

Neural dynamics of 3-D surface perception: Figure-ground separation and lightness perception

FRANK KELLY and STEPHEN GROSSBERG
Boston University, Boston, Massachusetts

This article develops the FACADE theory of three-dimensional (3-D) vision to simulate data concerning how two-dimensional pictures give rise to 3-D percepts of occluded and occluding surfaces. The theory suggests how geometrical and contrastive properties of an image can either cooperate or compete when forming the boundary and surface representations that subserve conscious visual percepts. Spatially long-range cooperation and short-range competition work together to separate boundaries of occluding figures from their occluded neighbors, thereby providing sensitivity to T-junctions without the need to assume that T-junction "detectors" exist. Both boundary and surface representations of occluded objects may be amodally completed, whereas the surface representations of unoccluded objects become visible through modal processes. Computer simulations include Bregman-Kanizsa figure-ground separation, Kanizsa stratification, and various lightness percepts, including the Munker-White, Benary cross, and checkerboard percepts.

Since the Paleolithic era, humans have endeavored to represent the three-dimensional (3-D) world with two-dimensional (2-D) pictures and line-drawings. The primary goal of the present article is to understand how a 2-D picture can generate a percept of a 3-D scene in which figure-ground separation of visual surfaces occurs. This is accomplished by developing the FACADE theory of biological vision (Grossberg, 1987, 1994, 1997). FACADE stands for the form-and-color-and-depth representations that the network constructs from two monocular retinal images and that are multiplexed together within the visual cortex.

We show how the FACADE model, which was designed to work with 3-D stereoscopic inputs, can also extract figure-ground relations (e.g., stratification of an object in depth or partial occlusion) from 2-D images. Grossberg (1997) qualitatively developed the FACADE network to better understand how a partially occluded object in a 2-D image can be perceptually completed behind an occluding object, even if the completed representation is not seen as a consciously visible color or contrast difference. Such a completion event has been termed

an *amodal* percept (Kanizsa, 1979; Michotte, Thines, & Crabbe, 1964), to distinguish it from *modal* percepts that do carry a perceptually visible sign. Even before these representations were given a name, it was known that certain areas of visual space had a dual representation, or "duo-representation" (Koffka, 1935); that is, in a region in which an occluding object overlaps an occluded object, the visual area where the objects intersect is twice represented, once as belonging to the occluder and once as part of the occluded object. This article quantitatively develops the theory in order to provide rigorous explanations and simulations of key figure-ground percepts that are derived from 2-D images.

Amodal representations are simulated for various examples, including the Kanizsa stratification display (Figure 1). To date, no model has shown quantitative simulations of how such representations can be created by the visual cortex. Another classic example involving amodal completion mechanisms is the Bregman-Kanizsa display (Figure 2) (Bregman, 1981; Kanizsa, 1979). Network simulations of this input demonstrate how these amodal representations are created and may be used to aid in the recognition of partially occluded objects. The model can also explain various lightness illusions, such as the Benary cross, Munker-White assimilation, and the checkerboard display (Figure 3). These simulation results suggest how 2-D and 3-D figure-ground relationships can be explained in a unified way by the FACADE model. The model does this by showing how contrastive and geometrical properties of images may be used by the visual system to create boundary and surface representations that are mutually consistent. It also clarifies how T-junction and X-junction sensitivity, often cited as being cues for occlusion and transparency, can be coded in a cortical network without explicit T-junction and X-junction operators. These results were briefly presented in Kelly and

F.K. was supported in part by the Defense Advanced Research Projects Agency and the Office of Naval Research (ONR N00014-95-1-0409), the National Science Foundation (NSF IRI 94-01659), and the Office of Naval Research (ONR N00014-92-J-1309 and ONR N00014-95-1-0657). S.G. was supported in part by the Defense Advanced Research Projects Agency and the Office of Naval Research (ONR N00014-95-1-0409), the National Science Foundation (NSF IRI 97-20333), and the Office of Naval Research (ONR N00014-95-1-0657). The authors acknowledge Dan Cruthirds, Niall McLoughlin, Rajeev Raizada, and Brad Rhodes for their assistance with the FACADE implementation and Robin Amos and Diana Meyers for their assistance with the manuscript preparation. Correspondence concerning this article should be addressed to S. Grossberg, Department of Cognitive and Neural Systems, Boston University, 677 Beacon Street, Boston, MA 02215 (e-mail: steve@cns.bu.edu).

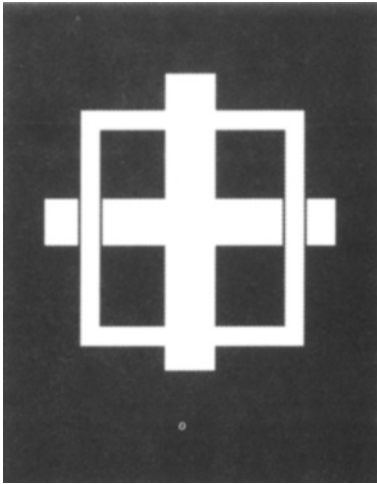


Figure 1. An example of perceptual stratification. From "Seeing and Thinking," by G. Kanizsa, 1985, *Rivista di psicologia*, 49, p. 8. Copyright 1985 by Praeger/Greenwood Publishing Group. Reprinted with permission.

Grossberg (1997, 1998). The percepts analyzed herein can be perceived either monocularly or binocularly. Grossberg and Kelly (1999) discuss related binocular properties of surface brightness perception.

QUALITIES OF FIGURE-GROUND PERCEPTION

The human visual system can perceive many different qualities of a surface: Texture, depth, orientation, lightness, illumination direction, opacity, color, movement, and occlusion relationships are just some of the general surface properties that can be perceived. Spatial or temporal changes in these surface properties can lead to differing segmentations of a visual scene in which certain objects are seen as a figure against a background. This section suggests how geometrical and contrastive scenic properties are employed by the visual system to allow us to separate figure from ground in 2-D pictures.

Lightness and Depth

Understanding how the visual system computes surface color and reflectance is an area of intense debate and research (Gilchrist, 1994). The perception of surface reflectance, or lightness, is affected by nearby or surrounding surfaces—for example, as during simultaneous contrast (Hering, 1920/1964). Surface lightness and contrast can also affect the perception of depth in paintings and natural scenes (O'Shea, Blackburn, & Ono, 1994). In the simple example of a cross (Figure 4a), the horizontal white bar is perceived as closer, and the two gray

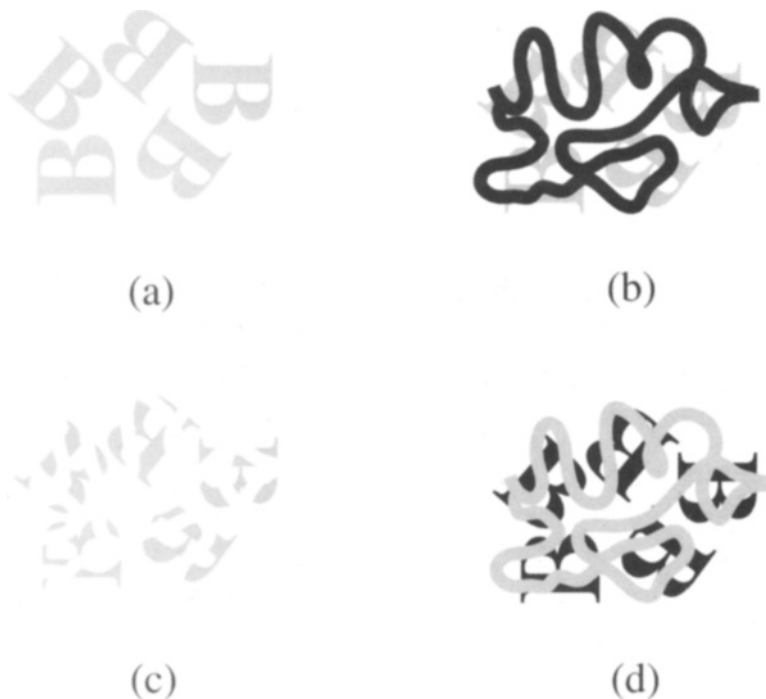


Figure 2. Bregman-Kanizsa display: (a) unoccluded Bs, (b) occluded B shapes, (c) B fragments, and (d) Occluded B shapes with different contrast. Panel c is from "Stereoscopic Depth: Its Relation to Image Segmentation, Grouping and the Recognition of Occluded Objects," by K. Nakayama, S. Shimojo, and G. H. Silverman, 1989, *Perception*, 18, p. 56. Copyright 1989 by Pion Ltd. Reprinted with permission.

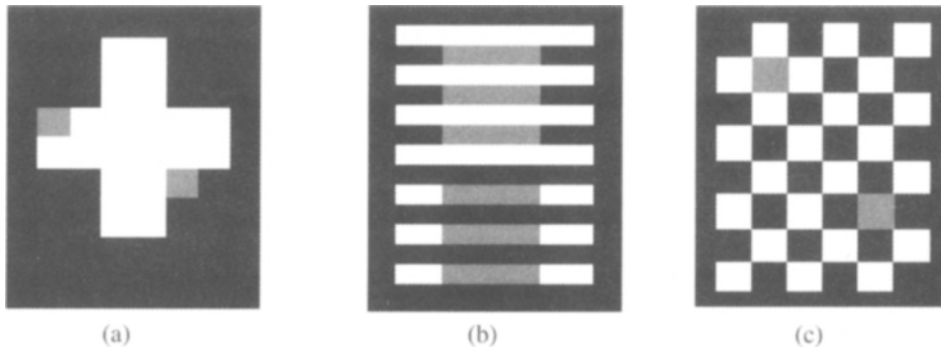


Figure 3. Lightness illusions that are often attributed to monocular depth cues: (a) Benary cross, (b) White assimilation display, (c) checkerboard pattern. See the text for details. Panel a is from “The Influence of Form on Brightness Contrast,” by W. Benary, in *A Source Book of Gestalt Psychology* (p. 104), by W. Ellis (Ed.), 1939, London: Routledge & Kegan Paul. Copyright 1939 by Routledge & Kegan Paul. Adapted with permission. Panel b is from “A New Effect of Pattern on Perceived Lightness,” by M. White, 1979, *Perception*, 8, p. 414. Copyright 1979 by Pion Ltd. Reprinted with permission.

vertical bars appear to be joined into a single larger bar that is partially occluded by the horizontal one. The vertical bar is said to be *amodally completed* behind the horizontal bar (Kanizsa, 1979), since we perceive the continuation of the gray bar without any modal or visible sign. Several researchers have proposed that geometric properties, such as T-junctions, are cues for figure-ground separation and amodal completion (Guzman, 1968/1984; Nakayama, Shimojo, & Ramachandran, 1990; Nakayama, Shimojo, & Silverman, 1989; Von Helmholtz, 1962; Watanabe & Cavanagh, 1993). T-junctions are created at the border between two overlapping lines or surfaces of different colors. In Figure 4a, four T-junctions are created where the white bar and the vertical gray stripes meet. The white bar boundary creates the top of the T-junction,

and each vertical gray stripe boundary forms the stem of a T. Traditionally, when figure-ground separation occurs, the T-junction is “split” so that the top is assigned to an occluding object and the stem is assigned to the partially occluded object (Nakayama et al., 1989).

Contrast can also influence perceived depth and figure-ground perception (Egusa, 1983). In particular, lighter or brighter objects appear closer while on a dark background. In Figure 4a, the geometric cues (T-junctions) indicate that the horizontal bar is occluding the vertical stripe. This cue is in agreement with the contrastive cue that the white object is closer than the gray one. These cues, in unison, result in a stable perceptual stratification of the white bar in front of the gray stripe. However, in Figure 4b, when one is asked which is perceived as be-

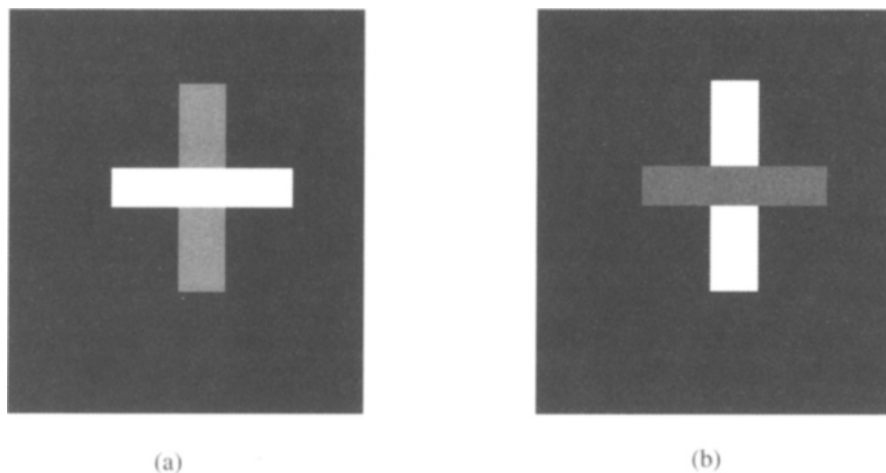


Figure 4. Pop-out and amodal completion: (a) pop-out of white bar and amodal completion of gray bar and (b) The same image with the bar colors interchanged.

ing closer, the perception is more bistable, since the geometric and the contrastive cues are no longer in agreement. The geometric relations (T-junctions) remain the same, but the change in relative contrast is a cue that the brighter vertical pieces are closer.

That these geometric and contrastive properties can cooperate or compete is also shown by the Kanizsa stratification (Kanizsa, 1985) images (Figure 1), wherein geometric and contrastive cues again lead to depthful percepts. Here the percept is one of a square weaving over and under the cross. This image is interesting because a single globally unambiguous figure-ground percept of one object's being in front (cross or thin outline square) does not occur. On the left and right arms of the cross in Figure 1, the contrastive vertical black lines are cues that the outline square is in front of the cross arms. The top and bottom regions consist of a homogeneously white figural area, but most observers perceive two figures, the cross arms in front of the thinner outline square. This is usually attributed to the fact that a thinner structure tends to be perceived behind a thicker one most of the time (Petter, 1956; Tommasi, Bressan, & Vallortigara, 1995). The figure-ground stratification percept is bistable through time, flipping intermittently between alternative cross-in-front and square-in-front percepts. We explain how this perceptual stratification of a homogeneously colored region occurs and how the visual system knows which depth to assign the surface color in different parts of the display.

So far, we have illustrated how lightness differences can affect depth. Other results suggest that depth can also affect perceived lightness. Schirillo, Reeves, and Arend (1990) showed that lightness matches are based on relationships among coplanar surfaces and not just retinally adjacent regions. Gilchrist (1977) formulated a computational rule called the *coplanar ratio hypothesis*, in which surface luminances would contrast with each other only if they were on the same plane. If they were on different planes, contrast would be partially *negated* (Benary, 1924; Wertheimer, 1923). However, Dalby, Saillant, and Wooten (1995) presented contradictory results. They suggested that experimental instructions given in previous reports confused the perception of lightness with that of brightness. Knill and Kersten (1991) showed how surface curvature (and subsequent perceptual illumination computations) could also affect lightness perception.

Such interactions of depth and lightness can also be seen in images that contain only monocular depth cues, such as the Benary cross (Figure 3a), the Munker-White display (Figure 3b), and the checkerboard pattern (Figure 3c). In the Benary cross (Benary, 1924; Wertheimer, 1923), the two small gray squares have the same physical reflectance but are seen as having different lightnesses: The top left gray square looks slightly darker than the bottom right gray square.

In the Munker-White assimilation display (Munker, 1970; White, 1979) of Figure 3b, all the gray sections are physically the same but are perceived to have different lightnesses. Owing to a simple simultaneous contrast ar-

gument, the top gray bars should be perceived as darker than the bottom gray bars, since they are adjacent to, and contrast with, mostly white areas. However, the opposite percept is obtained—hence, the label of being an assimilation illusion.

The top three gray bars in the Munker-White display percept may complete amodally behind the larger occluding white bars. The bottom three gray bars can also be perceived as a single gray surface occluded by black bars. It is also possible to see these gray bars as completing to form a transparent surface overlying the alternating black-white stripes. This assimilative lightness effect is elicited by monocular cues: The only depth cues are geometric and contrastive, not stereoscopic.

Agostini and Proffitt (1993) have suggested that if the top gray bars are seen as belonging on a black surface (the white stripes in front) and the bottom gray bars are seen as belonging to a white surface (the black bars in front), the resulting *coplanar* contrast explains the resulting illusion. Todorović (1997) has proposed qualitative rules for computing the perceived lightness in images containing T- and X-junctions—namely, the lightness of a region that has common borders with other regions and whose borders involve T- or X-junctions is predominately a function of the ratio of the luminance of the region and the luminance of colinear regions. For example, in the Munker-White display (Figure 3b), the gray lightness may be a result of contrast with colinear white bars on the bottom and the black background on top. Some research has, however, suggested that lightness differences are greater than what would be predicted purely by gray's contrasting with a white or a black surface (Anderson, 1997; but see Taya, Ehrenstein, & Cavonius, 1995). Possible short- and long-range mechanisms controlling the perception of the Munker-White display have also been discussed (Kingdom & Moulden, 1991; Moulden & Kingdom, 1989; Spehar, Gilchrist, & Arend, 1995).

Unlike the Benary and Munker-White displays, the checkerboard pattern in Figure 3c, which is a variant of the DeValois and DeValois (1988) checkerboard pattern that is due to Ennio Mingolla, is assimilative in nature, in that the gray patch contiguous with the white squares seems lighter than the gray patch connected to the black squares (see also Adelson, 1993). The Todorović (1997) X-junction rule breaks down here since here the percept does not rely on contrast with colinear squares.

Several authors (Anderson, 1997; Moulden & Kingdom, 1989) have endeavored to explain each lightness illusory display individually and qualitatively. This article shows how each illusion may result from the same set of computations performed by the visual cortex to separate the figure from the ground. Zaidi, Spehar, and Shy (1997) said that "given the present state of knowledge about visual neurophysiology, it is not possible to even speculate about possible physiological mechanisms for extracting T-junctions and inhibiting induced contrast" (p. 406). The quantitative computer simulations presented in this article provide concrete physiological underpinnings for

sensitivity to T-junctions and for how the figure-ground relations in visual cortex can affect perceived reflectance in 2-D images, as well as in 3-D images.

Amodal Completion and Recognition

Occlusion cues can be used in object recognition (Nakayama et al., 1989). In the Bregman-Kanizsa display (Figure 2b), when occluded by the black line, the partially occluded Bs are recognizable. However, if the occluder has the same color as the background (Figure 2c), the Bs are much harder to recognize. One mechanistic interpretation of this phenomena is that when the occluder has visible contrast with the background, it pops forward in front of the Bs, allowing the Bs to amodally complete behind the occluder. This completed representation is forwarded to the object recognition system (Grossberg, 1994). Without an occluder that contrasts with the background, no object surface is seen in front of the B, so the Bs cannot complete amodally and are harder to recognize. This work shows how, in addition to modal bound-

ary and surface outputs, amodal boundary and surface representations are also created.

THE FACADE MODEL OF THE VISUAL CORTEX

How Boundary Grouping Converts Multiple Scales Into Multiple Depths

This section reviews the FACADE theory by describing properties of the boundary contour system (BCS) and the feature contour system (FCS) and their interactions. The BCS creates an emergent 3-D boundary segmentation of edges, texture, shading, and stereo information at multiple spatial scales. The FCS compensates for variable illumination conditions and fills-in surface properties of brightness, color, depth, and form among the different spatial scales. Interactions between these complementary boundary and surface processes render them mutually consistent and thereby lead to properties of figure-ground separation. FACADE concepts are described at

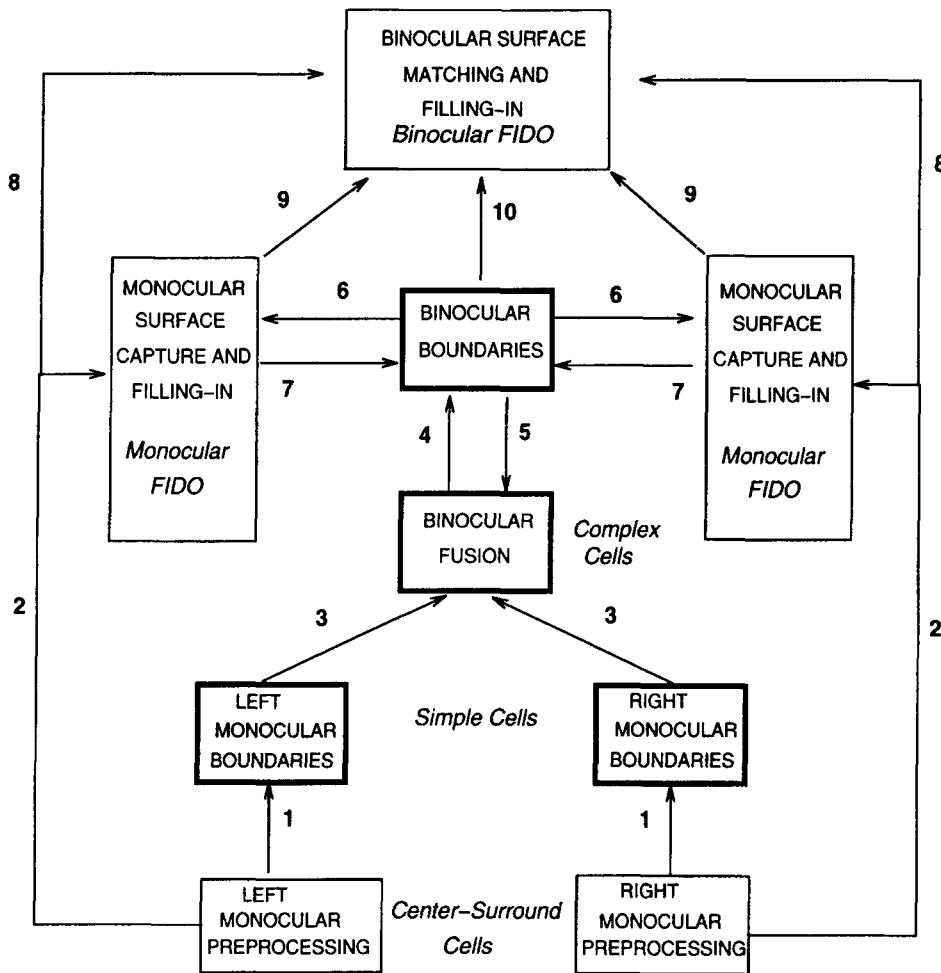


Figure 5. FACADE macrocircuit showing interactions of the boundary contour system and the feature contour system. See the text for details. FIDO, filling-in domain.

length in Grossberg (1994, 1997) and Grossberg and McLoughlin (1997). Here just enough detail is given to afford a self-contained exposition.

The model is mathematically defined in the Appendix, which can be found on line (<http://www.cns.bu.edu/Profiles/Grossberg>). Monocular processing of left-eye and right-eye inputs by the retina and the lateral geniculate nucleus discounts the illuminant and generates parallel signals to simple cells of the BCS via Pathways 1 and to monocular filling-in domains (FIDOs) of the FCS via Pathways 2 in Figure 5. Model simple cells have oriented receptive fields and come in multiple sizes. Simple cell outputs are binocularly combined at complex and complex endstopped (or hypercomplex) cells via Pathways 3. These interactions generate populations of disparity-sensitive cells that realize a size-disparity correlation. In particular, complex cells with larger receptive fields can binocularly fuse a broader range of disparities than can cells with smaller receptive fields (see Smallman & MacLeod, 1994, for a review). Competition across disparity at each position and among cells of a given size scale sharpens complex cell disparity tuning (Fahle & Westheimer, 1995). Spatial competition (endstopping) and orientational competition convert complex cell responses into spatially and orientationally sharper responses at hypercomplex cells.

Hypercomplex cell outputs activate BCS bipole cells via Pathway 4. These cells carry out long-range horizontal grouping and boundary completion. This grouping process collects together the outputs from all hypercomplex cells that are sensitive to a given depth range and inputs them to a shared set of bipole cells. The bipole cells, in turn, send excitatory feedback signals via Pathways 5 back to these hypercomplex cells at the same position and orientation and inhibitory feedback signals to hyper-

complex cells at the same and nearby positions and orientations. This feedback process binds together cells of multiple sizes into a BCS representation, or copy, that is sensitive to a prescribed range of depths. In this way, each BCS copy completes boundaries within a given depth range. Multiple BCS copies are formed, each corresponding to different (but possibly overlapping) depth ranges. This same feedback process also plays a key role in figure-ground separation, as we will now discuss.

T-Junction Sensitivity in the Boundary Contour System

The bipole cells that carry out long-range boundary completion are surrounded by an oriented receptive field with two parts (Figure 6). Each part receives inputs from a range of almost colinear orientations and positions. Bipole cells fire if both parts are simultaneously active, thereby ensuring that the cells do not complete beyond a line end unless there is another line end providing evidence for such a linkage. Cells with similar properties were reported by von der Heydt, Peterhans, and Baumgartner (1984) and are supported by many psychophysical data (e.g., Field, Hayes, & Hess, 1993; Shipley & Kellman, 1992).

Bipole cell outputs excite hypercomplex cells that code similar positions and orientations during the boundary completion process. This feedback spatially and orientationally sharpens the “fuzzy” outputs of the bipole cells. Feedback also inhibits other orientations and positions (Figure 6). The long-range bipole cooperation and the shorter range competition work together to give rise to T-junction sensitivity without the use of T-junction operators: Excitatory bipole feedback strengthens the boundary along the top of the T while inhibiting nearby stem boundary positions, because the top of the T receives more

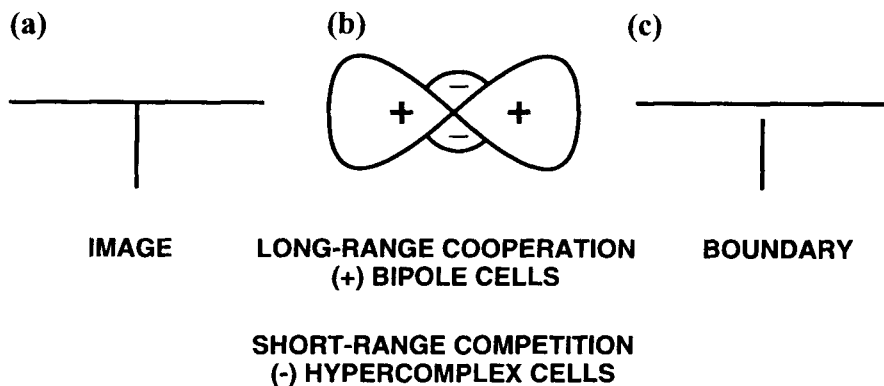


Figure 6. T-junction sensitivity. (a) T-junction in an image. (b) Bipole cells provide long-range cooperation (+), whereas hypercomplex cells provide short-range competition (-). (c) An end-gap in the vertical boundary arises. From “Cortical Dynamics of Three-Dimensional Figure-Ground Perception of Two-Dimensional Pictures,” by S. Grossberg, 1997, *Psychological Review*, 104, p. 641. Copyright 1997 by the American Psychological Association. Reprinted with permission.

support from its bipole cells than the stem receives from its bipole cells. As will be described below, this breaking of the tops from the stems creates gaps in the boundary, termed *end gaps*, which allow color to flow out of this figural region during the surface filling-in process.

Surface Capture and Binocular Surface Matching

The multiple depth-selective BCS copies are used to capture brightness and color signals within depth-selective FCS surface representations. The surface representations that make up the monocular FIDOs receive FCS brightness and color signals from a single eye. A different monocular FIDO preferentially interacts with each binocular BCS copy. In addition, BCS copies that represent nearby depth ranges may send convergent, albeit weaker, signals to each FIDO, thereby allowing a continuous change in perceived depth across a finite set of FIDOs.

Surface capture is achieved by a suitably defined interaction of BCS signals and illuminant-discounted FCS signals at the monocular FIDOs. Pathways 2 topographically input their monocular FCS signals to all the monocular FIDOs. Pathways 6 carry topographic boundary signals from each BCS copy to its FIDO. These boundary signals selectively capture those FCS inputs that are spatially coincident and orientationally aligned with the BCS boundaries. Other FCS inputs are suppressed by the BCS-FCS interaction.

The captured FCS inputs, and only these, can trigger diffusive filling-in of a surface representation on the corresponding FIDOs. Because this filled-in surface is activated by depth-selective BCS boundaries, it inherits the same depth as these boundaries. Not every filling-in event can generate a surface representation. Because activity spreads until it hits a boundary, only surfaces that are surrounded by a connected BCS boundary, or a fine web of such boundaries, are effectively filled in. The diffusion of activity dissipates across the FIDO otherwise.

An analysis of the outputs of the BCS and FCS subsystems has shown that too many boundary and surface fragments are formed as a result of the size-disparity correlation. These extra boundaries and surfaces are pruned by a process whereby the complementary boundary and surface properties interact to achieve a mutually consistent percept. Remarkably, many data about the perception of occluding and occluded objects may be explained as consequences of this pruning operation (see Grossberg, 1994, 1997, and Grossberg & McLoughlin, 1997).

Feedback from the FCS to the BCS is needed to achieve such boundary-surface consistency. A contrast-sensitive process at the monocular FIDOs detects the contours of successfully filled-in surface regions. These contour signals activate FCS-to-BCS feedback signals (Pathways 7), which further excite the BCS boundaries corresponding to their own positions and depths. The boundaries that activated the successfully filled-in surfaces are hereby strengthened. The feedback signals also inhibit redundant boundaries at their own positions and farther depths. This inhibition from near to far is the first example within the

theory of the *asymmetry between near and far*. The *boundary pruning* process spares the closest surface representations that successfully fill in at a given set of positions, while removing redundant copies of the boundaries of occluding objects that would otherwise form at farther depths. When the competition from these redundant occluding boundaries is removed, the boundaries of partially occluded objects can be amodally completed behind them on BCS copies that represent farther depths. Moreover, when the redundant occluding boundaries collapse, the redundant surfaces that they momentarily supported at the monocular FIDOs collapse. Occluding surfaces are hereby seen to lie in front of occluded surfaces.

The surface representations that are generated at the monocular FIDOs are depth selective, but they do not combine brightness and color signals from both eyes. Binocular combination of brightness and color signals takes place at the binocular FIDOs. Here monocular pre-processing signals from both eyes (Pathways 8) are binocularly matched. The surviving matched signals are pruned by inhibitory signals from the monocular FIDOs (Pathways 9). These inhibitory signals eliminate redundant FCS signals. They arise from the contrast-sensitive monocular FIDO outputs. In particular, monocular FIDO inputs to the binocular FIDOs inhibit the FCS signals at their own positions and farther depths. As a result, occluding objects cannot redundantly fill-in surface representations at multiple depths. This *surface pruning* process is the second instance in the theory of the *asymmetry between near and far*.

As in the case of the monocular FIDOs, the FCS signals to the binocular FIDOs can initiate filling-in only where they are spatially coincident and orientationally aligned with BCS boundaries. BCS-to-FCS Pathways 10 carry out depth-selective surface capture of the binocularly matched FCS signals that survive surface pruning. In all, the binocular FIDOs fill in FCS signals that (1) survive within-depth binocular FCS matching and across-depth FCS inhibition, (2) are spatially coincident and orientationally aligned with the BCS boundaries, and (3) are surrounded by a connected boundary or a fine web of such boundaries. At the binocular FIDOs, the BCS adds the boundaries of nearer depths to those that represent farther depths. This instance of the *asymmetry between near and far* is called *boundary enrichment*. These enriched boundaries prevent occluding objects from looking transparent by blocking filling-in of occluded objects behind them. The total filled-in surface representation across all binocular FIDOs represents the visible percept. It is called a FACADE representation because it combines together, or multiplexes, properties of form and color and depth.

The Calculation of Lightness and Depth

The separate surface representations that are formed by the FACADE model at multiple depths must be appropriately combined to give a calculation of relative depth and also of relative lightness. In the case in which

there is activity in only one of the depth-selective FIDO representations at any given position, the final network lightness output is calculated from that active position and depth. However, there are cases, as will be illustrated below, in which two or more FIDO representations at the same positions and very similar depths are simultaneously activated during the perception of an opaque surface. The activities of these FIDO representations are combined as follows to give a lightness and depth percept. First, FIDO activities at a particular depth are normalized. Then the final lightness percept is calculated by summing the normalized FIDO activities at nearby depths. FIDO activities that represent larger depth differences are not summed across depth. Their separate activities represent percepts of transparency.

The model combines FIDO outputs from both *on* cells and *off* cells in different ways to compute lightness and relative depth. For example, if a white object is represented in front of a black background, the white object will be represented in the near *on* FIDO and the black background will be represented in the far *off* FIDO. Thus, to calculate the relative depth of these regions, both *on* and *off* system outputs are used. The Three-Dimensional Modeling and Simulation section shows how these properties help to explain the percepts of depth and lightness in displays such as the Munker-White example.

Recognition of Occluded Objects and Perception of Opaque Occluding Objects

The binocular boundary and monocular FIDO stages in Figure 5 form percepts of the amodally completed boundaries and surfaces of partially occluded objects, as well as of the objects that occlude them. These processing stages are interpreted as occurring in the interstripes and thin stripes of cortical area V2. Modal, or visible, percepts are assumed to occur at the binocular FIDOs, where they represent the unoccluded parts of 3-D surfaces. These stages are interpreted as occurring in cortical area V4.

These distinct representations carry different types of information. The binocular boundaries and monocular FIDOs carry representations that can be used to recognize partially occluded objects. The binocular FIDOs cannot be used to recognize partially occluded objects because boundary enrichment at the binocular FIDOs mixes boundaries of occluding and occluded objects. In so doing, boundary enrichment prevents occluded objects from filling in behind their occluders. Thus, the ability to *recognize* occluded objects and to *see* opaque occluding objects and the unoccluded parts of partially occluded objects is represented at different processing stages.

In order to recognize perceptual properties, whether or not they are modally “seen,” several stages of FACADE processing are proposed to interact reciprocally with model cortical areas that are devoted to object recognition, which play the role of the inferotemporal (IT) cortex (Desimone, 1991; Desimone & Ungerleider, 1989; Mishkin, 1982; Perrett, Mistlin, & Chitty, 1987). Interactions between the object recognition (IT) system and

the binocular FIDOs (V4) are proposed to recognize the unoccluded visible parts of the 3-D surfaces. Interactions between the IT and the binocular boundaries and monocular FIDOs (V2) are proposed to recognize amodally completed occluded objects as well as occluders.

Why Are Both Modal and Amodal Surface Percepts Needed?

Both modal and amodal surface percepts occur in response to images like those in Figures 1–3. Do both types of percepts have functional utility? Grossberg (1997) suggested that their utility may be found in different sorts of recognition and action skills. For example, modal surface percepts may be used to recognize and reach unoccluded objects in the world. They let us know which objects are directly reachable and protect us from trying to reach through an occluder to an object that it occludes. Amodal surface percepts can be used to recognize partially occluded objects. They also provide a recognition signal—which is distinguishable from the modal signals—whereby to plan a reach around an occluder to an object that it occludes.

Evidence for the use of amodal representations in recognition and active touch has been presented by Streri, Spelke, and Rameix (1993) for adults, as well as for 4-month-old infants. Johnson and Aslin (1995) used preferential-looking tasks to provide evidence that 2-month-old infants can perceive occluded objects as being amodally completed. Consistent with the use of amodal surface representations for recognition, Kovács, Vogels, and Orban (1995) have shown that IT neurons that respond preferentially to certain filled object shapes also respond to those shapes when they are occluded by a visible occluder, but not when the occluder was invisible—that is, the same color as the background. Nakamura, Gattass, Desimone, and Ungerleider (1993) have found

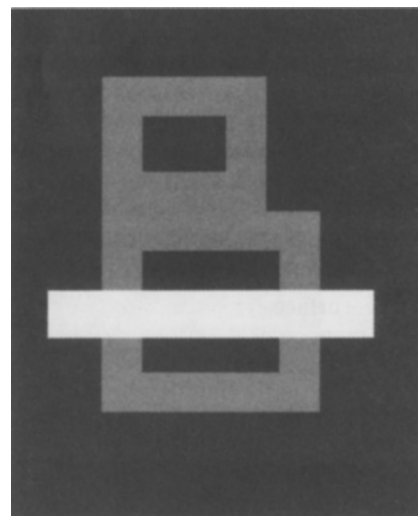


Figure 7. Bregman-Kanizsa simulation input pattern.

evidence for *by-pass* routes from V1 to V4, and from V2 to TEO, consistent with the proposal that amodal surface representations created by early stages of visual processing can be routed directly to object recognition centers and also to higher visual areas for further processing to create modal representations. Sekuler and Palmer (1992) have also shown that amodal representations develop over a longer time than modal percepts. This is consistent with the FACADE model's surface-to-boundary feedback and bipole completion, which require a small number of feedforward and feedback iterations to complete the modal and then the amodal surface percepts.

THREE-DIMENSIONAL MODELING AND SIMULATION

This section presents quantitative simulations of figure-ground separation and amodal completion in response to the Bregman-Kanizsa (Figure 7) and Kanizsa stratification displays, as well as simulations of the Benary, Munker-White, and checkerboard lightness illusions. In all FCS simulations of the monocular and binocular FIDOs, active cells are represented by using an activity-based scale with white (most active) or various shades of gray (less active). A lack of activity of FCS cells is rep-

resented by black-colored regions. Lighter areas of the percept are represented using more active *on* cells; however, darker image regions are not represented by the *on* cells. Darker regions are represented by more active *off* cells, whose activity is represented by nonblack values. Image lightness is calculated by measuring the double-opponent difference between the filled-in activities of *on* and *off* cells at each position. Owing to how the cell membrane equations respond to *on* and *off* inputs, all *on* and *off* output surface representations are normalized by dividing opponent activities (i.e., *on*-minus-*off*, *off*-minus-*on*) by the sum of these activities (*on*-plus-*off*). When near and far outputs are combined, they therefore have values between 0 and 1. See the Appendix for details.

Figure-Ground Separation and Amodal Completion in the Bregman-Kanizsa Percept

In this first simulation, the outputs of most stages of the FACADE model will be displayed to clarify how the model works. In other simulations, only the most important boundary and surface representations will be shown. The image is fed into the left and right monocular preprocessing stages. Figures 8a and 8b show the outputs of the *on* and *off* cells at the monocular preprocessing stages. Since left and right stream responses are identi-

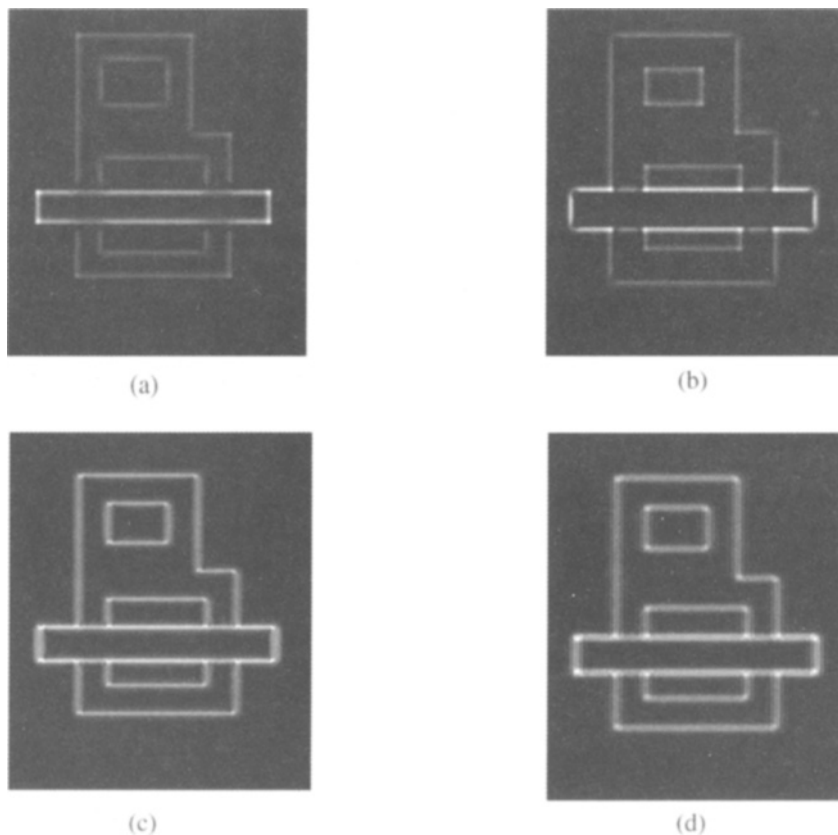


Figure 8. Output of the (a) *on* cell and (b) *off* cell monocular preprocessing stages. Output of the complex cell preprocessing stage: (c) large scale, disparity D0 and (d) disparity D1.

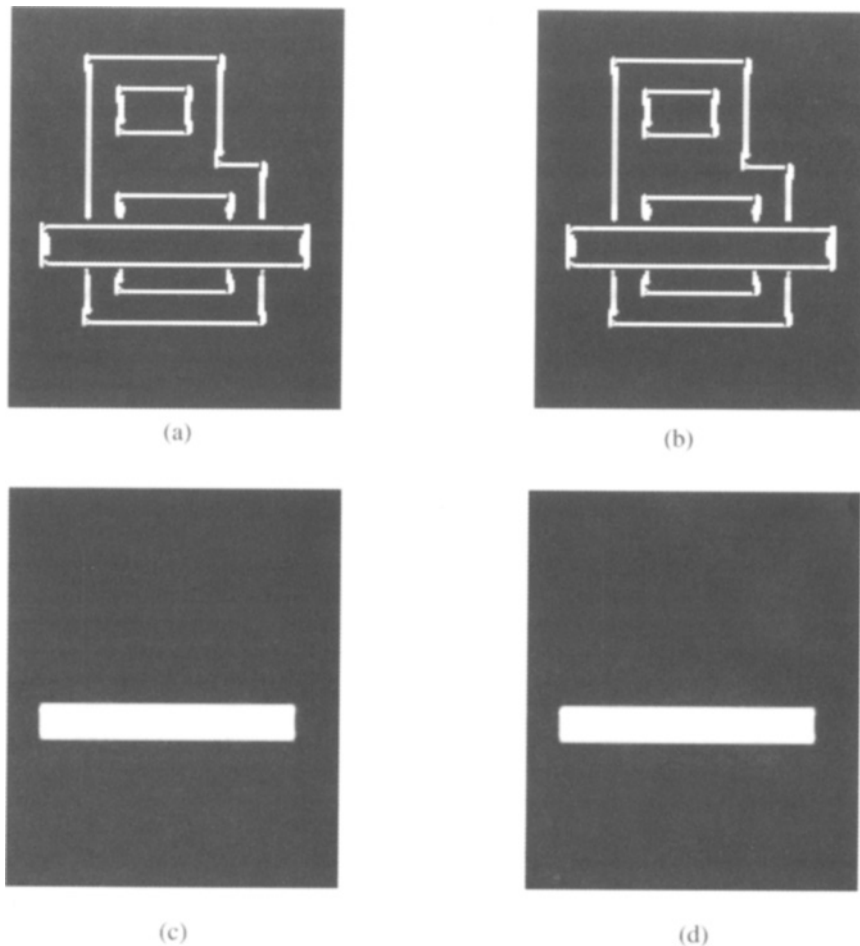


Figure 9. Binocular boundaries for monocular filling-in: (a) near depth and (b) far depth. Output of monocular FIDOs before boundary pruning occurs: (c) near depth and (d) far depth.

cal, Figures 8a and 8b show the *on* and *off* cell responses for only one of those streams. Simple cell processing is not shown. Figures 8c and 8d show complex cell stage outputs. Inhibition occurs across disparities within a scale and within a disparity across scales (from large to small scales) at the complex cells. As a result, the large-scale representation is active at zero disparity ($D0$), but the small-scale representation is active at a slightly farther disparity ($D1$).

Figures 9a and 9b show the output of the hypercomplex cells after spatial and orientational competition and subsequent bipole cell feedback act. Bipole feedback causes the breaking off of the tops from the stems of the T-junctions, since the tops of the T receive more support from the bipole cells. These binocular boundaries are used as filling-in barriers within monocular and binocular FIDOs. The end gaps in the boundary allow color to flow out of the corresponding regions and dissipate across space.

Figures 9c and 9d show the outputs of the monocular FIDOs before they activate surface-to-boundary feedback. Only the occluder regions, whose boundaries are fully closed, trap color and fill in. The end gaps in the B boundaries allow color to flow out of the partially occluded B region. Thus, in both the near-depth and the far-depth pools of the monocular FIDOs, the white occluder fills in, while the gray color flows out of the occluded regions, owing to the gaps in the boundary. Next, the near-depth monocular FIDOs send inhibitory signals to the BCS boundaries at farther depths and inhibit the occluder boundaries there. This allows far-depth bipole cells to amodally complete the occluded B boundaries, thereby removing the gaps that allowed color to flow out (Figure 10b). The near-depth boundaries are unaffected (Figure 10a). When the gaps in the B boundaries are closed, the entire B, including its occluded region, is filled in at the far depth pool (Figure 10d), thereby providing an amodal surface percept of a fully filled-in B at the monocular FIDOs. The filled-in occlud-

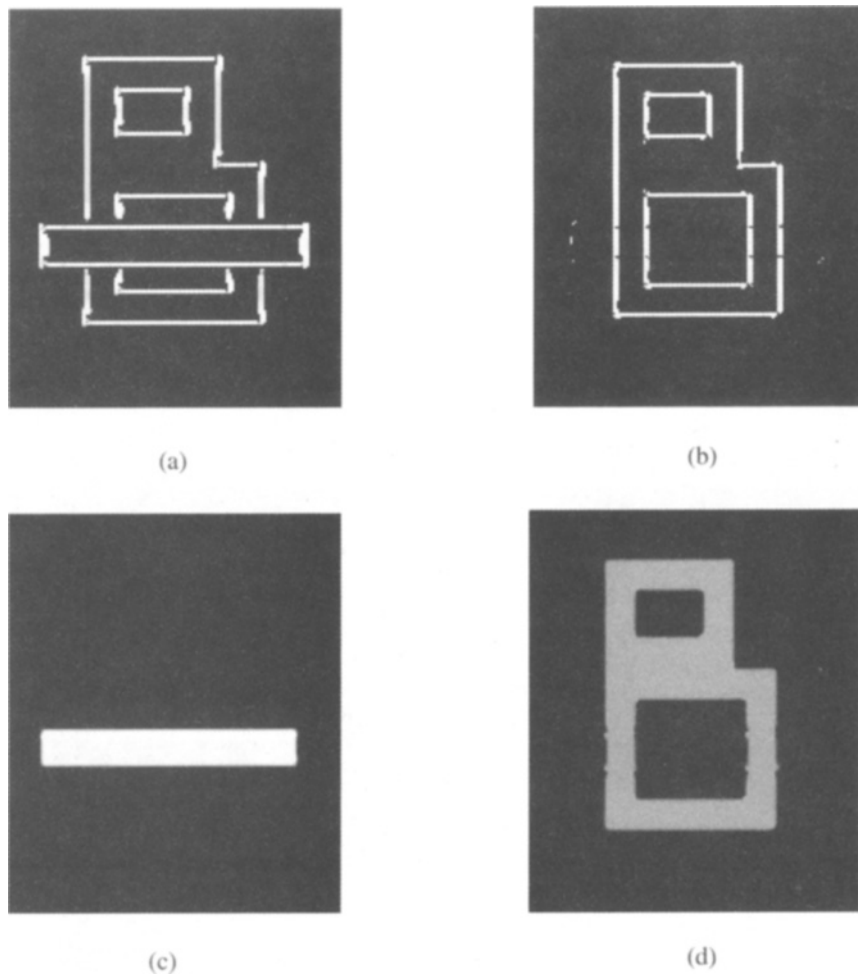


Figure 10. Amodal boundary and surface representations. Binocular boundaries after boundary pruning occurs: (a) near depth and (b) far depth. Amodal surface representations at the monocular filling-in domains: (c) near depth and (d) far depth.

ing white bar remains unchanged at the near-depth pool (Figure 10c). The amodal boundary (Figure 10b) and surface representations (Figure 10d) of the completed B are both used to recognize the B shape.

Modal percepts are represented at the binocular FCS. As was discussed earlier, two asymmetries between near and far are computed at the binocular FIDOS. The first asymmetry inhibits redundant filling-in signals. The near-depth monocular FIDO output (white horizontal bar in Figure 10c) hereby inhibits the corresponding filling-in signals at the far depth. As a result, the occluder's filling-in signal is removed from the far depth of the binocular FIDO (Figure 11b), and the occluding object is not seen at both the near- and the far-depth pools.

The second asymmetry is the addition of near boundaries to the far-boundary representation, as in Figure 11d. The near-boundary representation is the same as that in the monocular FIDO (Figure 11c). By combining these enriched binocular boundaries and pruned surface

inducers at the binocular FIDOS, the occluder fills in at the near depth (Figure 11e), but at the far depth, the gray B surface is filled in only within the regions that are unoccluded (Figure 11f). The resulting surface representations match the stratified percept of an occluder at a nearer depth than the object that it occludes.

Kanizsa Stratification Percept

Consider the Kanizsa stratification display in Figure 1. The thin vertical black lines create T-junctions with the cross. The stems of the T boundaries are broken by the bipole feedback, thus separating the thin outline square from the cross (see Figure 12a). At the top and bottom arms of the cross, vertical bipole cells link the sections of the cross arms together, thereby creating a T-junction with the sections of the square. The vertical bipole cells of the cross win out over the horizontal bipole cells of the squares. This happens because the cross is wider than the square. Thus, vertical bipole cells have more support

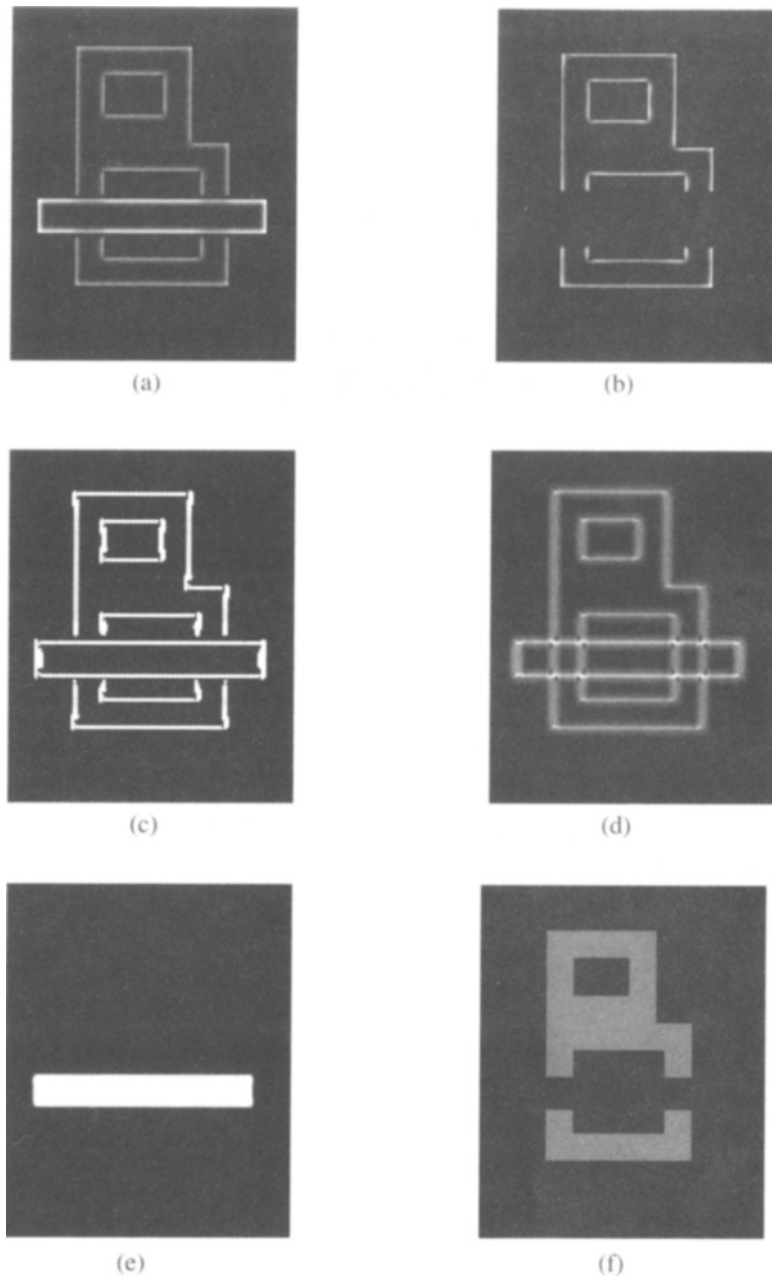


Figure 11. Enriched boundary and modal surface representations. Binocular filling-in domain (FIDO) filling-in signals at (a) near depth and (b) far depth. Enriched boundaries at the (c) near depth and (d) far depth. Binocular FIDO activity consisting of two modal surfaces at two different depths: (e) near depth and (f) far depth.

from their receptive fields than do the horizontal bipole cells at the cross-square intersection. The boundaries of the square are hereby inhibited, thereby creating end gaps. As a result, the cross arms pop in front and the square is seen behind the cross (Figures 12b and 12c).

The bistability of the stratification percept may be explained in the same way that the bistability of the Weisstein effect (Brown & Weisstein, 1988) was explained in

Grossberg (1994). This explanation used the habituated transmitters that occur in the Pathways 3 between complex cells and hypercomplex cells (Figure 5). Transmitter habituation helps to adapt active pathways and, thereby, to reset boundary groupings when their inputs shut off (Grossberg, 1997). This transmitter mechanism has been used to simulate psychophysical data about visual persistence, aftereffects, residual traces, and metacontrast

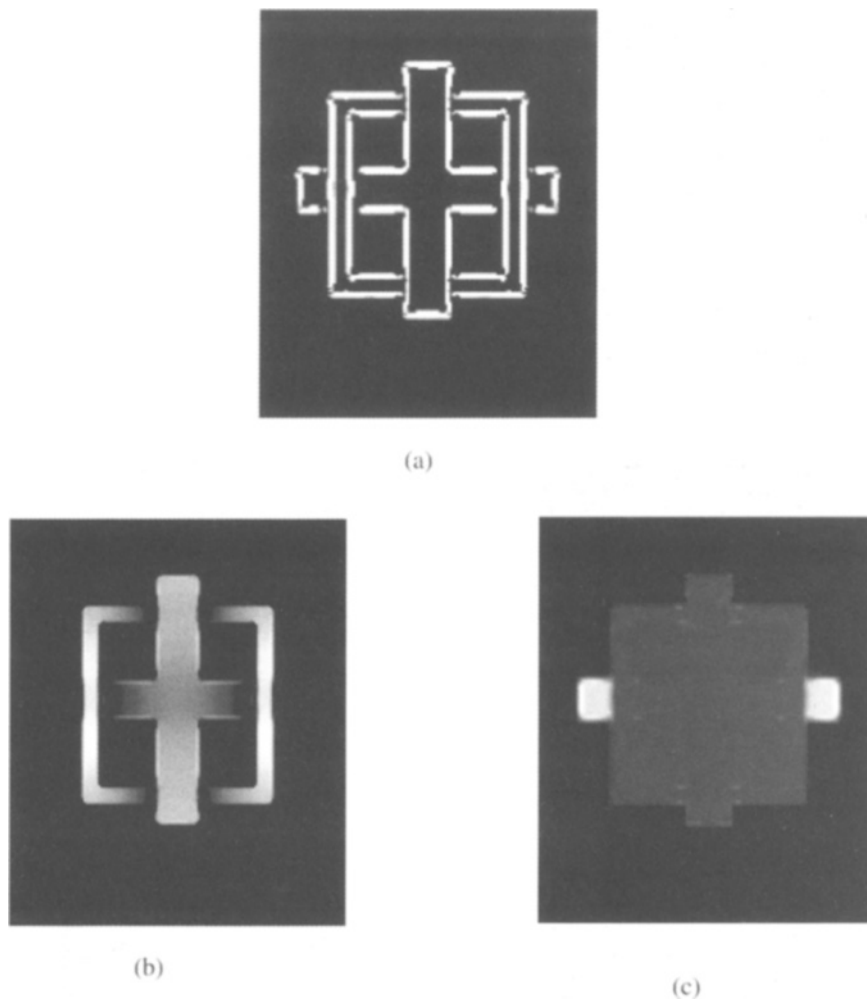


Figure 12. (a) Near-depth boundaries in response to the Kanizsa stratification image. Binocular filling-in domain activity at the (b) near depth and (c) far depth.

masking (Francis, 1997; Francis & Grossberg, 1996a, 1996b; Francis, Grossberg, & Mingolla, 1994), developmental data about the self-organization of opponent simple cells, complex cells, and orientation and ocular dominance columns within cortical area V1 (Grunewald & Grossberg, 1998; Olson & Grossberg, 1998), and neurophysiological data about area V1 cells (Abbott, Varela, Sen, & Nelson, 1997). The bistability of the stratification percept can hereby be traced to more basic functional requirements of the visual cortex.

Lightness Illusions: Benary Cross

Quantitative explanations of how the Benary, Munker-White, and checkerboard lightness illusions arise are now presented. FACADE theory suggests that these illusions are by-products of how the visual system solves the figure-ground problem. In particular, in the Benary and Munker-White displays, the contrastive illusion is explained by analyzing how the visual system interprets

whether the gray patch is solely on a white or a black background, thereby discounting the effect of other spatially congruent regions (cf. the coplanar ratio hypothesis of Gilchrist, 1977). In the checkerboard illusion, we show that as a result of how X-junction boundaries are grouped, extra end gaps are created, which allow more color flow that results in an overall assimilation effect.

The Benary cross (Figure 3a) leads to the near-depth boundary representation processing in Figure 13a. Here, the boundaries of the T-junction stems where the gray squares about the cross are broken to form end gaps. These boundaries allow color to fill in the entire cross at the near-depth monocular FIDO. Boundary pruning signals occur from this near-depth surface representation to the far-boundary representation via Pathways 7 in Figure 5. The cross boundaries are hereby inhibited at the far depth, as in Figure 13b. As a result, the cross boundaries, but not the T-junction stem boundaries, are removed at the far depth. The end gaps are no longer present here, but

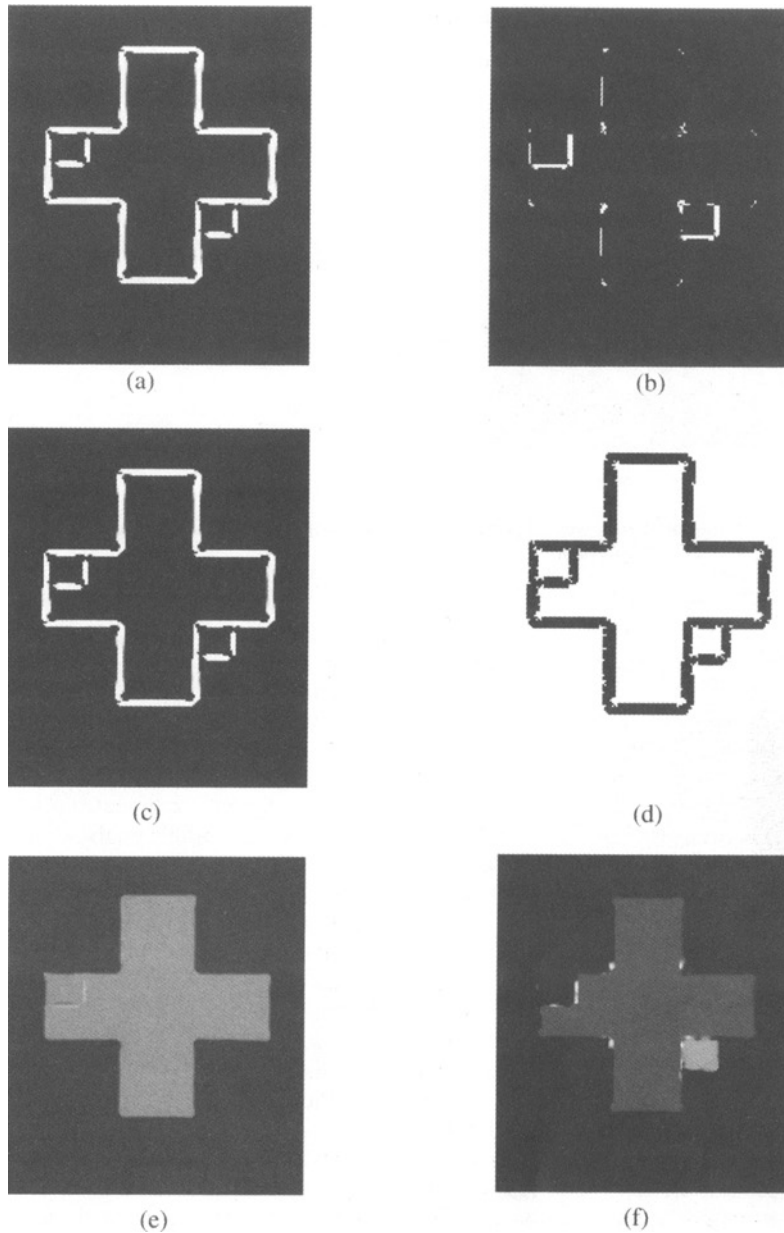


Figure 13. Benary cross binocular boundaries to monocular filling-in domains (FIDOs) after boundary pruning: (a) near depth and (b) far depth. Enriched boundaries to binocular FIDO: binocular FCS boundaries to binocular FIDO: (c) near depth and (d) far depth. Binocular FIDO output: (e) near depth and (f) far depth.

there are no connected boundary regions to trap color during filling-in. It is only at the binocular FIDOs, where near boundaries are added to the far boundaries by the boundary enrichment process (Pathways 10 in Figure 5), that fully closed boundaries are created at the far-depth plane. The binocular FIDO boundaries are shown in Figures 13c and 13d. At the near depth (Figure 13c), the end gaps caused by the breaking of T-junctions remain. At the far depth (Figure 13d), all end gaps are removed. Also in the binocular FIDOs, surface pruning inhibits the cross

filling-in signals at the far depth (Pathways 9 in Figure 5). Only the filling-in signals resulting from the gray squares remain. The modal outputs of the binocular FCS that are due to these filling-in generators within the boundaries of Figures 13c and 13d are shown in Figures 13e and 13f at the near- and the far-depth pools, respectively. Figure 13e shows how the end gaps in Figure 13c allow filling-in to spread through the entire cross. Figure 13e also shows how the bottom gray square fills in with black through the end gaps that abut the black background.

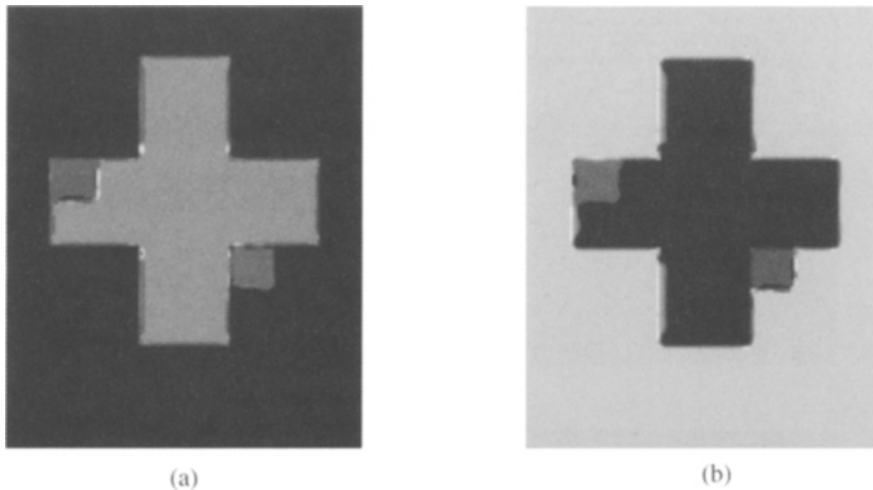


Figure 14. Benary cross combined near and far binocular filling-in domain outputs: (a) *on-minus-off* and (b) *off-minus-on*.

Figure 13f shows that the upper gray square fills in darker gray because some of its gray filling-in generators (at the black–gray border with the cross) are inhibited, owing to surface pruning. The remaining *on* cell generators (at the gray–white border) are outside the gray square boundaries, but inside the cross boundaries, and thus fill in the cross. The bottom gray square fills in lighter gray because its *on* filling-in signals at the gray–black border are not inhibited by the cross and are within the boundaries of the square.

Most people report a Benary cross percept of relative depth that is not nearly as compelling as that for the Bregman–Kanizsa display. They see two gray patches, one of which seems to be internal to the cross, the other external. We suggest that this ambiguity regarding depth occurs because the near and far FIDOs have some regions that are filled in at both the near- and the far-depth pools. To see this, we combine the near- and the far-depth pool representations to get the full modal percept. Figure 14a shows the filled-in *on-minus-off* representation, and Figure 14b the filled-in *off-minus-on* representation.

Owing to the coarseness of the image gray scale, the lightness illusion magnitude is not entirely clear from the output image. The final equilibrium values of the filled-in *on-minus-off* representation for each colored region are as follows: The magnitude of the “white” in the cross is 0.8; the gray on the white cross is 0.45; and the gray on the black background is 0.5. Consistent with the percept, the magnitude of the simulated illusion is quite small (around a 10% difference). The *off-minus-on* representation has similar values; however, high *off* magnitudes correspond to darker regions, and low *off* magnitudes correspond to lighter regions.

Lightness Illusions: Munker–White Assimilation

The Munker–White illusion in Figure 3b is considerably stronger than the Benary illusion. This may be be-

cause, unlike the case of the Benary cross, amodal completion of the gray patches occurs in this display. Figures 15a and 15b show the results of boundary formation after boundary pruning acts. At the near-depth boundaries (Figure 15a), the T-junction stems are entirely broken, thereby allowing white color signals to fill in all the bars. When pruning signals from the near-depth filled-in bars inhibit the far-depth horizontal bar boundaries, the vertical gray–white and gray–black boundaries can complete amodally behind the horizontal bars (Figure 15b).

In the monocular FIDOs, all seven horizontal bars fill in successfully at the near depth, but filling-in dissipates at the far depth, owing to the lack of connected boundary regions. Figures 15c–15f show the boundary and filling-in signals to the binocular FIDOs. At the near depth (Figure 15c), the T-junctions remain broken and allow color to flow along the length of the bars. At the far depth (Figure 15d), the addition of the near boundaries to the far ones creates connected boundary regions. The direction of the lightness illusion depends on the surface pruning process whereby far monocular FCS inputs (Pathways 8 in Figure 5) are inhibited by near monocular FIDO inputs (Pathways 9 in Figure 5). The near filled-in horizontal bars hereby inhibit their filling-in signals at the far depth. This leaves only the filling-in signals at the vertical gray–white or gray–black contours. Figure 15f shows these *on* filling-in signals at the far depth. Note that the *on* signals at the top three gray patches are larger than those at the bottom. The alignment of these FCS signals is also important. The top three pairs of FCS signals in Figure 15f are contained within the gray patch boundaries in Figure 15d and thus fill in these patches. The bottom three white–gray FCS signals, however, are contained within the boundaries of the white patches that about the gray patches and, therefore, do not contribute to the lightness of the bottom gray patches.

The simulated near-depth binocular FIDO activity profile is shown in Figure 16a. It consists of seven hori-

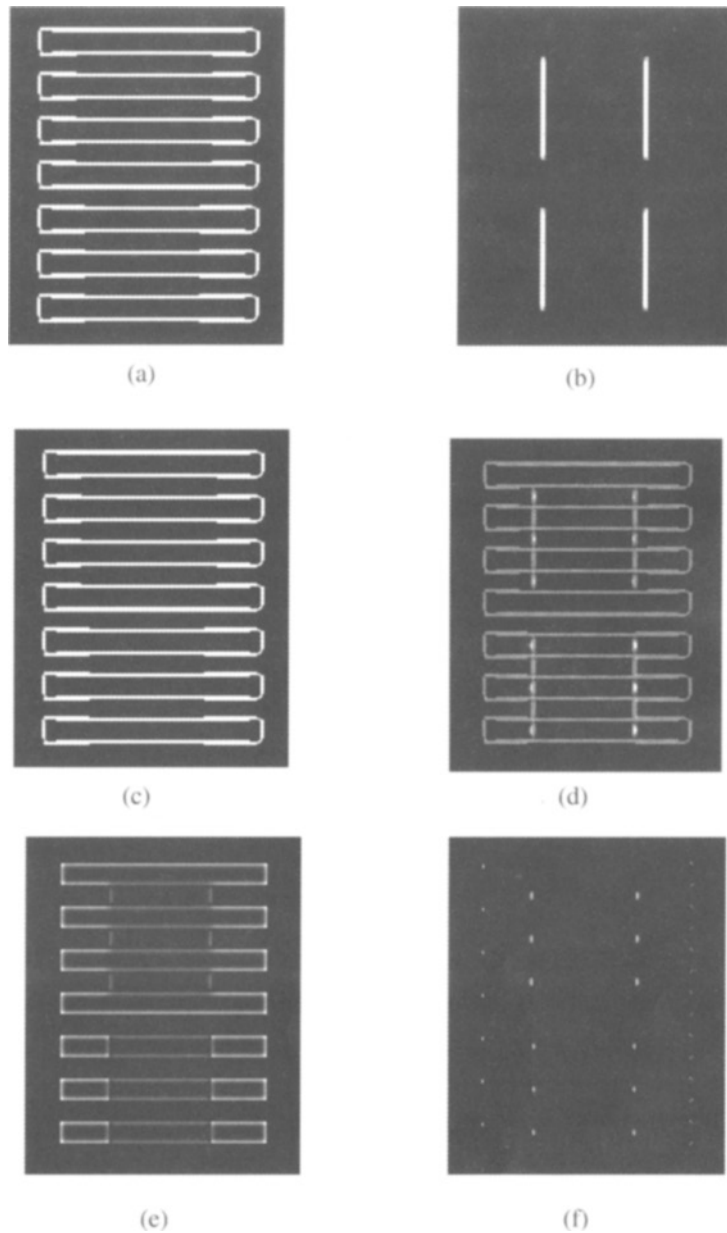


Figure 15. Munker-White binocular boundaries to monocular filling-in domains (FIDOs) after boundary pruning: (a) near depth and (b) far depth. Enriched binocular FIDO boundaries: (c) near depth and (d) far depth. Binocular FIDO filling-in signals: (e) near depth and (f) far depth.

zonal “occluding” bars. Figure 16b shows the corresponding far-depth binocular FIDO activity. Here, the top three gray patches fill in strongly, as do the white sections of the bottom three bars. When near and far representations are added together, the final simulated percept in Figure 16c is found. Then, the average activity of the filled-in gray bars on top is 0.6, whereas the gray bars on the bottom have an average filled-in value of 0.4, as in the Munker-White percept. Figures 16d and 16e show

the near and far filled-in *off* representations, and Figure 16f shows their combination. The *off* representation shows how the bottom gray sections can be perceived as darker. These model simulations suggest that the Munker-White assimilation is a misnomer, since the processes that give rise to the gray lightness differences in the simulations are primarily figure-ground and contrastive in nature.

The model clarifies how the long horizontal bars are perceived as being in front. However, for many observers,

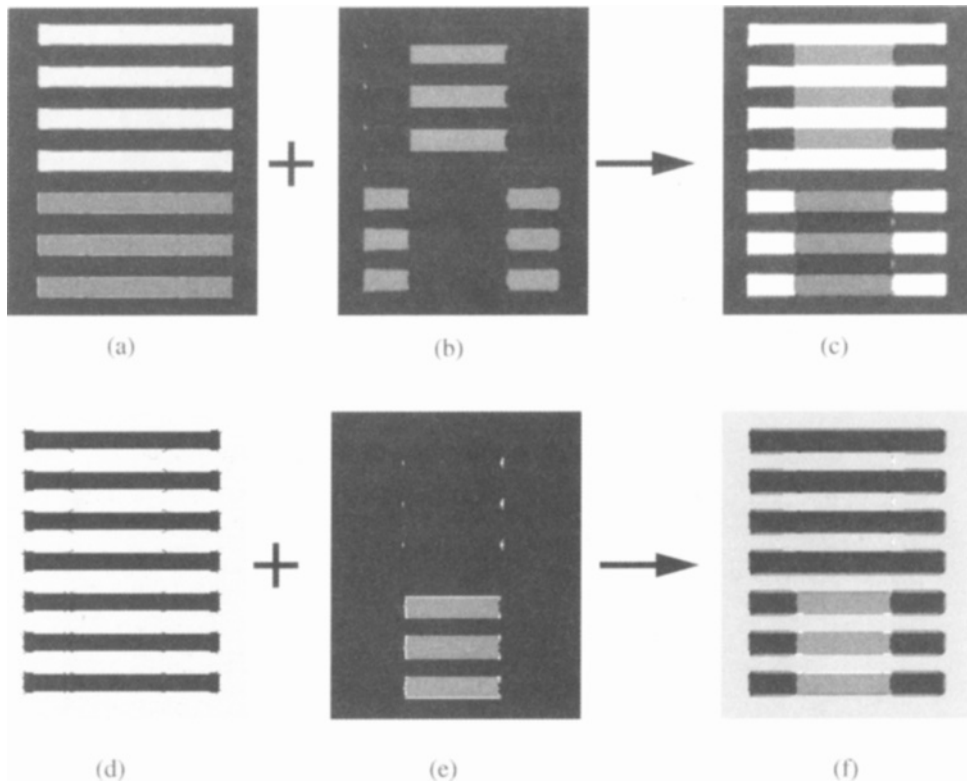


Figure 16. Munker–White binocular filling-in domain (FIDO) output of the *on* cells: (a) near depth, (b) far depth, and (c) combination of near and far depths. Binocular FIDO output of the *off* cells: (d) near depth, (e) far depth, and (f) combination of near and far depths.

the percept is bistable. One can see the gray patches at the top as being behind white occluders, but the gray patches on the bottom can also be seen as a transparent gray surface overlying the white bars. Such bistable representations can reorganize the output of FACADE, much as in response to the Kanizsa stratification image (Figure 1), to allow near and far representations to interchange and reorganize, using habituated transmitters as in the theory's explanation of the Weisstein effect (Grossberg, 1994).

Alternative figure–ground organization percepts of the Munker–White display can also be facilitated by attention shifts. In this way, one can more easily perceive the gray targets on the bottom three bars as a transparent gray filter overlying white bars. This percept is reminiscent of how the disk-and-checkerboard display of Kanizsa (1979) is perceived (see Figure 17a). As was noted by Kanizsa (1979), amodal completion behind the disks does not lead to the more “likely” perception of squares that the checkerboard would suggest. Instead, one is aware of, but does not see, a white cross and a black cross that are partially occluded by the gray disks. Similarly, in the bottom section of the Munker–White display (Figure 17b), when a gray transparent surface is seen to overlie the three horizontal bars, we suggest that subjects are amodally

aware of the continuation of the white surface color beneath the gray overlay. In the model, this amodal surface representation resides in the monocular FIDOs (Figure 5), whereas the visible-surface representations are computed in the binocular FIDOs. This percept illustrates the model hypothesis in the FACADE Model of the Visual Cortex section that distinct representations subserve modal and amodal perception.

We simulated such an attentional shift to the bottom area of the Munker–White display (Figure 17b) by strengthening the vertical white–gray contours. (See Grossberg, 1999, for an explanation of how attention can amplify a boundary grouping.) The T-junction stems that are defined by these vertical contours are now stronger than the T-junction tops and thus, as in Figure 6, cause breaks in the horizontal contours (see Figure 17c) and not the stems. Figures 17d and 17e show how the boundaries then develop over time. In particular, vertical boundaries are now completed by bipole grouping over the broken horizontal boundaries. Figure 18a shows the near-depth monocular FIDO output derived by using the boundaries from Figure 17e. Surface-to-boundary feedback then inhibits, or prunes, the same boundaries at the far depth and allows the horizontal bar boundaries to re-form. Subsequent filling-in of this far-depth monocular FIDO is

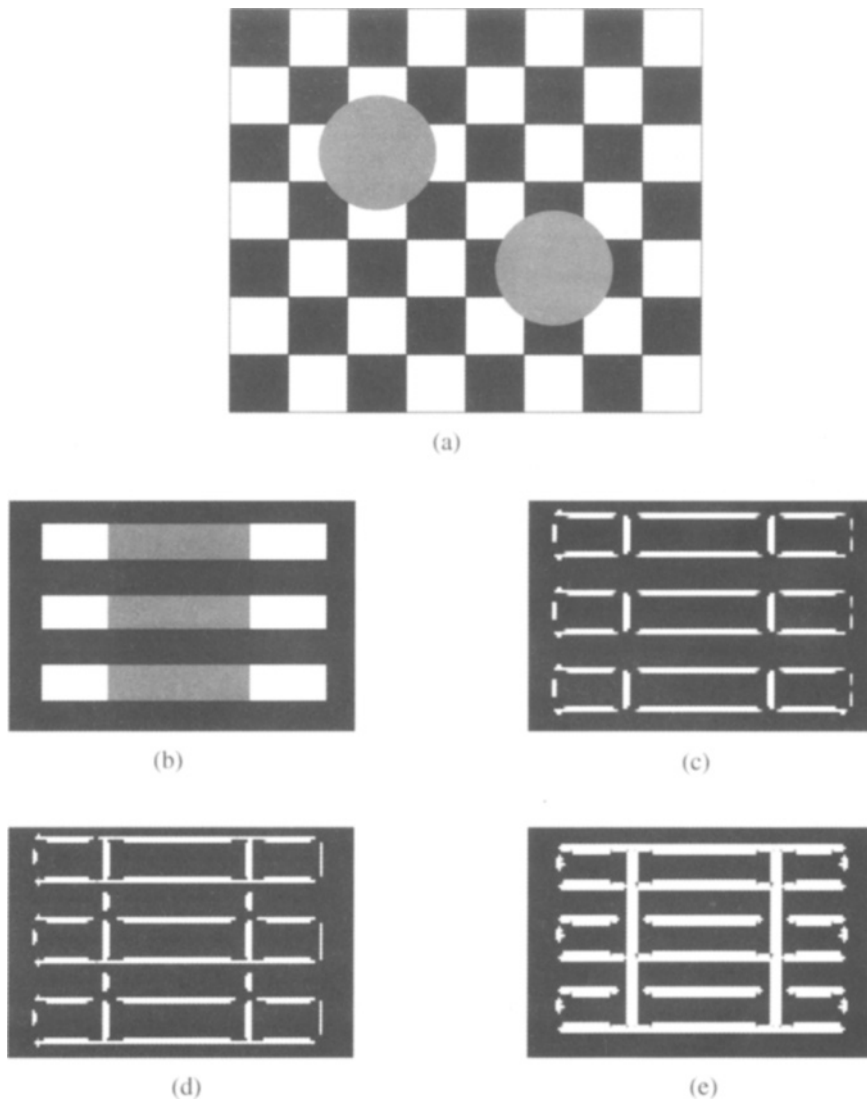


Figure 17. (a) Kanizsa example of amodal completion. (b) Bottom section of Munker-White display. (c) Boundary processing after attentional strengthening of vertical contours (Iteration 1), (d) boundary processing (iteration 2), and (e) boundary processing (equilibrium model at Iteration 3). Panel a is from *Organization of Vision: Essays on Gestalt Perception* (p. 97), by G. Kanizsa, 1979, New York: Praeger. Copyright 1979 by Greenwood Publishing Group. Reprinted with permission.



Figure 18. Munker-White display monocular filling-in domain output: (a) near depth and (b) far depth.

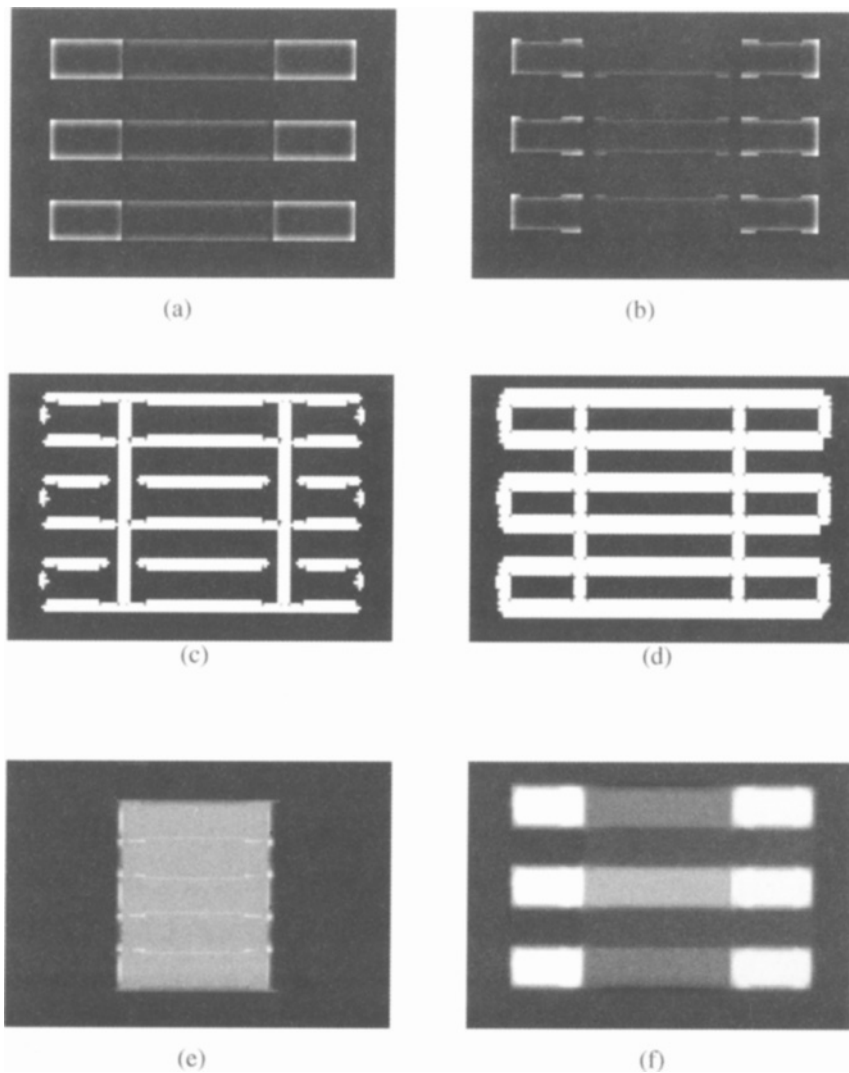


Figure 19. Munker-White binocular filling-in domain (FIDO) filling-in signals: (a) near depth and (b) far depth. Binocular FIDO boundaries: (c) near depth and (d) far depth. Filled-in binocular FIDO activity of *on* cells: (e) near depth and (f) far depth.

shown in Figure 18b. In all, two amodal surface representations are generated: a near representation that fills in a vertical band of gray color and a far representation of three light horizontal bars.

Near and far binocular FIDO filling-in signals are shown in Figures 19a and 19b after surface pruning occurs. Near and far binocular FIDO boundaries following boundary enrichment are shown in Figures 19c and 19d. Figures 19e and 19f show the near and far modal surface representations at the binocular FIDOs. The near binocular FIDO fills in a transparent gray surface (Figure 19e). In addition, the far FIDO filling-in signals can fill in the three gray regions only with gray, because the boundaries in Figure 19d prevent the white filling-in signal from entering. We suggest that in the percept of the gray transparent overlay in front of the bars, the disparity dif-

ference between near and far FIDO representations is greater than that in the previous percept of opaque surfaces. Because of this increased depth difference, near and far FIDO representations are not added together to achieve the final modal surface lightnesses but are perceived individually. In summary, although the model *sees* a gray region that is occluded by a gray transparent surface, as in Figures 19e and 19f, it *knows* that the horizontal bars are lighter, as in Figure 18b.

Lightness Illusions: Checkerboard

Agostini and Profitt (1993) proposed that the visual system computes the lightness of the gray patches in the Benary and Munker-White displays on the basis of *coplanarity* or *belongingness*. This view, however, has trouble explaining why the checkerboard illusion (Figure 3c)

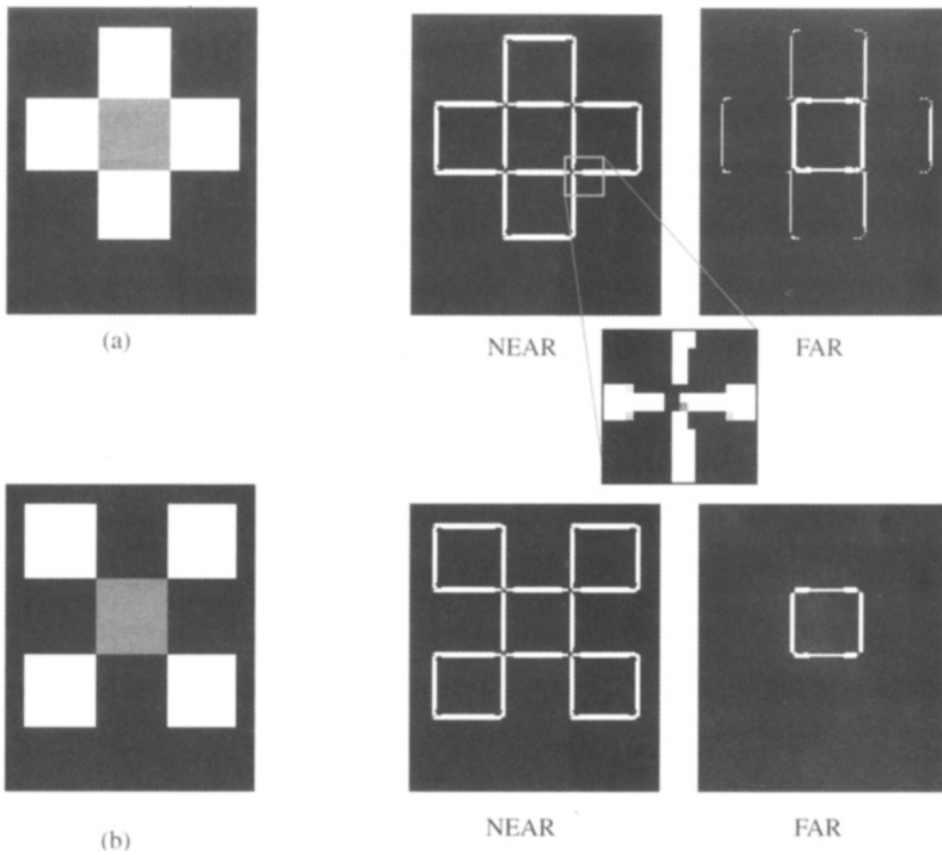


Figure 20. Checkerboard near and far boundaries to monocular filling-in domains after boundary pruning: (a) the cross display and (b) the X display.

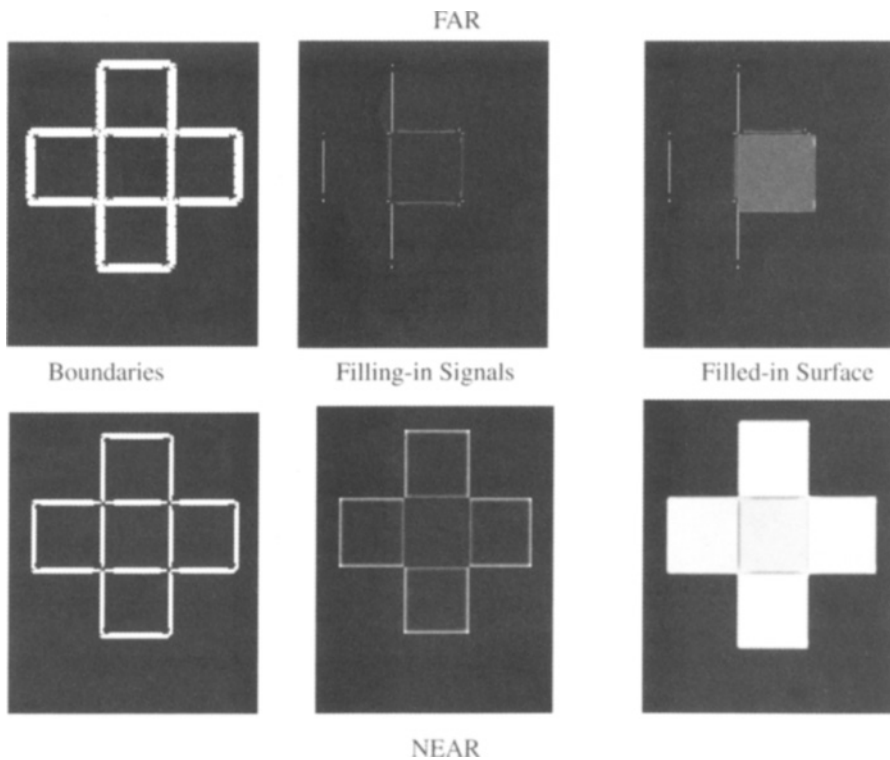


Figure 21. Cross display near and far enriched boundaries, filling-in signals, and filled-in binocular filling-in domain surface values.

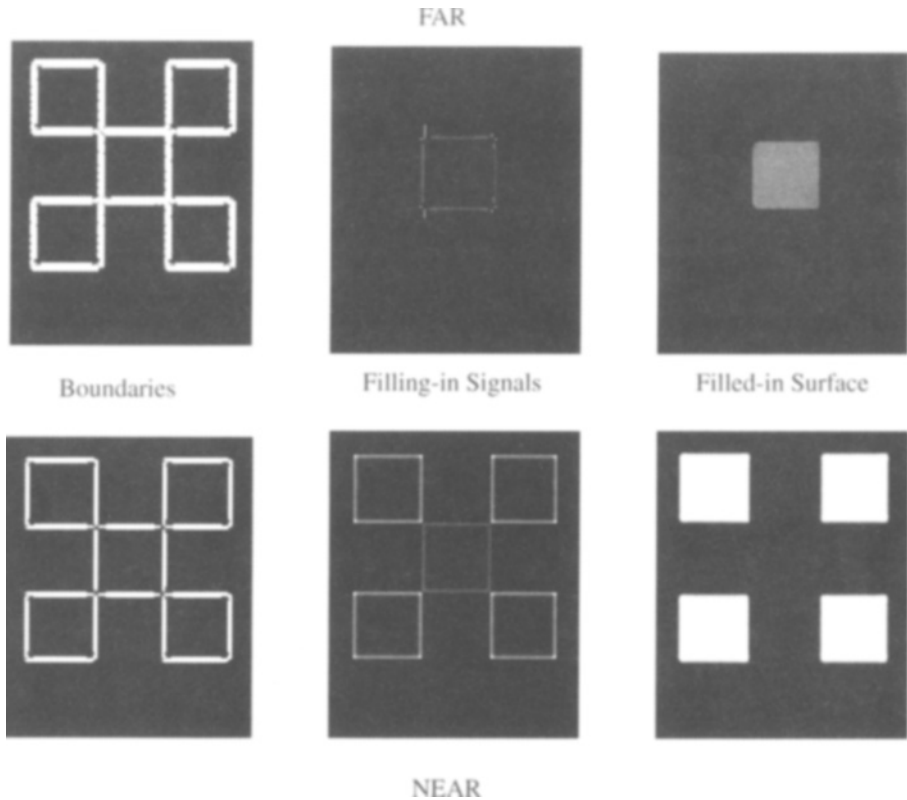


Figure 22. X display near and far enriched boundaries, filling-in signals, and filled-in binocular surface values.

is assimilative: The gray patch that belongs to the white cross (in the upper left-hand corner) is lighter than the gray patch that belongs to the black cross (in the lower right-hand corner).

We propose that the contrastive effect—which is rate limiting in the Munker–White percept—is outweighed

by a process that fills in more white (or black) at cells that code a disparity that lies just behind the gray region. This filling-in occurs at the binocular FCS, the source of the modal percept, and thus when seen in conjunction with the grays, it makes the grays on the white background seem lighter than the grays on the black background. This

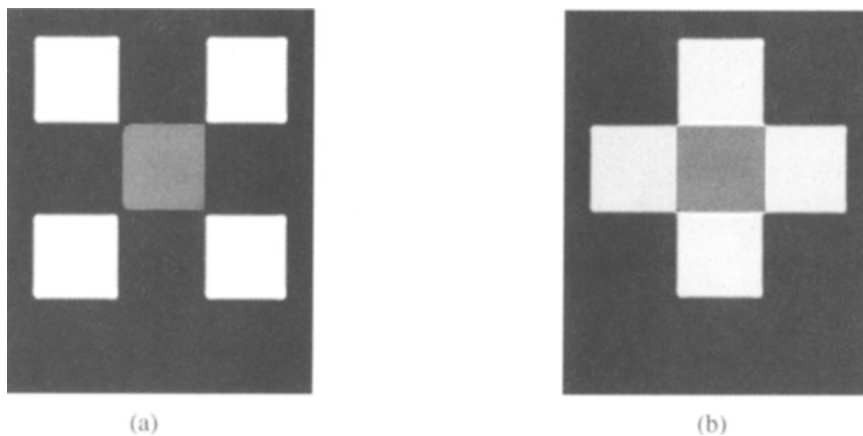


Figure 23. Final combined (near+far) filled-in binocular filling-in domain surface activities for (a) X and (b) cross sections of the checkerboard display.

extra black or white filling-in behind the gray patches results from the presence of X-junctions in the image, which create end gaps that allow more color to flow than in the Benary or Munker-White illusion. The following figures illustrate these processes.

We simulated the checkerboard display in two parts in order to compensate for the relative sparseness of model cells, relative to cells in the visual cortex, and to make the simulation more tractable. In all other respects, we used the same network parameters as those in the other simulations. Figure 20 shows the boundary signals for these two subsections of the checkerboard display. Display Figure 20a is called the cross display, and Figure 20b the X display. In both displays, boundaries are broken at X-junctions. In the cross display, the end gaps (Figure 20, near) allow white color filling-in signals from the four surrounding squares to flow into the central region to create a fully filled-in white cross. When the cross boundary pruning signals are fed back to the far-depth cross boundaries, these boundaries are inhibited, and the central square boundaries remain as a fully connected region (Figure 20a, far). In the X display, the X-junction boundary breaks (Figure 20b, near) and allows the gray signal to flow out and dissipate into the surround, while the four white squares fill in. When boundary pruning signals from the four square near-depth surfaces are fed back to the far-depth square boundaries, only the boundaries around the central square survive (Figure 20b, far). The amodal surface percepts that are created by the boundaries in Figure 20 are as follows: For the cross display, a white cross surface is present at the near depth, and a gray square patch at the far depth; for the X display, four white square surfaces are present at the near depth, and a gray square patch at the far depth.

Figure 21 shows the binocular boundaries, filling-in signals, and filled-in binocular FIDO values for the cross. The filled-in values in Figure 21 show that much of the white filling-in signal spreads into the central square at the near depth. Figure 22 shows the same quantities for the X display. The near-depth boundary representation is the same as that in Figure 20b. At the far depth, boundary enrichment re-forms the X-junctions by the addition of near boundaries. The binocular FIDO receives the same near-depth FCS signals as the monocular FIDO. Surface pruning removes the white cross FCS signals from the far depth, leaving only the gray square FCS signals. The filled-in surfaces show four white surfaces at the near depth and the gray square at the far depth. Figure 23a shows that adding the near and far equilibrium values of the X display in Figure 22 adds a gray square (far) to black (near). Figure 23b shows that adding the near and far equilibrium values of the cross display adds a gray square (far) to the white filling-in of the central cross patch (near). The gray patch in Figure 23a has activity of 0.45, whereas the gray patch in Figure 23b has activity of 0.55, thereby demonstrating the assimilation that is seen in the checkerboard percept.

CONCLUSION

This article shows how further development and quantitative simulations of FACADE lead to explanations of data on figure-ground separation, amodal completion, and lightness perception. The lightness percepts illustrate how the direction and amplitude of each effect can depend on a context-sensitive interplay of the boundary and surface processes that separate figure from ground. Some of these properties may be modeled by using neural filters, as is illustrated by the work of Blakeslee and McCourt (1997). On the other hand, explaining the full set of properties also requires an analysis of 3-D figure-ground and surface formation mechanisms. In particular, the model suggests how a wide range of percepts may arise as emergent properties of such ecologically vital processes as the size-disparity correlation, surface capture, and the asymmetry between near and far—including boundary and surface pruning and boundary enrichment—when these processes are activated by visual images and scenes.

REFERENCES

- ABBOTT, L. F., VARELA, J. A., SEN, K., & NELSON, S. B. (1997). Synaptic depression and cortical gain control. *Science*, *275*, 220-224.
- ADELSON, E. H. (1993). Perceptual organization and the judgment of brightness. *Science*, *262*, 2042-2044.
- AGOSTINI, T., & PROFFITT, D. R. (1993). Perceptual organization evokes simultaneous lightness contrast. *Perception*, *22*, 263-272.
- ANDERSON, B. L. (1997). A theory of illusory lightness and transparency in monocular and binocular images: The role of contour junctions. *Perception*, *26*, 419-454.
- BENARY, W. (1924). Beobachtungen zu einem Experiment über Helligkeitskontrast. *Psychologische Forschung*, *5*, 131-142. [Translated as "The influence of form on brightness contrast," in W. Ellis (Ed.), *A source book of Gestalt psychology*, 1939, London: Routledge & Kegan Paul]
- BLAKESLEE, B., & MCCOURT, M. E. (1997). Similar mechanisms underlying simultaneous brightness contrast and grating induction. *Vision Research*, *37*, 2849-2869.
- 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.
- BROWN, J. M., & WEISSTEIN, N. (1988). A spatial frequency effect on perceived depth. *Perception & Psychophysics*, *44*, 157-166.
- DALBY, T. A., SAILLANT, M. L., & WOOTEN, B. R. (1995). The relation of lightness and stereoscopic depth in a simple viewing situation. *Perception & Psychophysics*, *57*, 318-332.
- DESIMONE, R. (1991). Face-selective cells in the temporal cortex of monkeys. *Journal of Cognitive Neuroscience*, *3*, 1-8.
- DESIMONE, R., & UNGERLEIDER, L. G. (1989). Neural mechanisms of visual processing in monkeys. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 2, pp. 267-299). Amsterdam: Elsevier.
- DE VALOIS, R. L., & DE VALOIS, K. K. (1988). *Spatial vision*. New York: Oxford University Press.
- EGUSA, H. (1983). Effects of brightness, hue and saturation on perceived depth between adjacent regions in the visual field. *Perception*, *12*, 167-175.
- FAHLE, M., & WESTHEIMER, G. (1995). On the time-course of inhibition in the stereoscopic perception of rows of dots. *Vision Research*, *35*, 1393-1399.
- FIELD, D. J., HAYES, A., & HESS, R. F. (1993). Contour integration by the human visual system: Evidence for a local "association field." *Vision Research*, *33*, 173-193.

- FRANCIS, G. (1997). Cortical dynamics of lateral inhibition: Metacontrast masking. *Psychological Review*, **104**, 572-594.
- FRANCIS, G., & GROSSBERG, S. (1996a). Cortical dynamics of boundary segmentation and reset: Persistence, afterimages, and residual traces. *Perception*, **25**, 543-567.
- FRANCIS, G., & GROSSBERG, S. (1996b). Cortical dynamics of form and motion integration: Persistence, apparent motion, and illusory contours. *Vision Research*, **36**, 149-173.
- FRANCIS, G., GROSSBERG, S., & MINGOLLA, E. (1994). Cortical dynamics of feature binding and reset: Control of visual persistence. *Vision Research*, **34**, 1089-1104.
- GILCHRIST, A. (1977). Perceived lightness depends on perceived spatial arrangement. *Science*, **195**, 185-187.
- GILCHRIST, A. (1994). Introduction: Absolute versus relative theories of lightness perception. In A. Gilchrist (Ed.), *Lightness, brightness and transparency* (pp. 1). Hillsdale, NJ: Erlbaum.
- GROSSBERG, S. (1987). Cortical dynamics of three-dimensional form, color, and brightness perception: II. Binocular theory. *Perception & Psychophysics*, **41**, 117-158.
- GROSSBERG, S. (1994). 3-D vision and figure-ground separation by visual cortex. *Perception & Psychophysics*, **55**, 48-120.
- GROSSBERG, S. (1997). Cortical dynamics of three-dimensional figure-ground perception of two-dimensional pictures. *Psychological Review*, **104**, 618-658.
- GROSSBERG, S. (1999). How does the cerebral cortex work? Learning, attention, and grouping by the laminar circuits of visual cortex. *Spatial Vision*, **12**, 163-186.
- GROSSBERG, S., & KELLY, F. J. (1999). Neural dynamics of binocular brightness perception. *Vision Research*, **39**, 3796-3816.
- GROSSBERG, S., & McLOUGHLIN, N. (1997). Cortical dynamics of 3-D surface perception: Binocular and half-occluded scenic images. *Neural Networks*, **10**, 1583-1605.
- GRUNEWALD, A., & GROSSBERG, S. (1998). Self-organization of binocular disparity tuning by reciprocal corticogeniculate interactions. *Journal of Cognitive Neuroscience*, **10**, 199-215.
- GUZMAN, A. (1984). Decomposition of a visual scene into three dimensional borders. In O. Fischein (Ed.), *Artificial Intelligence* (Information Technology Series, Vol. VI, pp. 310-335). Reston, VA: AFIPS Press. (Original work published 1968)
- HERING, E. (1964). *Outlines of a theory of the light sense* (L. M. Hurvich & D. Jameson, Trans.). Cambridge, MA: Harvard University Press. (Original work published 1920)
- JOHNSON, S. P., & ASLIN, R. N. (1995). Perception of object unity in 2-month-old infants. *Developmental Psychology*, **31**, 739-745.
- KANIZSA, G. (1979). *Organization in vision: Essays on Gestalt perception*. New York: Praeger/Greenwood Publishing Group.
- KANIZSA, G. (1985). Seeing and thinking. *Rivista di Psicologia*, **49**, 7-30.
- KELLY, F. J., & GROSSBERG, S. (1997). Cortical dynamics of 3-D pop-out and amodal completion. *Investigative Ophthalmology & Visual Science*, **38**, S969.
- KELLY, F. J., & GROSSBERG, S. (1998). How does visual cortex determine surface brightness during figure-ground perception. *Investigative Ophthalmology & Visual Science*, **39**, S326.
- KINGDOM, F., & MOULDEN, B. (1991). White's effect and assimilation. *Vision Research*, **31**, 151-159.
- KNILL, D. C., & KERSTEN, D. (1991). Apparent surface curvature affects lightness perception. *Nature*, **351**, 228-230.
- KOFFKA, K. (1935). *Principles of Gestalt psychology*. New York: Harcourt, Brace.
- KOVÁCS, G., VOGELS, R., & ORBAN, G. A. (1995). Selectivity of macaque inferior temporal neurons for partially occluded shapes. *Journal of Neuroscience*, **15**, 1984-1997.
- MICHOTTE, A., THINES, G., & CRABBE, G. (1964). *Les compléments amodaux des structures perceptives*. Louvain: Publications Universitaires de Louvain.
- MISHKIN, M. (1982). A memory system in the monkey. *Philosophical Transactions of the Royal Society of London: Series B*, **298**, 85-95.
- MOULDEN, B., & KINGDOM, F. (1989). White's effect: A dual mechanism. *Vision Research*, **29**, 1245-1259.
- MÜNCKER, H. (1970). *Farbige Gitter, Abbildung auf der Netzhaut und Übertragungstheoretische Beschreibung der Farbwahrnehmung*. Habilitationsschrift. Ludwig-Maximilians-Universität, Munich.
- NAKAMURA, H., GATTASS, R., DESIMONE, R., & UNGERLEIDER, L. G. (1993). The modular organization of projections from areas V1 and V2 to areas V4 and TEO in macaques. *Journal of Neuroscience*, **13**, 3681-3691.
- 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.
- OLSON, S. J., & GROSSBERG, S. (1998). A neural network model for the development of simple and complex cell receptive fields within cortical maps of orientation and ocular dominance. *Neural Networks*, **11**, 189-208.
- O'SHEA, R. P., BLACKBURN, S. G., & ONO, H. (1994). Contrast as a depth cue. *Vision Research*, **34**, 1595-1604.
- PERRETT, D. I., MISTLIN, A. J., & CHITTY, A. J. (1987). Visual cells responsive to faces. *Trends in Neurosciences*, **10**, 358-364.
- PETTER, G. (1956). Nuove ricerche sperimentali sulla totalizzazione percettiva. *Rivista di Psicologia*, **50**, 213-227.
- SCHIRILLO, J., REEVES, A., & AREND, L. (1990). Perceived lightness, but not brightness, of achromatic surfaces depends on perceived depth information. *Perception & Psychophysics*, **48**, 82-90.
- SEKULER, A. B., & PALMER, S. E. (1992). Perception of partly occluded objects: A microgenetic analysis. *Journal of Experimental Psychology: General*, **121**, 95-111.
- SHIPLEY, T. F., & KELLMAN, P. J. (1992). Strength of visual interpolation depends on the ratio of physically specified to total edge length. *Perception & Psychophysics*, **52**, 97-106.
- SMALLMAN, H. S., & MACLEOD, D. I. A. (1994). Size-disparity correlation in stereopsis at contrast threshold. *Journal of the Optical Society of America A*, **11**, 2169-2183.
- SPEHAR, B., GILCHRIST, A., & AREND, L. (1995). The critical role of relative luminance relations in White's effect and grating induction. *Vision Research*, **35**, 2603-2614.
- STRERI, A., SPELKE, E., & RAMEIX, E. (1993). Modality-specific and amodal aspects of object perception in infancy: The case of active touch. *Cognition*, **47**, 251-279.
- TAYA, R., EHRENSTEIN, W. H., & CAVONIUS, R. C. (1995). Varying the strength of the Munker-White effect by stereoscopic viewing. *Perception*, **24**, 685-694.
- TODOROVIĆ, D. (1997). Lightness and junctions. *Perception*, **26**, 379-394.
- TOMMASI, L., BRESSAN, P., & VALLORTIGARA, G. (1995). Solving occlusion indeterminacy in chromatically homogeneous patterns. *Perception*, **24**, 391-403.
- VON DER HEYDT, R., PETERHANS, E., & BAUMGARTNER, G. (1984). Illusory contours and cortical neuron responses. *Science*, **224**, 1260-1262.
- VON HELMHOLTZ, H. (1962). *Treatise on physiological optics* (J. P. C. Southall, Trans.). New York: Dover.
- WATANABE, T., & CAVANAGH, P. (1993). Transparent surfaces defined by implicit X junctions. *Vision Research*, **33**, 2339-2346.
- WERTHEIMER, M. (1923). Untersuchungen zur Lehre von der Gestalt: II. *Psychologische Forschung*, **4**, 301-350. (Translated as "Laws of organization in perceptual forms," in W. D. Ellis, Ed., 1939, *A Source Book of Gestalt Psychology*. London: Routledge and Kegan Paul).
- WHITE, M. (1979). A new effect of pattern on perceived lightness. *Perception*, **8**, 413-416.
- ZAIDI, Q., SPEHAR, B., & SHY, M. (1997). Induced effects of backgrounds and foregrounds in three-dimensional configurations: The role of T-junctions. *Perception*, **26**, 395-408.