of the image appear closer than its high spatial frequency region. This property contributes to percepts 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, then the regions with the higher frequency appear closer in depth than those containing the lower frequency, as illustrated in Figure 4.

A comparison of the opposite dependence between spatial frequency and depth in Figures 3 and 4 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*.

These data challenge theories to explain how the expected relationship between spatial frequency and depth, as shown in Figure 3, may be reversed by boundary segmentation processes, as shown in Figure 4, to influence which parts of an image or scene will appear as figure and which as ground. Relative depth may also be influenced by other factors 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 models such 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 (1987b), that distinguish early processing of stereo disparity from later processing of surface depth.

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 corresponding 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 (Bregman, 1981; Kanizsa, 1979). A comparison of Figures 5(b) and 5(c) shows that the existence of the occluding black sinewy shape in front of the occluded B's is needed to readily recognize them as B's.

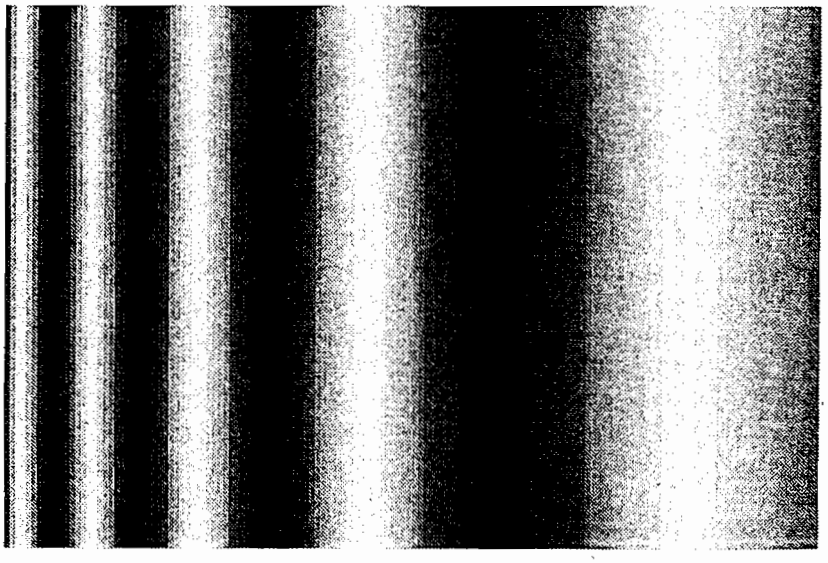
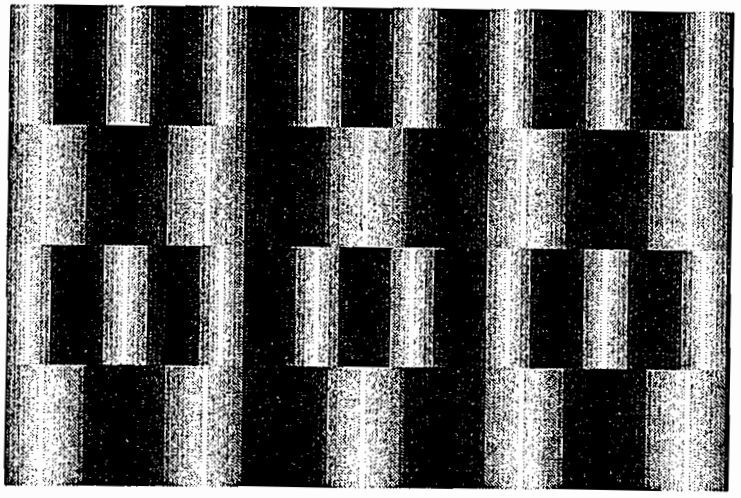
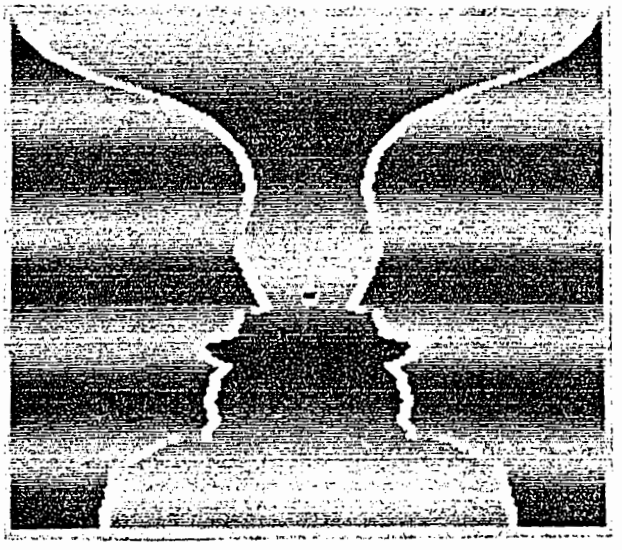


FIGURE 3. The higher spatial frequencies appear to be further away than the lower spatial frequencies.



(a)



(b)

FIGURE 4. The higher spatial frequencies appear to be closer than the lower spatial frequencies. (a) Adapted from Brown and Weisstein (1988b). (b) Reprinted with permission from Klymenko and Weisstein (1986).

How does a 2-D image create a 3-D percept of occluding objects in front of occluded objects, as in Figure 5(b)? How are the occluded objects recognized in Figure 5(b), but not Figure 5(c), even though they are equally well seen in both? A comparison of Figures 5(b) and 5(c) illustrates that properties of form, color, and depth interact to generate a percept, and that this interaction may, as in Figure 5(b), or may not, as in Figure 5(c), 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 purposes of recognition in response to Figure 5(b), but not 5(c), even though the occluded regions are not seen in either Figure 5(b) or 5(c).

The theory suggests how 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 B's from flowing "behind" the black occluder and thereby rendering it transparent, much as the nearer B boundary in the DaVinci stereopsis display of Figure 2 prevents filling-in of the surface BD into the region AB.

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

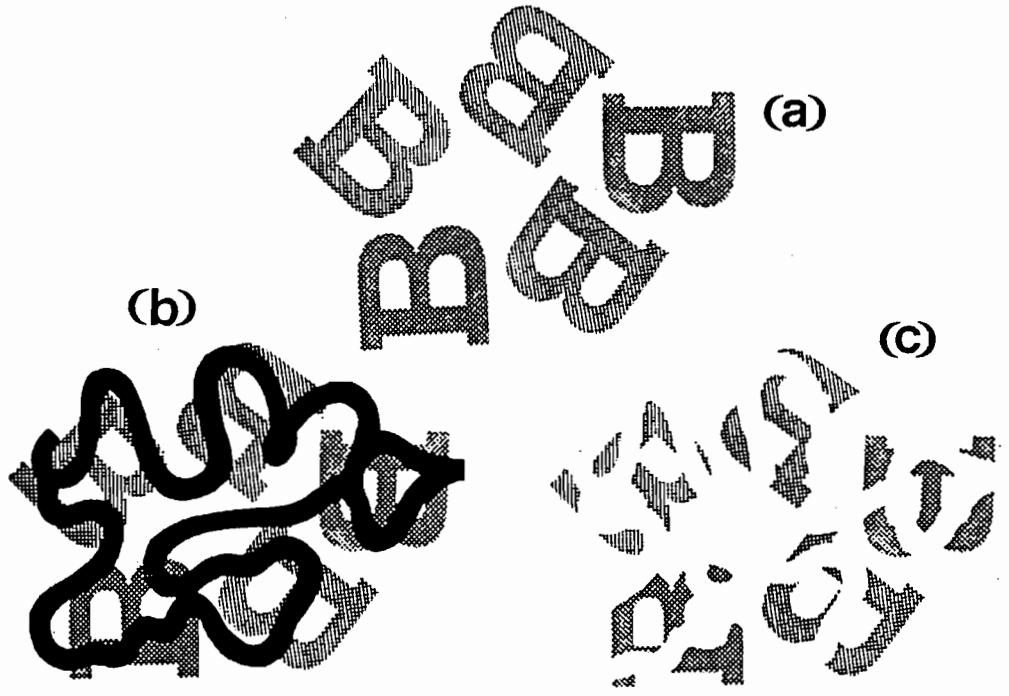


FIGURE 5. Role of occluding region in recognition of occluded letters: (a) Upper case "B" letters; (b) same, except partially hidden by a black snake-like occluder; (c) same, except occluder is white, and therefore merges with the remainder of the white background. Although the exposed portions of the letters are identical in (b) and (c), they are much better recognized in (b). (Reprinted with permission from Nakayama, Shimojo, & Silverman (1989).)

the case of the Bregman-Kanizsa B's, an interaction between the boundary segmentations of *differently colored regions* generates a 3-D percept from a 2-D image. We need to analyse 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.

5. OCCLUDED BOUNDARY COMPLETION AND RECOGNITION WITHOUT FILLING-IN

Even if the shared boundaries between occluder and B shapes in Figure 5(b) 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 6 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 (de-

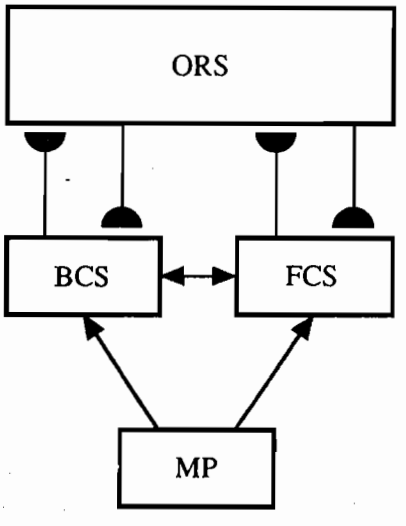


FIGURE 6. Completed boundaries within the Boundary Contour System (BCS) can be recognized within the Object Recognition System (ORS) via direct BCS → ORS interactions whether or not they are seen in the Feature Contour System (FCS) by separating two regions with different filled-in brightnesses or colors. The FCS → BCS interactions are introduced in this article.

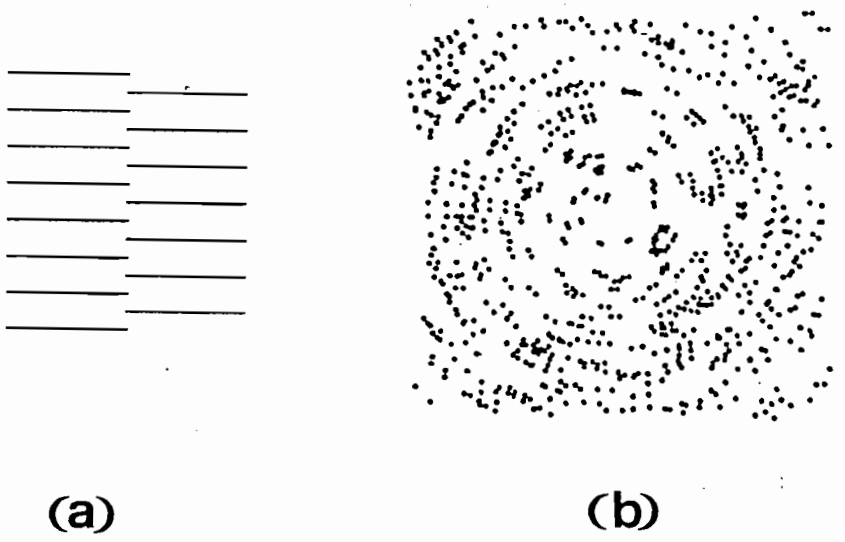


FIGURE 7. (a) The vertical line is easily recognized in the absence of a vertically oriented contrast difference. (b) A Glass pattern. The emergent circular pattern is recognized without being seen. (Reprinted with permission from Glass & Switkes (1976).)

noted FCS). The completed BCS boundary can directly activate the Object 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 prestriate 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. In Figures 7(a) and 7(b), for example, the vertical illusory boundary and the circular illusory groupings are vivid even though they do not correspond to large perceived contrast differences.

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). The theory suggests that 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. The theory traces this asymmetry between near and far to the same mechanism that prevents the nearer surface AB in the DaVinci 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. This mechanism is summarized in Figure 8. It is called *BF Intercopies* because BCS boundaries from multiple disparities converge on FCS filling-in domains in a partially ordered manner, such that boundary segmentations which correspond to a given depth obstruct filling-in of surface representations that correspond to that depth and all farther depths.

As noted, addition of boundaries from near to far surfaces is assumed 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 theory is to explain how this can happen in response to a 2-D pic-

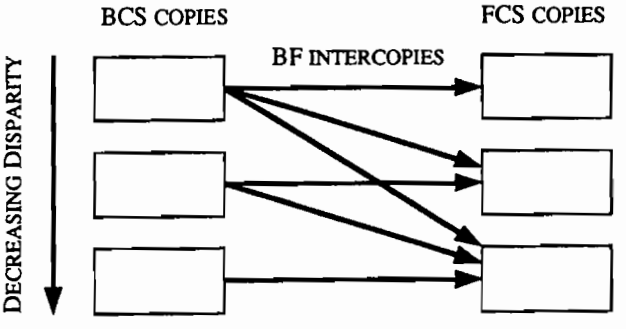


FIGURE 8. Each FCS copy receives inhibitory boundary-gating signals, or BF Intercopies, from one or more BCS copies. The BF Intercopy inputs are partially ordered from larger disparity to smaller disparity BCS copies. Each FCS copy contains three pairs of opponent Filling-In Domains (FIDOs).

ture 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. Along the way, the theory offers an explanation of why some surfaces do look transparent.

6. OVERVIEW OF FACADE THEORY

Perceptual properties such as those summarized above illustrate how 3-D segmentations and surface representations 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. The theory develops an earlier theory of 3-D preattentive vision that was introduced in Grossberg (1987a, 1987b). This theory has been called FACADE Theory because it suggests how visual representations of Form-And-Color-And-DEpth (FACADES) are generated in area V4 of the prestriate visual cortex (Figure 9). The theory describes the neural architecture of two parallel subsystems, the BCS and the 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, 1987a, 1987b, 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, 1987a, 1987b; Grossberg & Mingolla, 1985a; Grossberg & Todorović, 1988; Grossberg & Wyse, 1991, 1992).

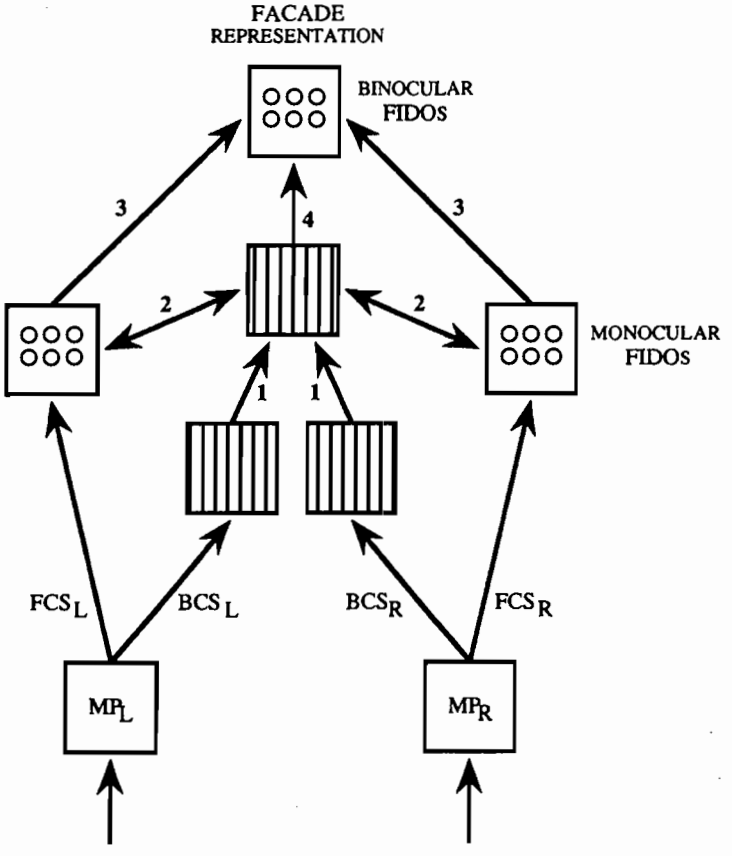


FIGURE 9. Macrocircuit of monocular and binocular interactions of the Boundary Contour System (BCS) and the Feature Contour System (FCS): Left eye and right eye monocular preprocessing stages (MP_L and MP_R) send parallel pathways to the BCS (boxes with vertical lines, designating oriented responses) and the FCS (boxes with three pairs of circles, designating opponent colors). The monocular signals BCS_L and BCS_R activate simple cells which, in turn, activate bottom-up pathways, labelled 1, to generate a binocular boundary segmentation using the complex, hypercomplex, and bipole cell interactions of Figure 10. The binocular segmentation generates output signals to the monocular Filling-In Domains, or FIDOs, of the FCS via pathways labelled 2. This interaction selects binocularly consistent FCS signals, and suppresses the binocularly inconsistent FCS signals. The surviving FCS signals activate the binocular FIDOs via pathways 3, where they interact with the binocular BCS segmentation to fill-in a multiple-scale surface representation of Form-And-Color-And-DEpth, or FACADE. Processing stages MP_L and MP_R 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 Inter stripes, 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 labelled 2, 3, and 4, and among FCS and BCS copies, are described in the text.

The BCS has been used to analyse and predict neurobiological data concerning the parvocellular processing stream from the 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 analyse 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 orientation, disparity, color, and motion 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 analysed elsewhere in terms of a Motion BCS (Grossberg & Mingolla, 1992; Grossberg & Rudd, 1989, 1992). In order 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 this paper.

Many experimental and modelling articles that have been 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; 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, Frisby, & Mayhew, 1989; Todd & Akerstrom, 1987), interactions between filling-in of brightness or color and illusory contour formation (Dresp, Lorenceau, & Bonnet, 1990; Kellman & Shipley, 1991; Nakayama et al., 1990; Prinzmetal, 1990; Prinzmetal & Boaz, 1989; Ramachandran, 1992; Shipley & Kellman, 1992; Takeichi, Shimojo, & Watanabe, 1992; Watanabe & Sato, 1989; Watanabe & Takeichi, 1990), interactions between depth, emergent segmentation and filling-in (Meyer & Dougherty, 1987; Nakayama et al., 1990; Takeichi et al., 1992; Watanabe & Cavanagh, 1992), orientation-specific luminance after-effects (Mikaelian, Linton, & Phillips, 1990), transient dynamics of filling-in (Arrington, 1992; 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 analysed 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 data about 3-D visual perception. These interactions are consistent with the previous theory and build upon it. Several investigators have described experimental evidence for the existence of interactions between scales; for example, Tolhurst (1972), Watt (1987), and Wilson, Blake, and Halpern (1991). 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. The predictive power of these rules derives from their ability to explain a large body of otherwise intractable perceptual data in a unified way. Different sets of experiments lend greater support to some rules than to others. Taken together, the rules as a whole are supported by a large body of perceptual data. 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 must be realized. These earliest possible stages are used in the neural predictions described in the theory.

In previous articles, the Static BCS was used to suggested 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 10). Functional roles for additional cell properties, such as end-stopped simple cells (Grossberg & Mingolla, 1992) 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 have been supported by subsequent neurophysiological experiments (von der Heydt et al., 1984; Peterhans & von der Heydt, 1989). The interactions within the simple-complex-hypercomplex cell module defines a static oriented contrast-sensitive filter, called the SOC Filter. This filter compensates for uncertainties of positional localization in the output of

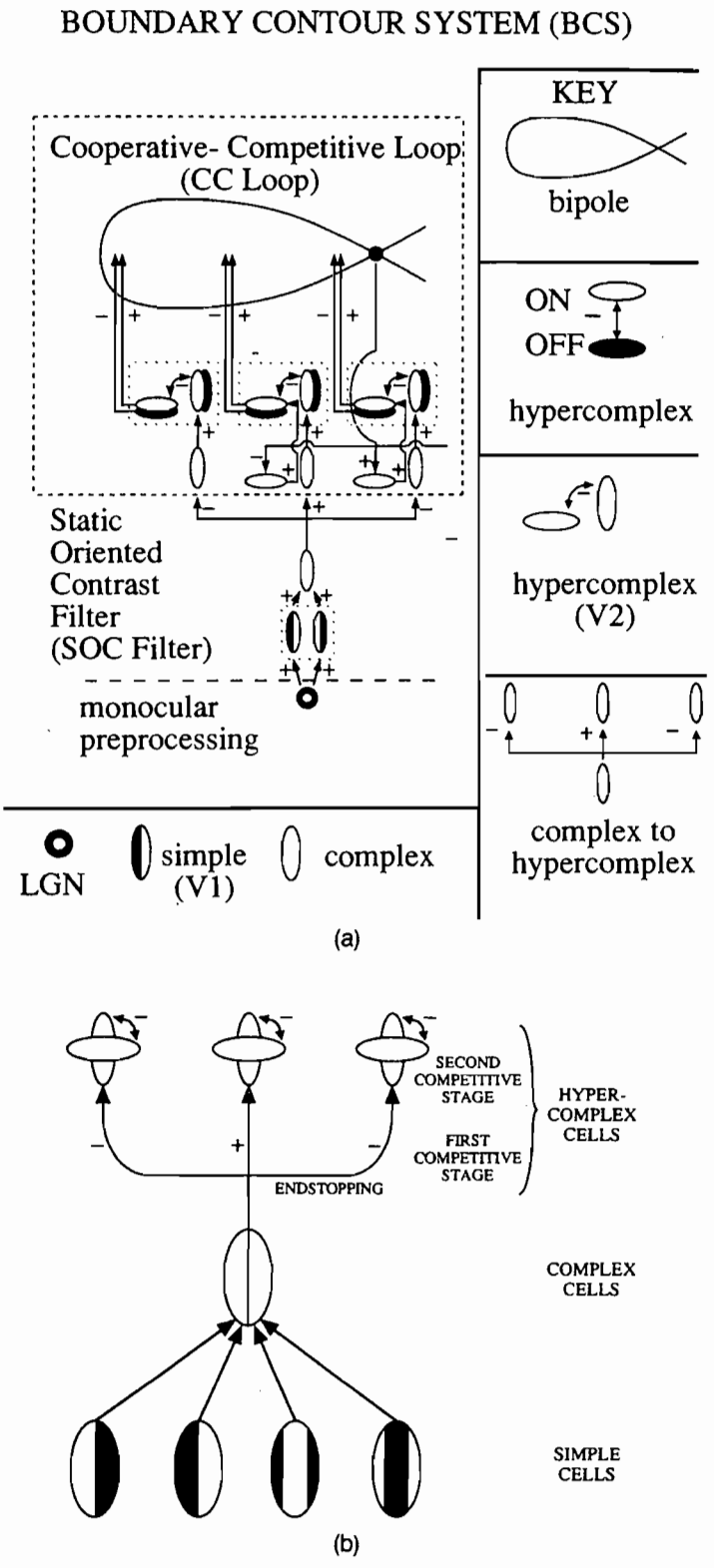


FIGURE 10. (a) The monocular Boundary Contour System of Grossberg and Mingolla (1985b). The circuit is divided into a static oriented contrast-sensitive filter (SOC Filter) followed by a cooperative-competitive feedback network (CC Loop). Multiple copies of the circuit are used, each corresponding to a different range of receptive field sizes. Each copy models interactions of simple cells, complex cells, hypercomplex cells, and bipole cells; (b) 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 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 endstopping, 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.

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, that 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 11).

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 matching of left and right image contrasts, as many AI vision theories had proposed. Rather, it is based upon 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 arise. 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, 1987b). Interactions of cortical ocular dominance columns, self-organizing feature maps, and monocular BCS mechanisms were shown to 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 was shown to obtain for the maximal disparity at which a scale of a given size can binocularly fuse monocular pairs of boundaries (Grossberg, 1987b; Grossberg & Marshall, 1989). Many data about binocular vision were comprehensible within this binocular BCS theory. The data explained by the present extension of FACADE Theory were not.

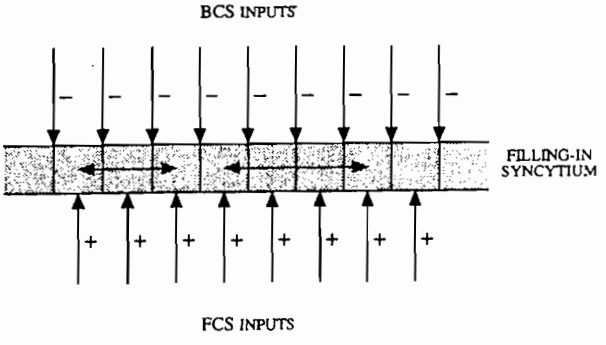


FIGURE 11. A monocular syncytium within the FCS. The Feature Contour signals are output signals from a shunting on-center off-surround network that discounts the illuminant. These signals activate cells that permit rapid electrotonic diffusion of activity, or potential, across their cell membranes, except at those membranes which receive Boundary Contour signals. The gap junctions at these membranes respond to the BC signals with an increase in resistance that decreases diffusion across them. Thus FC signals rapidly fill-in across syncytium cells until they reach a BCS boundary or are attenuated by their spatial spread.

7. INTERSCALE AND INTERSTREAM INTERACTIONS

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 analysed 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 data about 3-D visual perception.

The theory posits the existence of five new types of interactions which complement, and are consistent with, previously defined BCS and FCS mechanisms. 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 8 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.

The first interaction takes place among the complex cells of the BCS. Inhibitory competitive interactions are assumed to occur between complex cells that code different disparities at the same position and scale. These interactions are called *BB Intrascales*. As a result of this interaction, 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.

Interactions called *BB Interscales* are also predicted to occur. These 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 effect of these cooperative interactions on the competitive interactions of the SOC Filter (Figure 10), 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). These interactions are predicted to occur between the cortical Interblobs and Interstripes (Figure 1).

In the theory developed in Grossberg (1987b), each disparity-sensitive 3-D boundary segmentation, or BCS copy, interacts with a Monocular Filling-In-DOmain (FIDO) of the FCS, along the BCS → FCS pathways that are denoted in Figure 9 by 2. These BCS signals select those monocular brightness and color signals, labelled 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.

In addition, the theory herein posits that 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 DaVinci stereopsis in Section 2.6. These connected regions are assumed to occur at the Monocular FIDOs of Figure 9. The theory develops the hypothesis that the filled-in connected domains, which represent those monocular surface representations that are binocularly consistent, are the ones that 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 excite BCS cells corresponding to the same disparity and position, while inhibiting BCS cells corresponding to smaller disparities at that position. These FB Intercopies inhibit the BCS boundaries of any occluded region that occur at the same positions as the boundaries of an occluding region, such as the boundaries of the gray B's 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.

A possible neural locus for these BF Intracopies and FB Intercopies derives from 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 ↔ FCS interactions. The earliest possible cortical stage at which this could occur is 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 and FB Intracopy interaction is between the Blobs and Interblobs. The next possible cortical stage is between the Thin Stripes and Interstripes.

In addition to these BF and FB interactions, *FF Intercopies* are predicted to occur along the pathways labeled 3 in Figure 9. Excitatory output signals are generated, as in the case of FB Intercopies, at the boundaries of filled-in connected regions of the Monocular FIDOs. These excitatory signals activate Binocular FIDOs that correspond to the same disparity and position. These excitatory signals activate the filling-in of the 3-D surface representation. In addition, inhibitory signals suppress Binocular FIDOs corresponding 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. These FF Intercopies occur within the Blob cortical stream. They are initiated at, or later than, the same cortical stage that gives rise to FB Intercopies. They have their excitatory and inhibitory effects no later than area V4.

The final new interactions are called *BF Intercopies.* These are the BCS → FCS boundary signals from a given disparity and position that add to the BCS boundaries of all smaller disparities at that position (Figure 8) in order to prevent all nearer occluding surfaces from appearing transparent due to filling-in of their positions by the brightness and colors of farther occluded surfaces.

We now sketch an explanation of the data summarized in Sections 2–4. A more detailed analysis of theoretical mechanisms is provided in Grossberg (1992).

8. AN EXPLANATION OF BREGMAN-KANIZSA FIGURE-GROUND SEPARATION AND COMPLETION

First let us consider how the occluded gray B's in Figure 5 are seen and recognized on a surface behind the occluding black bands. Consider the image in Figure

12(a). 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 12(b) and 13(b)). 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 in Grossberg (1992) 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. 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 12(c) and 13(c), 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. (Such a multiscale disparity-sensitive competition was computationally simulated in Grossberg & 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 larger scales. This property is due to the size-disparity correlation (Richards & Kaye, 1974; Schor & Tyler, 1981; Schor & Wood, 1983; Schor, et al., 1984; Tyler, 1975, 1983). The smaller scale in Figure 12(c) was chosen so that it cannot fuse D_1 but it can 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 respond to the image contours. This can happen because there are no smaller scale complex cells that can fuse disparity D_1, and thus no BB Intrascale competition from disparity D_1 to D_2. Thus Figure 12(c) 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 fusable disparities winning over smaller ones.

Figures 12(d) and 13(d) show that end gaps, or holes in the boundary, are formed at the B boundaries as a result of CC Loop feedback. Both top-down bipole-to-hypercomplex competition between positions and hypercomplex-to-hypercomplex competition between orientations help to create these end gaps (see Figure 10).

In Figures 12(e) and 13(e), binocular BCS boundaries interact with monocular FCS signals via BF In-

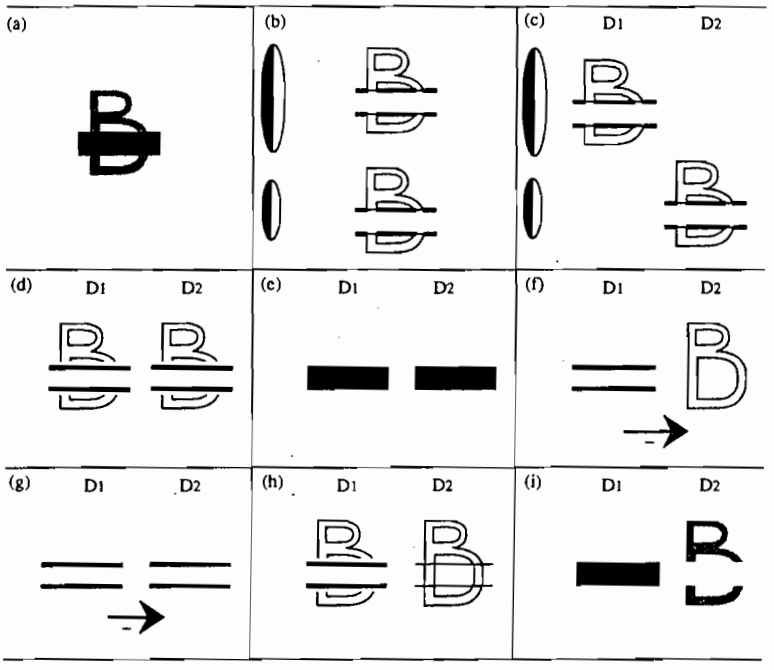


FIGURE 12. Bregman-Kanizsa figure-ground separation: (a) image; (b) monocular simple cell activations in the BCS; (c) complex cell activations after BB Intrascale competition from disparity D_1 to D_2; (d) CC Loop boundary segmentation at higher-order hypercomplex cells after end gaps form; (e) filling-in of connected components in monocular 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.

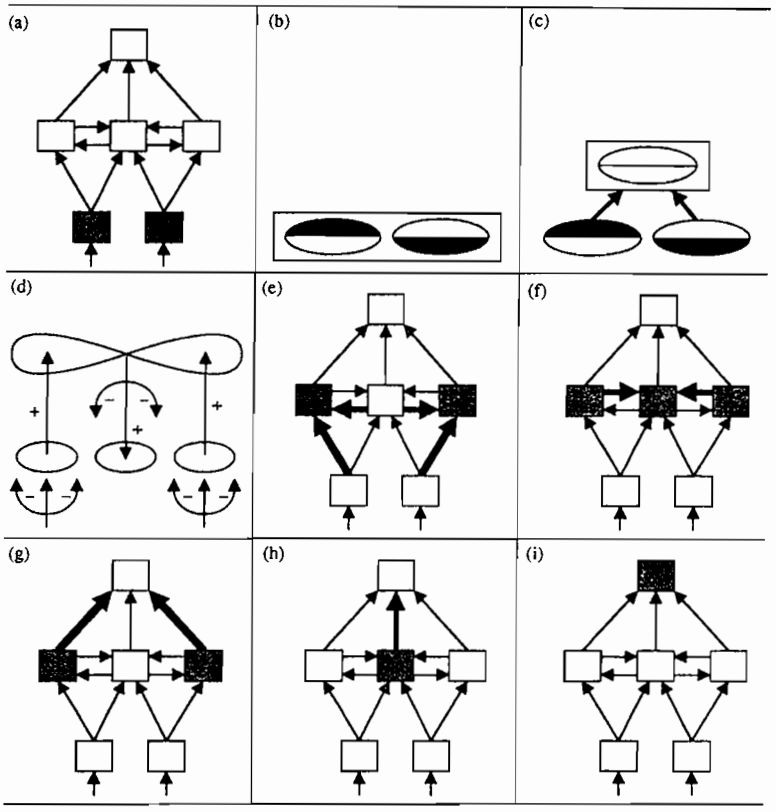


FIGURE 13. Active network stages during processing of a 3-D scene: (a) discounting of the illuminant occurs in the monocular preprocessing stages, notably the lateral geniculate nucleus; (b) simple cell activation; (c) complex cell activation; (d) emergent boundary segmentation by hypercomplex-bipole cell feedback in the CC Loop; (e) filling-in of the monocular syncytia by monocular FCS signals that are consistent with the binocular BCS segmentation; (f) FB Intercopies inhibit boundaries at smaller scales and disparities; (g) FF Intercopies excite filling-in of the corresponding binocular syncytia and inhibit monocular FCS signals at smaller disparities; (i) the final multiscale filled-in surface representation of Form-And-Color-And-DEpth emerges within the binocular syncytia.

tracopies to select those monocular FCS signals that are consistent with the binocular BCS boundaries. BCS boundaries hereby act as *filling-in generators* within the FCS; see the pathways labelled 2 in Figure 9. All other monocular FCS signals are suppressed. The selected FCS signals fill-in their respective filling-in domains, or syncytia. If end gaps in the regions exist, as in Figure 12(d), then the filling-in signals cross the gaps and dissipate across space unless they are contained by other nearby boundaries. Figure 12(e) 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 12(f) and 13(f). Output signals are hereby generated only at the boundaries of the black occluder. These FCS output signals activate parallel pathways that influence both the BCS and the FCS. The FB 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 12(f). These completed B boundaries generate direct BCS → ORS signals, as in Figure 6. Thus a completed letter B can be recognized at the ORS, even if only its unoccluded surfaces are seen at the FCS; cf., Biederman (1987).

Why is the letter B not completely seen at the FCS? This is due partly to FF Intercopies. As shown in Figures 12(g) and 13(g), FF Intercopies give rise to excitatory output pathways from both the left eye and right eye monocular filling-in domains. These output signals arise at contours of the filled-in connected components of the monocular filling-in-domains. They thus delineate both the locations and the perceptual qualities of monocular surface components that are consistent with the binocular BCS segmentation. These monocular output signals are binocularly matched at the binocular filling-in domains of the FCS. 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 FACADE at the binocular filling-in domain (Figure 9). In addition, FF Intercopies inhibit all the FCS signals at their position which correspond to smaller disparities. These inhibitory interactions may

possibly be triggered within the binocular filling-in domains as part of an on-center off-surround response to the excitatory FF Intercopies. 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.

Why cannot FCS signals from smaller disparities, but different positions, fill-in behind a nearer occluding surface? This is due partly to BF Intercopies, which add their boundary signals to the binocular syncytia of smaller disparities, as in Figures 12(h) and 13(h). These BF Intercopies are inhibitory signals, just like the FB Intercopies. Inhibitory signals to an FCS syncytium create barriers to filling-in at their target cells (Cohen & Grossberg, 1984; Grossberg, 1987a; Grossberg & Todorović, 1988). As a result, in Figure 12(h), complete boundaries of both the occluding band and the occluded B exist at the smaller disparity.

The BF Intercopies and FF Intercopies of Figures 12(g) and 12(h) work together to generate the binocular filling-in events shown in Figures 12(i) and 13(i). The B surface is filled-in at disparity D_2 only where it is not occluded, due to BF Intercopies. The occluding surface is not filled-in at all at disparity D_2, due to FF Intercopies. 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.

9. AN EXPLANATION OF DAVINCI STEREOPSIS

The same mechanisms can now be used to explain the 3-D percept of the DaVinci stereopsis image in Figure 2, with one addition: the interaction of near-zero disparity cells will be emphasized. Figure 14 outlines the main steps of the explanation. Figure 14(a) 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 8 for edges on the same planar surface. These fused boundaries are represented in Figure 14(b). 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 14(b) shows the complex cell activations at both scales and disparities.

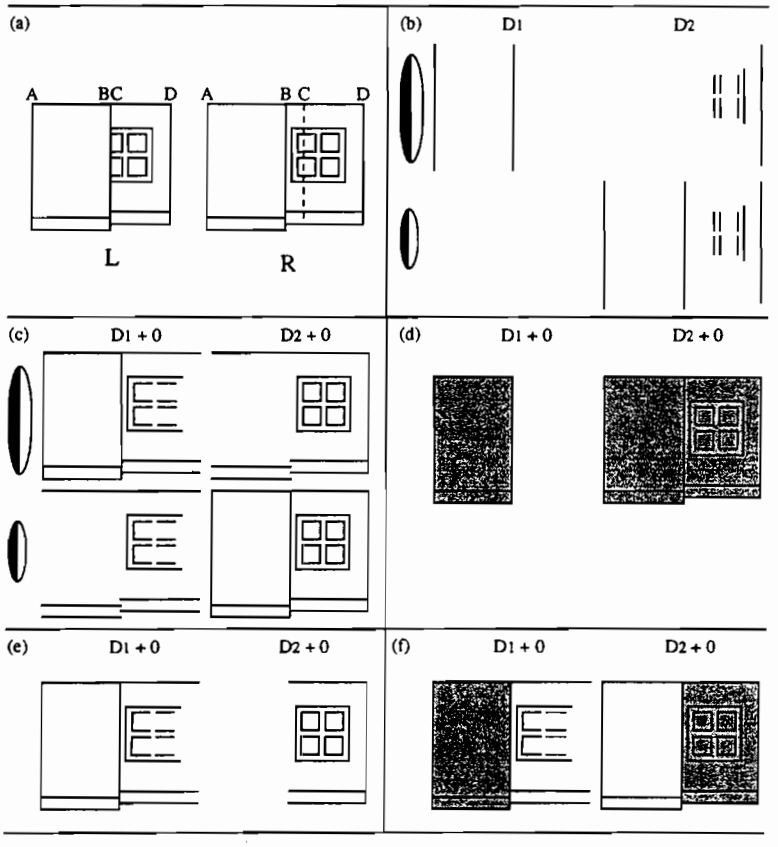


FIGURE 14. (a) Left and right eye views of a scene. Region AB is closer than region BCD, and region BC is monocularly viewed; (b) binocularly fused complex cell responses of nonzero disparity cells at two scales and disparities; (c) combination of fused non-zero disparity responses with near-zero disparity 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 interscales from connected regions; (f) overlay of final BCS boundary representation and filled-in surface representations at the binocular syncytia.

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 14(b) 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 in Figure 14(c) 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 14(c) assumes that allelotropia has deformed the binocularly 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 (1987b).

The CC Loop does not substantially change the boundary representation of Figure 14(c) 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 14(d) indicates the regions of Figure 14(c) that can be successfully filled-in within the monocular syncytia, as in Figure 13(e). Figure 14(e) describes the boundaries that survive the inhibition due to FB Intercopies, as in Figure 13(f). A similar inhibition of FCS signals for region AB occurs at disparity D_2 due to FF Intercopies, as in Figure 13(g). Figure 14(f) shows the effect of BF Intercopies on the final connected boundary segmentations, as in Figure 13(h), and the final filling-in of the binocular syncytia, as in Figure 13(i). 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.

10. AN EXPLANATION OF THE CLOSER APPEARANCE OF HIGHER SPATIAL FREQUENCIES THAN LOWER SPATIAL FREQUENCIES

An explanation of the depthful spatial frequency percepts that were described in Section 3 can also be derived from these mechanisms. The explanation begins by noting 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 15(a)). As a result, the vertically oriented complex cells that are activated by the high spatial frequency sinusoid inhibit the contiguous vertically oriented hypercomplex cells that are activated by the low spatial frequency sinusoid, more than conversely (Figure 10). End gaps hereby begin to form at these locations (Figure 15(b)). 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 Grossberg & Mingolla (1987a) for computer simulations of boundary webs).

The asymmetric inhibition of hypercomplex cells at the first competitive stage (Figure 10) enables the higher-order hypercomplex cells at the second competitive stage to form end cuts that bound the high frequency sinusoids (Figure 15(b)). 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 15(c)). 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 15(d)). These completed low spatial frequency boundaries can be recognized via the direct BCS → ORS pathway (Figure 6). 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 15(e)). Hence, the high 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. Habituative transmitter gates exist in the pathways to the hypercomplex cells of the second competitive stage and in the bipole cell feedback pathways (Grossberg, 1987a). These habituative transmitter gates help to trigger reset of a boundary segmentation when stimulus conditions

change (Francis, Grossberg, & Mingolla, 1992; Grossberg, 1991). 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 6) to influence such a bistable depth percept. A shift in spatial 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.

11. CONCLUDING REMARKS: BCS AND FCS FROM A COMPUTATIONAL AND PAINTERLY PERSPECTIVE

This paper outlines a solution to the classical 3-D figure-ground problem of biological vision. It does so within the context of a neural theory of biological vision, called FACADE Theory, that predicts the types and ordering of interactions, as in Figure 13, that may occur within and between the two parvocellular cortical processing streams from LGN to cortical area V4 (Figure 1). In so doing, the theory suggests explanations of a large and paradoxical database from visual psychophysics and neurobiology that has not been explained by alternative theories.

These explanations are derived from an ecological analysis of how monocularly viewed parts of an image fill-in the appropriate depth from contiguous binocularly viewed parts, as during DaVinci stereopsis. The explanations can be developed as part of an analysis of how the two parvocellular processing streams that join LGN to V4 interact to generate a multiplexed representation of FACADE within area V4. The two parvocellular streams are modelled by a BCS and an FCS. The BCS generates emergent boundary segmentations that combine edge, texture, shading, and stereo information. The FCS discounts the illuminant and fills-in surface properties of brightness, color, and depth. The ensemble of all surface representations constitutes the FACADE representation. The BCS and FCS interact reciprocally via adaptive filters with an ORS, interpreted to occur in inferotemporal cortex, to bind these segmentation and surface properties together.

It is shown how interactions between BCS and FCS, especially partially ordered interactions form larger scales and disparities to smaller scales and disparities, inhibit spurious boundary and surface signals. Additional new ideas include the observations that filled-in connected regions at a given disparity inhibit the boundaries and features of smaller disparity representations; near-zero disparity cell pools and nonzero disparity cell pools interact to generate boundary segmentations; the cortical magnification factor helps to convert different disparity computations at different foveal eccentricities into a planar surface representation; multiple receptive field sizes cooperate to generate positionally accurate segmentations; and networks of simple cells, complex cells, hypercomplex cells, higher-order hypercomplex cells, and bipole cells generate boundary segmentations that organize surface representations into ecologically useful 3-D percepts.

From a broader perspective, it is important to keep in mind that the BCS and the FCS compute complementary ways of representing the visual world (Grossberg et al., 1989). BCS/FCS complementarity, and the intertwined interactions that are used to resolve it in a higher synthesis of visual representation, have been harmoniously illustrated and commented upon in the work of many artists. Matisse was a particularly great observer and innovator in resolving this complementarity through esthetic means (Elderfield, 1992). When Matisse wrote about "the external conflict between drawing and color," we can interpret his insights in terms of the complementary properties of BCS segmentations and FCS filling-in. The Fauvist movement made explicit through a painterly technique the realization that one could draw with color; that color strips could directly group into emergent segmentations which, in turn, could support the colors themselves as part of vivid surface representations, undimmed by lines whose sole purpose would be to segment them. Matisse's famous paintings of 1905, The Open Window and the Roofs of Collioure, masterfully blended direct applications of discrete color strips and continuous surface colors to generate harmonious surface representations in which both types of strokes are unified.

Matisse's lifelong interest in these complementary processes was vividly expressed towards the end of his life in his 1947 book "Jazz." Here he wrote as follows about the exquisite paper cutout maquettes that he created during this period: "Instead of drawing an outline and filling in the color . . . I am drawing directly in color." Put somewhat more technically, the bifurcation from the Monocular Preprocessing Stages (MP_L and MP_R) in Figure 9, at which the illuminant is discounted, to both the BCS and the FCS, show how one can "draw directly in color" by activating both BCS segmentations and FCS filling-in of surface properties, instead of drawing an outline, which primarily activates the BCS before "filling-in the color" by primarily activating the FCS. One of the unexpected pleasures of studying FA-

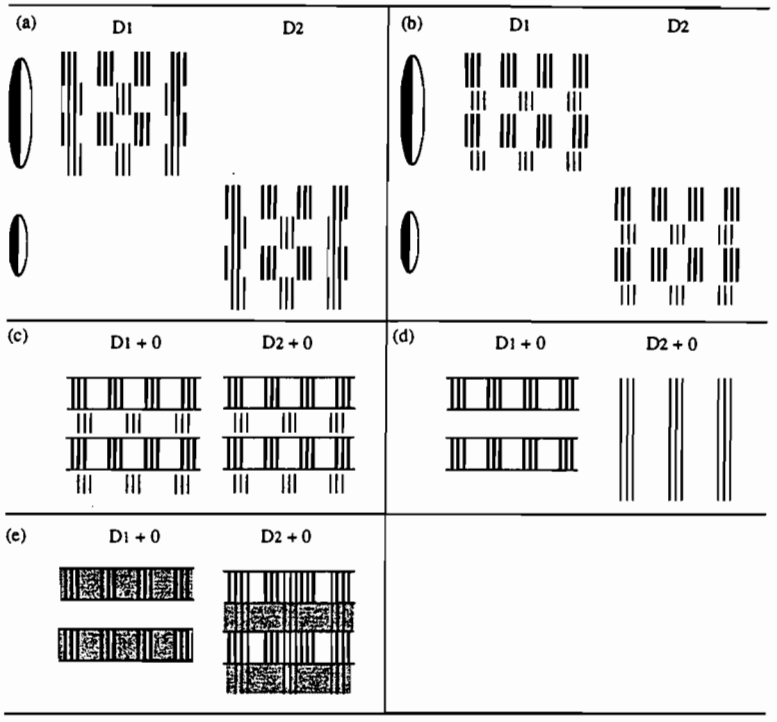


FIGURE 15. Why high spatial frequency inputs appear closer than low spatial frequency inputs: (a) Complex cells at the larger scale and disparity respond more strongly to the higher spatial frequency. The larger scale, smaller disparity cells do not respond due to inhibition from BB intrascales. The 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 abut the high spatial frequency activations of higher-order hypercomplex cells; (d) FB intercopies from the filled-in large disparity FCS copy to the smaller disparity BCS cells inhibit the high spatial frequency responses there, and enable the smaller disparity 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.

CADE Theory derives from the light it sheds on the scientific basis of painterly discoveries through the ages.

REFERENCES

Arrington, K. (1992). *Neural network models for color and brightness perception and binocular rivalry.* Unpublished doctoral dissertation, Boston University, Boston, MA.

Beck, J., Graham, N., & Sutter, A. (1991). Lightness differences and the perceived segregation of regions and populations. *Perception and Psychophysics,* **49,** 257–269.

Beck, J., Prazdny, K., & Ivry, R. (1984). The perception of transparency with achromatic colors. *Perception and Psychophysics,* **35,** 407–422.

Beck, J., Rosenfeld, A., & Ivry, R. (1990). Line segregation. *Spatial Vision,* **4,** 75–101.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review,* **94,** 115–147.

Bregman, A. L. (1981). Asking the "what for" question in auditory perception. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 99–118). Hillsdale, NJ: Erlbaum Associates.

Brown, J. M., & Weisstein, N. (1988a). A phantom context effect: Visual phantoms enhance target visibility. *Perception and Psychophysics,* **43,** 53–56.

Brown, J. M., & Weisstein, N. (1988b). A spatial frequency effect on perceived depth. *Perception and 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., & Mehanian, C. (1989). Invariant recognition of cluttered scenes by a self-organizing ART architecture: CORT-X boundary segmentation. *Neural Networks,* **2,** 169–181.

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

Cruthirds, D., Gove, A., Grossberg, S., & Mingolla, E. (1991). Preattentive texture segmentation and grouping by the boundary contour system. *Proceedings of the International Joint Conference on Neural Networks, Seattle,* **I,** 655–660.

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.

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.

Elderfield, J. (1992). *Henri Matisse: A retrospective.* New York: The Museum of Modern Art.

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 chro-

matic detection by luminance contours. *Journal of the Optical Society of America A*, **8**, 394–403.

Francis, G., Grossberg, S., & Mingolla, E. (1992). Cortical dynamics of feature binding and reset: Control of visual persistence. Submitted for publication.

Gibson, J. J. (1950). *Perception of the visual world.* Boston, MA: Houghton Mifflin.

Gillam, B., & Borsting, E. (1988). The role of monocular regions in stereoscopic displays. *Perception*, **17**, 603–608.

Glass, L., & Switkes, E. (1976). Pattern recognition in humans: Correlations which cannot be perceived. *Perception*, **5**, 67–72.

Graham, N., Beck, J., & Sutter, A. (1992). Nonlinear processes in spatio-frequency channel models of perceived texture segregation: Effects of sign and amount of contrast. *Vision Research*, in press.

Grossberg, S. (1980). How does a brain build a cognitive code? *Psychological Review*, **87**, 1–51.

Grossberg, S. (1984). Outline of a theory of brightness, color, and form perception. In E. Degreef & J. van Buggenhaut (Eds.), *Trends in mathematical psychology* (pp. 5559–5586). Amsterdam: Elsevier/North-Holland.

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

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

Grossberg, S. (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 and Psychophysics*, **49**, 117–141.

Grossberg, S. (1992). 3-D vision and figure-ground separation by visual cortex (Tech. Rep. No. CAS/CNS-TR-92-019). Boston, MA: Boston University.

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 and Psychophysics*, **38**, 141–171.

Grossberg, S., & Mingolla, E. (1987a). Neural dynamics of surface perception: Boundary webs, illuminants, and shape-from-shading. *Computer Vision, Graphics, and 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, IV* (pp. 177–184). Piscataway, NJ: IEEE Service Center.

Grossberg, S., & Mingolla, E. (1992). Neural dynamics of motion perception: Direction fields, apertures, and resonant grouping. *Perception and Psychophysics*, 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., & Todorović, D. (1988). Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena. *Perception and Psychophysics*, **43**, 241–277.

Grossberg, S., & Wyse, L. (1991). Invariant recognition of cluttered scenes by a self-organizing ART architecture: Figure-ground separation. *Neural Networks*, **4**(6), 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 (Eds.), *Neural networks for vision and image processing* (pp. 161–194). Cambridge, MA: MIT Press.

Humphreys, G. W., Quinlan, P. T., & Riddoch, N. J. (1989). Grouping processes in visual search: Effects with single- and combined-feature targets. *Journal of Experimental Psychology: General*, **118**, 258–279.

Kanizsa, 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.

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.

Klymenko, V., & Weisstein, N. (1986). Spatial frequency differences can determine figure-ground organization. *Journal of Experimental Psychology: Human Perception and Performance*, **12**, 321–330.

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

Lawson, R. B., & Gulick, W. L. (1967). Stereopsis and anomalous contour. *Vision Research*, **1**, 271–297.

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.

Marr, D., & Poggio, T. (1979). A computational theory of human stereo vision. *Proceedings of the Royal Society of London (B)*, **204**, 301–328.

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., DaPos, O., & Cavedon, A. (1985). Balanced and unbalanced, complete and partial transparency. *Perception and 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 and Performance*, **13**, 353–360.

Meyer, G. E., & Senecal, M. (1983). The illusion of transparency and chromatic subjective contours. *Perception and Psychophysics*, **34**, 58–64.

Mikaelian, H. H., Linton, M. J., & Phillips, M. (1990). Orientation-specific luminance after-effects. *Perception and Psychophysics*, **47**, 575–582.

Mishkin, M. (1982). A memory system in the monkey. *Philosophical Transactions Royal Society of London B*, **298**, 85–95.

Mishkin, M., & Appenzeller, T. (1987). The anatomy of memory. *Scientific American*, **256**, 80–89.

Nakayama, K., & Shimojo, S. (1990). DaVinci 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.

Paradiso, M. A., & Nakayama, K. (1991). Brightness perception and filling-in. *Vision Research*, **31**, 1121–1236.

Peterhans, E., & von der Heydt, R. (1989). Mechanisms of contour perception in monkey visual cortex, II: Contours bridging gaps. *The Journal of Neuroscience*, **9**, 1749–1763.

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.

Prinzmetal, W. (1990). Neon colors illuminate reading units. *Journal of Experimental Psychology: Human Perception and 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.

Ramachandran, V. S. (1992). Perception: A biological perspective. In G. A. Carpenter & S. Grossberg, S. (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.

Regan, D., Erkelens, C. J., & Collewijn, H. (1986). Visual field defects for vergence eye movements and for stereomotion perception. *Investigative Ophthalmology and Visual Science*, **27**, 806–819.

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.

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. (Ed.). (1990). *Computational neuroscience*. Cambridge, MA: MIT Press.

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**, 576–5778.

Shipley, T. F., & Kellman, P. J. (1992). Perception of partly occluded objects and illusory figures: Evidence for an identity hypothesis. Submitted for publication.

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

Takeichi, H., Shimojo, S., & Watanabe, T. (1992). Neonflank and illusory contour: Interaction between the two processes leads to color filling-in. *Perception*, **21**, 313–324.

Takeichi, H., Watanabe, T., & Shimojo, S. (1992). Illusory occluding contours and surface formation by depth propagation. *Perception*, **21**, 177–184.

Tausch, R. (1953). Die beidaugige Raumwahrnehmung-ein Prozess auf Grund der Korrespondenz und Disparation von Gestalten anstelle der Korrespondenz oder Disparation einzelner Netzhautelemente. *Zeitschrift fur experimentelle und angewandte Psychologie*, **1**, 394–421.

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

Ts'o, D. Y. (1989). The functional organization and connectivity of color processing. In D. M.-K. Lam & Gilbert, C. D. (Eds.), *Neural mechanisms of visual perception: Proceedings of the Retina Research Foundation Symposia, Volume* 2 (pp. 87–115). The Woodlands, TX: Portfolio Publishing Co.

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. Cuiffreda (Eds.), *Vergence eye movements* (pp. 199–295). Boston, MA: Butterworths.

von der Heydt, R., Peterhans, E., & Baumgartner, G. (1984). Illusory contours and cortical neuron responses. *Science*, **224**, 1260–1262.

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

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 and Psychophysics*, **45**, 427–430.

Watanabe, T., & Takeichi, H. (1990). The relation between color spreading and illusory contours. *Perception and 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.

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

Wheatstone, C. (1838). On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society (London)*, **128**, 371–394.

Wilde, K. (1950). Der Punktreiheneffekt und die Rolle der binocularen Querdisparation beim Tiefenshen. *Psychologische Forschung*, **23**, 223–262.

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.

Zeki, S. (1983a). Colour coding in the cerebral cortex: The reaction of cells in monkey visual cortex to wavelengths and colours. *Neuroscience*, **9**, 741–765.

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.