COMPUTER VISION, GRAPHICS, AND IMAGE PROCESSING 37, 116–165 (1987)

Neural Dynamics of Surface Perception: Boundary Webs, Illuminants, and Shape-from-Shading

STEPHEN GROSSBERG* AND ENNIO MINGOLLA†

Center for Adaptive Systems, Department of Mathematics, Boston University, 111 Cummington Street, Boston, Massachusetts 02215

Received July 5, 1986

A real-time visual processing theory is used to provide a new approach to the analysis of surface perception, notably shape-from-shading. The theory elsewhere has been used to explain data about boundary detection and completion, textural segmentation, depth perception, color and brightness perception, and striate-prestriate cortical interactions. Neural network interactions within a multiple scale boundary contour (BC) system and feature contour (FC) system are used to explain these phenomena. Each spatial scale of the BC system contains a hierarchy of orientationally tuned interactions, which can be divided into two successive subsystems called the OC filter and the CC loop. The OC filter contains two successive stages of oriented receptive fields which are sensitive to different properties of image contrasts. The OC filter generates inputs to the CC loop, which contains successive stages of spatially short-range competitive interactions and spatially long-range cooperative interactions. Feedback between the competitive and cooperative stages synthesizes a coherent, multiple scale structural representation of a smoothly shaded image, called a *boundary web*. Such a boundary web regulates multiple-scale filling-in reactions within the FC system which generate a percept of form-and-color-in-depth. Computer simulations establish key properties of a boundary web representation: nesting of boundary web reactions across spatial scales, coherent completion and regularization of boundary webs across incomplete image data, and relative insensitivity of boundary webs to illumination level and highlights. The theory clarifies data about interactions between brightness and depth percepts, transparency, influences of highlights on perceived surface glossiness, and shape-from-texture gradients. The total network suggests a new approach to the design of computer vision systems, and promises to provide a universal set of rules for 3D perceptual grouping of scenic edges, textures, and smoothly shaded regions.

1. THE PROBLEM OF MODELING SURFACE PERCEPTION

The perception of surfaces is arguably the single most important competence of the visual systems of terrestrial mammals. The ground that affords support for locomotion and the bodies of predators, prey, and conspecifics are bounded by surfaces of various shapes and appearances. The state of scientific understanding of the processes of surface perception, however, is such that few if any pages are devoted to the subject in many textbooks on visual perception. A number of recent experimental studies have examined important aspects of surface perception, but as yet no coherent psychological theory with broad explanatory power has emerged.

*Supported in part by the Air Force Office of Scientific Research (AFOSR 86-0149 and AFOSR F49620-86-C-0037), the Army Research Office (ARO DAAG-29-85-K-0095), and the National Science Foundation (NSF DMS-84-13119).

†Supported in part by the Air Force Office of Scientific Research (AFOSR 86-0149).

0734-189X/87 $3.00

Moreover, as we will argue later in this paper, many of the computational models of surface perception to date have failed to solve or in some cases even identify some of the most fundamental problems in dealing with surfaces. As a result, these models lack both biological plausibility and the capacity to be readily implemented in functional, real-time vision machines. The present article begins to extend a rapidly growing, neurally based theory of 3D form perception to the domain of surface perception.

For decades before the subject became fashionable, Gibson studied the perception of surfaces and their properties. If his only contribution had been to emphasize the importance of gradients, whether of motion parallax, luminance, or textural properties, to the perception of surfaces, that contribution would by itself have been important [22]. A major goal of this paper is accordingly to introduce and analyze the notion of a *boundary web*, wherein the same mechanisms responsible for detecting discrete luminance or textural boundaries are also capable of synthesizing form-sensitive compartments, or dynamically generated coordinate systems, in response to static scenic gradients.

Gibson, however, studied more than gradients, going on to identify up to nine surface properties accessible by the visual and haptic systems [24, 69]. These nine properties, including several relating to surface reflectance, illumination conditions, and the relative transparency, opacity, or luminosity of surfaces, are deemed to be perceivable simultaneously with the shape or form or, as Gibson preferred, the layout of surfaces. The present paper will not address all of the phenomenological richness of Gibson's classification. We will, however, present a mechanistically plausible theory which takes seriously the need to provide a coherent account of how we can, in different but in equally real ways, perceive both the structural, shape-related properties of surfaces, and the more evanescent modes of surface appearance. Our theory is a natural one for such a development in three dimensions (3D), because its analysis of boundary completion and textural grouping has already led to the articulation of the analogous distinction in two dimensions, namely that of seeing contrast in luminance or hue versus recognizing form or pattern [35, 37].

With this addition, the theory has been shown competent to analyse perceptual data about surface perception, boundary completion, textural segmentation, motion-induced segmentation, stereopsis, hypercuity, monocular and binocular brightness perception, binocular rivalry, McCullough effect, border distinctness, lateral inhibition within spatial frequency channels, control of object superiority effect by emergent features, metacontrast, transparency, stereoscopic capture, and illusory features. In addition, the theory has suggested explanations and a number of new predictions about the organization and reciprocal interactions among the hypercolumns, blobs, and stripes of the visual cortical areas $V1$, $V2$, and $V4$, and has successfully predicted several recently discovered perceptual and neural phenomena [15, 16, 34–37].

Our analyses of these phenomena have led to a qualitatively new computational theory of how visual systems are designed, including the identification of several new uncertainty principles which visual interactions are designed to surmount. These developments have led to a number of revolutionary conclusions, whose paradoxical nature is best perceived when they are expressed simply and without technical caveats or interpretations. These conclusions include: All boundaries are invisible. All line ends are illusory. Boundaries are formed discontinuously. Orienta-

tionally fuzzy computations are the basis for orientationally sharp segmentations. Positionally fuzzy computations are the basis for positionally sharp percepts of 3-dimensional space.

This paper is divided into two parts. The first part describes the issues to be probed and briefly reviews our previously existing theoretical machinery. The second part extends the theory to cope with surface perception of shaded images. The main goal of this paper is to outline this neurally based theory of the perception of surfaces, including the process of deriving shape from shading. Because what we take to be the explanatory requirements of such a theory differ from the requirements of most computational models, Section 2 describes some criteria that we hope to meet in a fully developed theory of surface shape and quality perception. Section 3 then reviews several important computational models for recovering shape from shading, and criticizes them from the standpoint of the criteria articulated in Section 2. In Section 4 we examine some of the perceptual data on perception of solid shape from shading, in the light both of theoretical criteria and of predictions from the computational models. Section 5 begins to summarize key aspects of our theory as it has already been developed to deal with real-time boundary detection and textural segmentation phenomena. This brief review of our theory leads to a discussion explicitly linking previously described mechanisms and concepts to the requirements of the extended theory as it treats surface perception. Several types of perceptual data are reanalysed from this perspective. The results of computer simulations of boundary webs are presented in Sections 13, 21, and 22.

2. THE EXPLANATORY REQUIREMENTS OF A THEORY OF SURFACE PERCEPTION

Consideration of the perception of extended, smoothly curved surfaces in 3D, especially with respect to a source of visual information as volatile as shading, requires great care in the choice of desirable goals and attributes for a theory. For instance, we claim that to approach the problem of surface perception in shaded images as simply one of shape-from-shading is so incomplete as to be misleading. In the first place, "shading" is an imprecise term. Taken to refer to available luminance distributions in an image, its usage is relatively noncontroversial. Used in that sense, shading is rarely deemed to be a sufficient basis upon which to completely determine shape. Most existing computational models recover shape from information about image shading, illumination conditions, and surface reflectances. Such analyses tend to obscure the fact that determination of the values of "shading" variables at an area of an image—such as illumination conditions at an area on a surface, pigmentation or texture of that area, or presence of cast shadows—is an *achievement* of the visual system, simultaneous with the determination of the surface shape depicted within that area of the image [1, 4, 12, 20, 25–27, 47].

We suggest that an adequate theory of surface perception must simultaneously explain how the invariant properties of surface shape are usually perceived with high fidelity, despite gross perturbations of surface appearance, without assuming knowledge about highly changeable variables such as illumination conditions, or treating them as noise.

3. COMPUTATIONAL MODELS FOR RECOVERING SHAPE-FROM-SHADING

A number of models for recovering shape-from-shading, while differing in certain details, do not meet these requirements [39–41, 43, 65, 81]. In such models, an

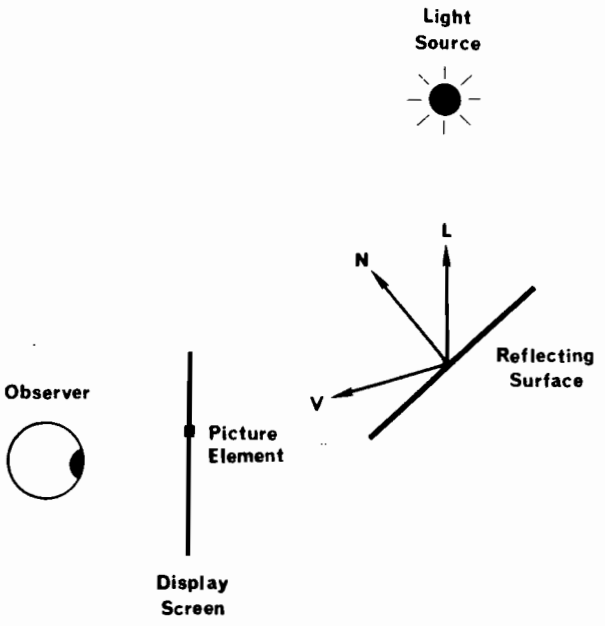

FIG. 1. The geometric variables needed for a model to simulate an image of a surface in three dimensional space. Symbols L, V, and N are all unit vectors from a point on the reflecting surface; L lies in the direction of the light source; V lies in the direction of a picture element and the point of observation; and N is perpendicular to the surface. (Adapted with permission from Todd and Mingolla [78].)

attempt is made to mathematically invert the image formation process. Algorithms are proposed to reconstruct the environment most consistent with image data from inferences about what is known or assumed about image generation. Figure 1 displays a model of shaded image generation, involving parameters for positions of eye, illuminated surface, and illumination source, as well as illuminant intensity and surface reflectance. Elaborate computer graphics implementations of such models are by now capable of making stunningly realistic synthetic images of a variety of surfaces.

What makes recovery of shape-from-shading a problem is precisely that many combinations of image generation parameters can result in any given pixel intensity. The computational analyses for recovering shape just cited all attempt to somehow reduce the number of unknown parameters or to find constraints on combinations of their values so that the number of equations to be solved does not exceed the number of unknowns that can be measured from image data. Typical moves for accomplishing this goal are: to assume knowledge of illuminant direction; to first independently measure illuminant direction from the image itself; to assume that no illumination edges (e.g., cast shadows) fall on a surface; to assume knowledge of type of reflectance for each surface; to assume that all surface reflectances are Lambertian.

The foregoing computational procedures invoke various consistency requirements across local parameter estimates; and special handling of boundary conditions or image extrema can cause the effects of certain local gradient measures on recovered shape to propagate across large image regions. Nevertheless, these procedures for

recovering shape take as their primary analytical units the local values of image intensity gradients, and aim to produce estimates of local surface orientation and/or depth as their outputs.

A more global type of analysis has been contributed in a series of papers by Koenderink and van Doorn [51–53]. While these papers do not propose an algorithm or procedure for actually recovering shape, they provide analyses of the kinds of invariant image structures characteristic of particular shapes and illumination conditions. That is, instead of ending the analysis of image generation with techniques for determining individual pixel intensities along the lines described in Fig. 1, they proceed to analyze the resulting configurations of luminance extrema and inflection points, and the connectivities of isophotes (contours of equal luminance). Tacit in their approach is the implication that the major perceptual variables for recovering shape are not parametric, ratio-scaled estimates of local luminance gradients, but the more categorical or ordinal global structures which are induced by those gradients. This characteristic is desirable from the standpoint of noise tolerance and receiver-induced nonlinearities in measurement, insofar as the topology of isophotes is likely to remain relatively intact while local gradient measures are severely distorted. Similarly, Koenderink and van Doorn claim that their analysis holds for a wide variety of illumination conditions and, with certain extensions, surface reflectance functions. For example, one does not have to know illuminant direction (beyond ruling out degeneracies) in order to apply their analysis to infer from isophote configuration to qualitative category of surface shape. Thus, Koenderink and van Doorn's analysis is both more modest concerning what can or ought to be computed from images and more robust in its handling of fluctuations induced by measurement noise or changes in environmental conditions.

Neither type of approach reviewed in this section indicates how to explain the simultaneous apprehension of shape and of the surface appearances which are perceived as occurring on a shape. We will argue that an analysis of the *interactions* between the mechanisms determining shape and appearance is needed to fully understand the design of either type of mechanism. For example, in many images—such as the monocularly perceived gradients of Gibson [22] or the binocularly perceived stereograms of Kaufman [48] and Julesz [44]—small regions of the image are perceptually ambiguous in isolation, yet can be unambiguously perceived as part of a surface when their monocular or binocular contexts are also processed. Thus global aspects of form can alter the appearance of local image regions. Appearance and shape are not independently processed, despite the fact that appearance can vary greatly on an invariant shape [64]. The hypothesis that independent modules process distinct visual properties such as appearance and shape is fundamentally flawed [57].

4. PERCEPTUAL DATA ON PERCEPTION OF SOLID SHAPE IN SHADED IMAGES

Two of the central issues for experimental work on surface perception are: (1) Is the visual system's analysis of shape based on some prior knowledge or independent assessment of illuminant direction? and (2) Must knowledge about surface reflectance precede recovery of shape?

A common belief about the first issue is that the answer is affirmative. Classically, this conclusion has been based on some rather striking reversals of perceived concavity and convexity that are achievable by simply rotating an image 180° in the

frontoparallel plane [14]. We believe that the significance of this demonstration has been overstated, however. In the first place, many aspects of perceived shape are preserved even over this gross transformation. Switching concavity and convexity does not change Gaussian curvature, and both minima and maxima of perceived depth remain extrema, though of opposite kind, after the image is rotated. More importantly, however, is the inapplicability of the symmetric reversal of illuminant direction to the more relevant and common cases. That is, what are the consequences of misconstruing illuminant elevation by, say, 20°? What if clouds or irregularly spaced lamps produce conditions with no single prevailing direction of illumination, wherein the directions from which most illumination falls on surfaces varies in continuously graded and unpredictable ways?

In a direct experimental test of the first issue, Mingolla and Todd [64] found only weak or nonexistent correlations between errors in judgments of direction of illumination and judgments of surface orientation for subjects viewing computer generated shaded images of ellipsoids. The computed angles of elevation of point source illumination varied by 30° in this study. The assumption that prior knowledge or computation of illuminant direction is required for perception of shape was not supported by these data.

Regarding the second issue, Todd and Mingolla [78] found that highlights enhanced the perception of curvature in computer generated displays of cylindrical surfaces, and Mingolla and Todd [64] found that highlights had no effect on surface orientation judgments for ellipsoid surfaces. In none of these experiments did subjects have any information about surface reflectance besides the image itself. Beck and Prazdny [9] showed subjects a number of photographs of vases, some of which had been retouched with various perturbations of contours and highlights. They concluded that the perception of glossiness is a direct response to local visual stimulation rather than an inference based on stored memories. It appears that the human visual system determines values of glossiness and direction of illumination with varying accuracy, but in any event *concurrently* in the logical and perhaps temporal sense, with its determination of surface shape.

Mingolla and Todd [64] reported other findings of relevance to modeling of shape perception. Using a procedure for fitting ellipsoid shape and orientation parameters to subject errors, they reconstructed the perceived shapes and orientations of ellipsoids for all subjects. In virtually every case, the orientations of the axes of the recovered ellipsoids were more aligned with the axes of the display screen and the line of sight than were the computed axes for the displayed ellipsoids. That is, subjects experienced more symmetrical figures than were shown according the computer graphics shading model. Thus global, configurational processes were clearly at work.

5. BOUNDARY CONTOURS AND FEATURE CONTOURS

The basic concepts of our theory are described at length in Cohen and Grossberg [15, 16], Grossberg [34, 35], and Grossberg and Mingolla [36, 37]. This and the next two sections provide a brief overview of the theory, highlighting those concepts relevant to our analysis of surface perception.

Human observers cannot distinguish those parts of a percept that are derived directly from retinal stimulation from those parts that are occluded by retinal veins

and the blind spot. Such a perceptual synthesis accomplishes two tasks: boundaries induced by discontinuities in retinal stimulation have to be matched and completed over the obstructions, and the relevant featural quality (brightness or color) must be filled in. By characterizing the different processing rules that are obeyed, we have argued that boundary completion and featural filling-in are mechanistically different processes.

While a great many perceptual phenomena can be used to argue for the two processes, we have highlighted three. The Land [56] color and brightness experiments indicate that perceived color and brightness are based primarily on measurements taken at scenic gradients, notably edges. Stabilized image experiments [54, 82] have shown that when luminance or color edges are stabilized, perceived colors flow throughout formerly bounded regions. Finally, reverse-contrast illusory figures [36, 66, 71] show that the boundary completion process (Fig. 2) is insensitive to the direction of contrast (e.g., light-dark vs dark-light). Such data support rules which distinguish between two systems (Fig. 3), which we have called the boundary contour system (BC system) and the feature contour system (FC system).

The BC system is sensitive to the amount and orientation of scenic contrast, but not to its direction of contrast. The FC system is sensitive to both the amount and the direction of scenic contrast. In the BC system, boundary completion occurs *inward* in an *oriented* fashion from *pairs* of inducing BC generators, as in Fig. 2. In the FC system, featural filling-in occurs *outward* in an *unoriented* fashion from *individual* inducing FC generators. FC filling-in signals spread until they hit a boundary induced within the FC system or are attenuated by their own spatial spread. This filling-in reaction is described by a spatial diffusion process. It performs spatial averaging of activation whose spatial bandwidth is dynamically controlled by the strength and density of boundary signals from the BC system.

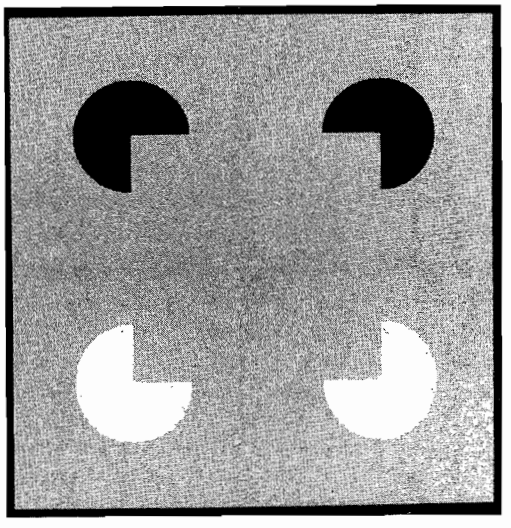

FIG. 2. An illusory square is induced by two black and two white figures on a gray background. Illusory contours can thus join edges with opposite directions of contrast. (This effect may be weakened by photographic reproduction processes. Reprinted by permission from Grossberg and Mingolla [36].)

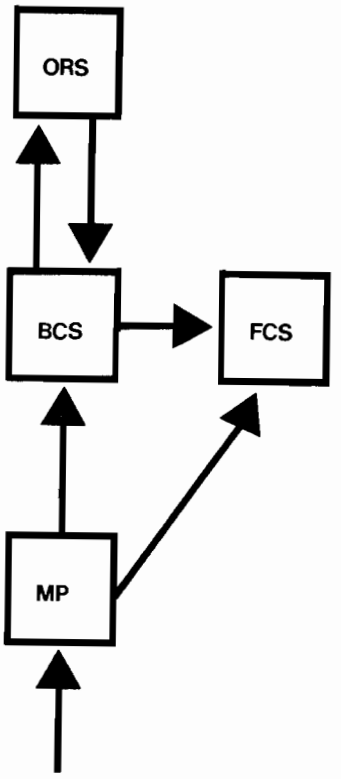

FIG. 3. A macrocircuit of processing stages: Monocular preprocessed signals (MP) are sent independently to both the boundary contour system (BCS) and the feature contour system (FCS). The BCS pre-attentively generates coherent boundary structures from these MP signals. These structures send outputs to both the FCS and the object recognition system (ORS). The ORS, in turn, rapidly sends top-down learned template signals to the BCS. These template signals can modify the pre-attentively completed boundary structures using learned information. The BCS passes these modifications along to the FCS. The signals from the BCS organize the FCS into perceptual regions wherein filling-in of visible brightnesses and colors can occur. This filling-in process is activated by signals from the MP stage.

6. HIERARCHICAL RESOLUTION OF UNCERTAINTY: ORIENTED CONTRAST DETECTORS

To date, our characterization of the BC and FC systems has primarily been based upon their reactions to images built up from configurations of step-function luminance edges. Even in such sharply defined image environments, fundamental measurement uncertainties exist that require compensatory processing at later processing stages. From the very earliest stages of BC system processing, image contrasts are grouped and regrouped in order to generate BC configurations of ever greater global coherence and structural invariance.

For example, even the oriented masks at the earliest stage of BC system processing regroup image contrasts (Fig. 4). Such masks are oriented *local contrast* detectors, rather than edge detectors. This property enables them to fire in response to a wide variety of spatially nonuniform image contrasts that do not contain edges, as well as in response to edges. In particular, such oriented masks can respond to spatially nonuniform densities of unoriented textural elements, such as dots. They can also respond to spatially nonuniform densities of surface gradients. Thus by

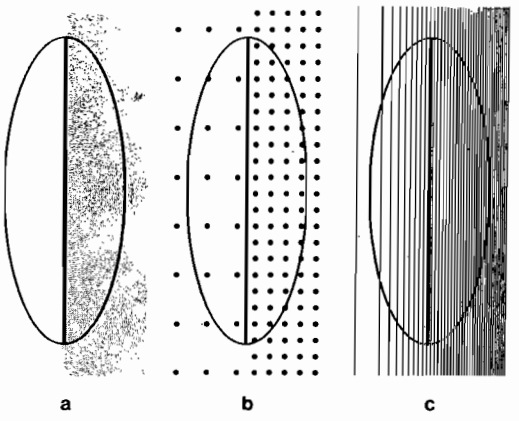

FIG. 4. Oriented masks respond to amount of luminance contrast over their elongated axis of symmetry, regardless of whether image contrasts are generated by (a) luminance step functions, (b) differences in textural distribution, or (c) smooth luminance gradients (indicated by the spacings of the lines).

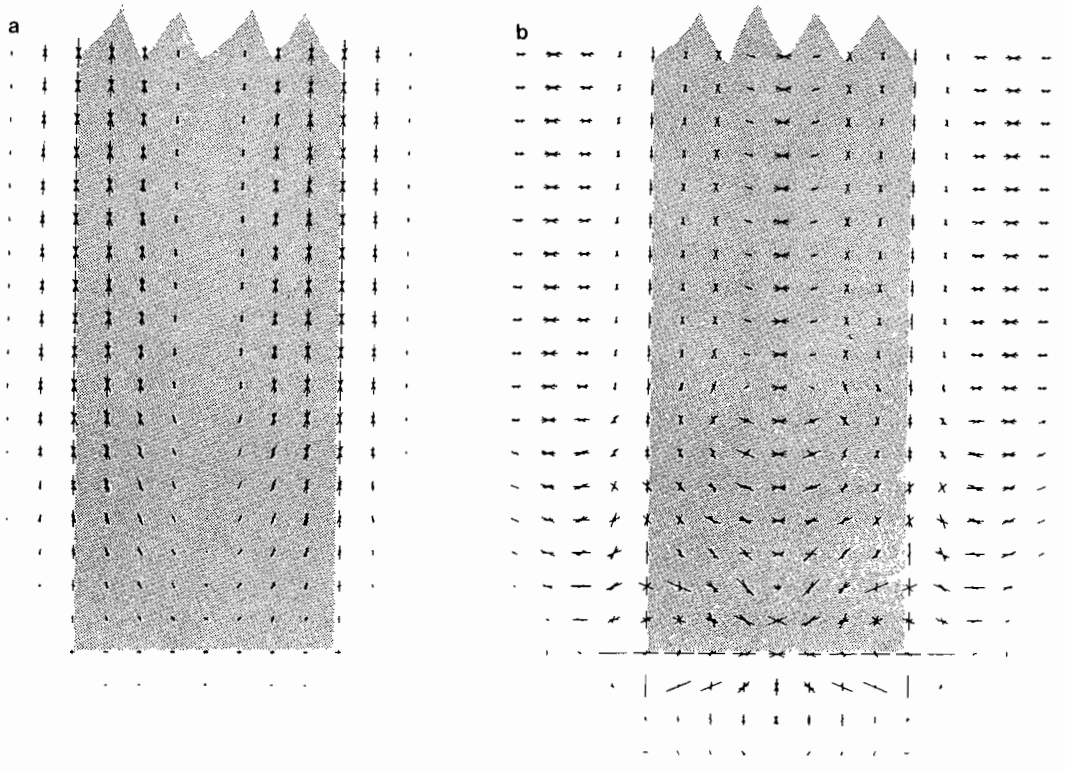

FIG. 5. (a) An orientation field: Lengths and orientations of lines encode the relative sizes of the activations and orientations of the input masks at the corresponding positions. The input pattern, which is a vertical line end as seen by the receptive fields, corresponds to the shaded area. Each mask has total exterior dimensions of 16 × 8 units, with a unit length being the distance between two adjacent lattice positions. (b) Responses of a network with two stages of short-range competition to the orientation field of Fig. 5a: A process called end cutting generates horizontal activations at line end locations that receive small and orientationally ambiguous input activations.

sacrificing a certain amount of spatial resolution in order to detect oriented local contrasts, these masks achieve a general detection characteristic which works equally well in response to boundaries, textures, and surfaces.

The fact that these receptive fields are *oriented* greatly reduces the number of possible groupings into which their target cells can enter. On the other hand, in order to detect oriented local contrasts, these receptive fields must be elongated along their preferred axis of symmetry. Then the cells can preferentially detect differences of average contrast across this axis of symmetry, yet can remain silent in response to differences of average contrast that are perpendicular to the axis of symmetry. Such receptive field elongation creates even greater positional uncertainty about the exact locations within the receptive field of the image contrasts which fire the cell. This positional uncertainty becomes acute during the processing of image line ends and corners. Figure 5a illustrates the breakdown of oriented receptive field sensitivity at a line end. In summary, there exists an uncertainty principle which says: Orientational "certainty" implies positional "uncertainty" at line ends and corners.

We have shown how two subsequent stages of spatially short-range competition can compensate for this breakdown through what we have called an *end cutting* process (Fig. 5b).

7. THE OC FILTER AND THE SHORT-RANGE COMPETITIVE STAGES

The processing stages that are hypothesized to generate end cuts are summarized in Fig. 6. First, oriented receptive fields of like position and orientation, but opposite direction-of-contrast, cooperate at the next processing stage to activate cells whose receptive fields are sensitive to the same position and orientation as themselves, but are insensitive to direction-of-contrast. These target cells maintain their sensitivity to *amount* of oriented contrast, but not to the *direction* of this oriented contrast. Such model cells, which play the role of complex cells in Area 17 of the visual cortex, pool inputs from receptive fields with opposite directions-of-contrast in order to generate boundary detectors which can detect broadest possible range of luminance or chromatic contrasts [35, 75]. These two successive stages of oriented contrast-sensitive cells are called the OC filter [37].

The output from the OC filter successively activates two types of short-range competitive interaction whose net effect is to generate end cuts. First, a cell of prescribed orientation excites like-oriented cells corresponding to its location and inhibits like-oriented cells corresponding to nearby locations at the next processing stage. In other words, an on-center off-surround organization of like-oriented cell interactions exists around each perceptual location. The outputs from this competitive mechanism interact with the second competitive mechanism. Here, cells compete that represent different orientations, notably perpendicular orientations, at the same perceptual location. This competition defines a push–pull opponent process. If a given orientation is excited, then its perpendicular orientation is inhibited. If a given orientation is inhibited, then its perpendicular orientation is excited via disinhibition.

These competitive rules generate end cuts as follows. The strong vertical activations along the edges of a scenic line, as in Fig. 5a, inhibit the weak vertical activations near the line end. These inhibited vertical activations, in turn, disinhibit horizontal activations near the line end, as in Fig. 5b. Thus the positional uncer-

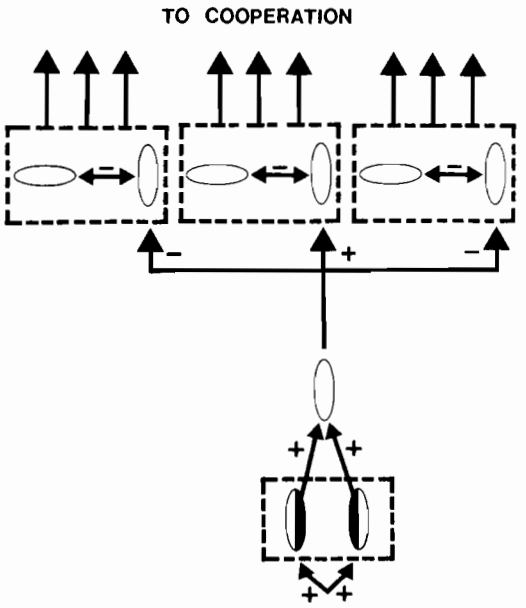

FIG. 6. Early stages of boundary contour processing: At each position cells exist with elongated receptive fields of various sizes which are sensitive to orientation, amount-of-contrast, and direction-of-contrast. Pairs of such cells sensitive to like orientation but opposite directions-of-contrast (lower dashed box) input to cells that are sensitive to orientation and amount-of-contrast but not to direction-of-contrast (white ellipses). These cells, in turn, excite like-oriented cells corresponding to the same position and inhibit like-oriented cells corresponding to nearby positions at the first competitive stage (upper dashed boxes). At this stage, cells corresponding to the same position but different orientations inhibit each other via a push–pull competitive interaction.

tainty at line ends that is caused by orientational tuning is eliminated by the interaction of two short-range competitive mechanisms.

8. LONG-RANGE COOPERATION: BOUNDARY COMPLETION

The outputs from the competition input to a spatially long-range cooperative process, called the *boundary completion* process. This cooperative process helps to build up sharp coherent global boundaries from noisy local boundary fragments. In the first stage of this boundary completion process, outputs from the second competitive stage corresponding to (approximately) like-oriented cells that are (approximately) aligned across perceptual space cooperate to begin the synthesis of an intervening boundary. For example, such a boundary completion process can span the retinal blind spot and the faded stabilized images of retinal veins. The same boundary completion process completes the sides of the Kanizsa square in Fig. 2. To understand further details about this boundary completion process, it is important to understand that the boundary completion process overcome a different type of informational uncertainty than that depicted in Fig. 5.

This type of uncertainty is clarified by considering Figs. 7. Figure 7a shows that the tendency to form boundaries that are perpendicular to line ends is a strong one; the illusory boundary forms sharp corners to keep the boundary perpendicular to the inducing line ends. Figure 7b shows, however, that the boundary completion process can generate a boundary that is not perpendicular to the inducing line ends

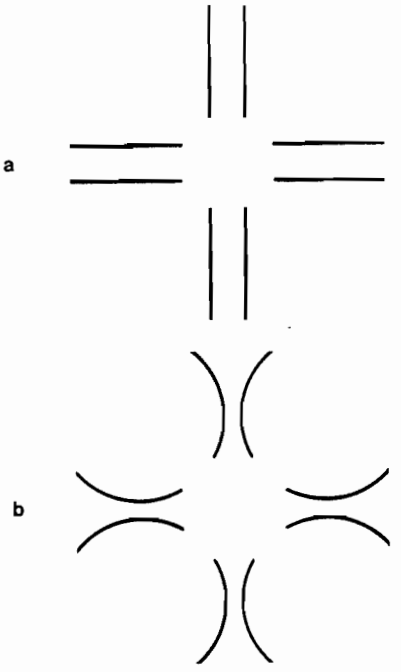

FIG. 7. (a) An illusory square generated by lines with orientations that are perpendicular to the illusory contour. (b) An illusory square can also be generated by lines with orientations that are not exactly perpendicular to the illusory contour.

in certain circumstances. A comparison of Figs. 7a and b indicates the nature of the other problem of uncertain measurement. Figures 7a and b show that boundary completion can occur within a *band* of orientations. These orientations include the orientations that are perpendicular to their inducing line ends (Fig. 7a), as well as nearby orientations that are not perpendicular to their inducing line ends (Fig. 7b). Figure 5b illustrates the existence of such a band of end cuts at the end of a scenic line. Such a band of possible orientations increases the probability that spatially separated boundary fragments can group cooperatively into a global boundary. If only a single orientation at each spatial location were activated, then the probability that these orientations could precisely line up across perceptual space to initiate boundary completion would be small. Bands of orientations facilitate the initiation of the perceptual grouping process that controls boundary completion.

This orientational uncertainty can, however, cause a serious loss of acuity in the absence of compensatory processes. If *all* orientations in each band could cooperate with *all* approximately aligned orientations in nearby bands, then a fuzzy band of completed boundaries, rather than a single sharp boundary, could be generated. The existence of such fuzzy boundaries would severely impair visual acuity. Figure 7 illustrates that only a single sharp boundary becomes visible despite the existence of orientational bands. How does the nervous system resolve the uncertainty produced by the existence of orientational bands? How is a single global boundary chosen from among the many possible boundaries that fall within the local orientational bandwidths?

Our answer to these questions suggests a basic reason why the cooperative stage of BC system processing sends feedback signals to the second competitive stage of BC system processing. This cooperative feedback provides a particular grouping of orientations with competitive advantage over other possible groupings.

9. BOUNDARY CHOICES BY A COOPERATIVE–COMPETITIVE FEEDBACK NETWORK: THE CC LOOP

We assume that pairs of similarly oriented and spatially aligned cells are needed to activate the cooperative cells that subserve boundary completion (Fig. 8a). These cells, in turn, feed back excitatory signals to like-oriented cells at the stage where competition occurs between orientations at each position. In Fig. 8a, for example, feedback signals are triggered in pathway 2 if sufficient activation simultaneously occurs in both of the feedforward pathways labelled 1. Then both pathways labelled

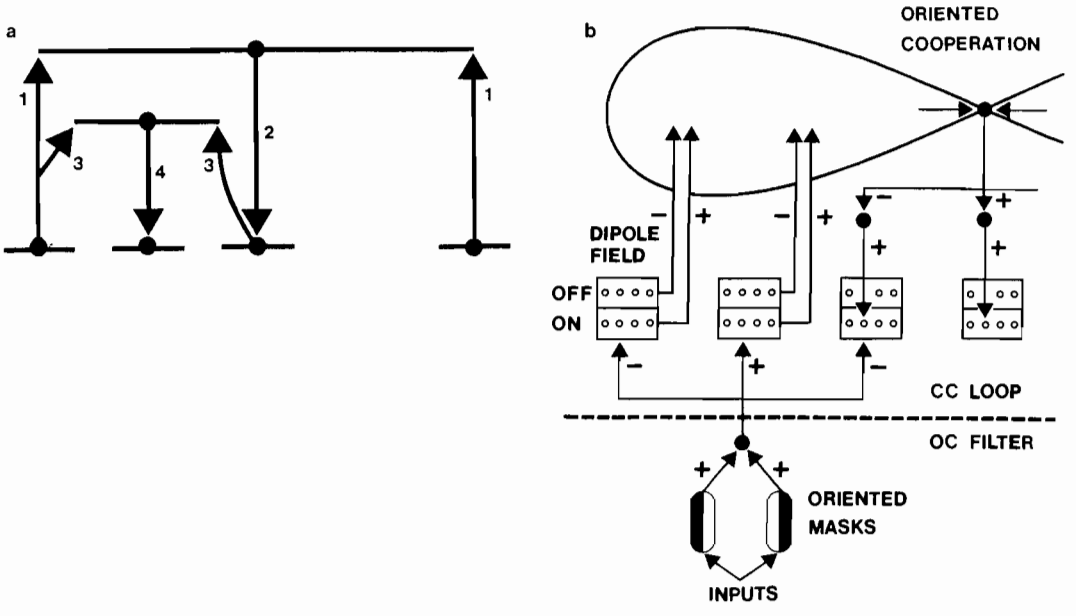

FIG. 8. An overview of cooperative feedback: (a) The pair of pathways 1 activate positive boundary completion feedback along pathway 2. Then pathways such as 3 activate positive feedback along pathways such as 4. Rapid completion of a sharp boundary between the locations of pathways 1 can thereby be generated by a spatially discontinuous bisection process. (b) Circuit diagram of the boundary contour system: Inputs activate oriented masks which cooperate at each position and orientation before feeding into an on-center off-surround interaction. This interaction excites like-orientations at the same position and inhibits like-orientations at nearby positions. The affected cells are on-cells within a dipole field. On-cells at a fixed position compete among orientations. On-cells also inhibit off-cells which represent the same position and orientation. Off-cells at each position, in turn, compete among orientations. Both on-cells and off-cells are tonically active. Net excitation of an on-cell excites a similarly oriented cooperative receptive field at a location corresponding to that of the on-cell. Net excitation of an off-cell inhibits a similarly oriented cooperative receptive field of a bipole cell at a location corresponding to that of the off-cell. Thus, bottom-up excitation of a vertical on-cell, by inhibiting the horizontal on-cell at that position, disinhibits the horizontal off-cell at that position, which in turn inhibits (almost) horizontally oriented cooperative receptive fields that include its position. Sufficiently strong net positive activation of both receptive fields of a cooperative cell enables it to generate feedback via an on-center off-surround interaction among like-oriented cells. On-cells which receive the most favorable combination of bottom-up signals and top-down signals generate the emergent perceptual grouping.

3 can trigger feedback in pathway 4. The feedback cycle, once initiated, can rapidly complete an oriented boundary between pairs of inducing scenic contrasts via a spatially discontinuous bisection process.

This boundary completion process realizes a type of real-time statistical decision theory. Each cooperative cell is sensitive to the position, orientation, density, and size of the inputs that it receives from the second competitive processing stage. Each cooperative cell performs like a type of statistical "and" gate, since it can fire feedback signals to the second competitive processing stage only if both of its branches are sufficiently activated. We call such cells *bipole* cells. The entire cooperative–competitive feedback network is called a CC loop. The CC loop can choose a sharp boundary from a band of possible boundaries for the following reason [36, 37].

As in Fig. 5b, certain orientations at a given position are more strongly activated than other orientations. Suppose that the cells which encode a particular orientation at two or more approximately aligned positions can more strongly activate their target cooperative cells than can the cells which encode other orientations. Then competitive cells of similar orientation at intervening positions will receive more intense excitatory feedback from their cooperative cells. This excitatory feedback enhances the activation of these cells relative to the activation of cells which encode other orientations. This advantage enables the favored orientation to suppress alternative orientations due to the orientational competition (Fig. 6). Cooperative feedback hereby provides the network with contrast-enhancing properties that enable it to choose a single sharp boundary from among a band of possible boundaries by using the short-range competitive interactions. Figure 8b joins together the OC filter with the CC loop to describe the BC system circuit which Grossberg and Mingolla [37, 38] used to simulate properties of perceptual grouping.

The ability of the CC loop to group fuzzy local orientational bands into sharp global boundaries illustrates how the imposition of simple perceptual constraints can lead to unsuspected mechanistic conclusions. The need to generate end cuts (Fig. 5) led to the hypothesis that orientational competition occurs corresponding to each perceptual location (Fig. 6) at a prescribed stage of boundary processing. Once orientational competition is available, the cooperative process which it feeds can use the *same* orientational competition to also make sharp boundary choices and to choose coherent textural segmentations. The same competitive laws are also used to define disparity-sensitive binocular boundaries, to suppress binocularly generated double images, and to generate properties of hyperacuity [35]. These competitive stages, which are new to our theory, have proved to be a critical ingredient for explaining a wide range of perceptual and neural data about vision.

The receptive fields of cooperative cells are taken to have the bipolar shape indicated in Fig. 9a. The spatial averaging within each bipole receptive field regroups inputs from the competitive layers using different rules than those whereby the oriented receptive fields in Fig. 4 regroup their inputs. Both types of regrouping enable their respective cells to process a large set of scenic images, but both types of regrouping also generate potential sources of uncertainty which are corrected by interactions with competitive mechanisms. In particular, a horizontally oriented branch of a bipole receptive field is excited by (approximately) horizontally tuned competitive cells, but is inhibited by (approximately) vertically tuned competitive cells within its receptive field (Fig. 8b). A bipole receptive field emits a net

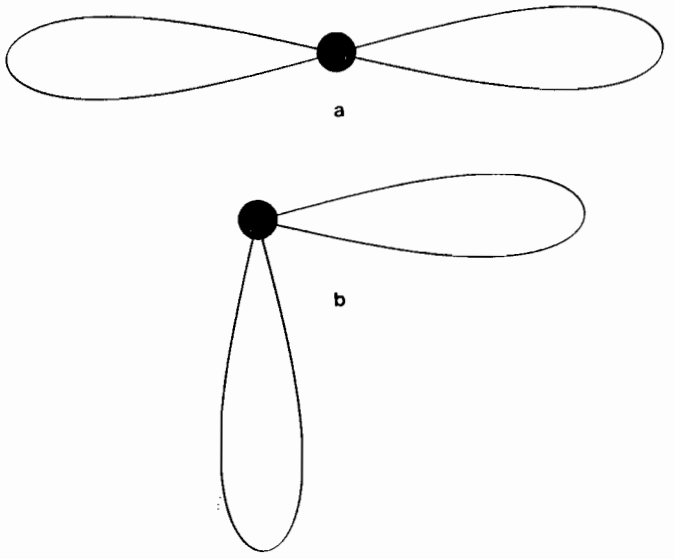

FIG. 9. Bipole cells: (a) The receptive fields of the bipole cells used in our simulations have colinear branches. (b) Bipole cell receptive fields may have noncolinear branches. Such cells would be sensitive to corners, as well as to the angle subtended at these corners.

excitatory signal to its cell body only if its net input from all oriented competitive cells is positive. A bipole cell body can fire positive feedback signals towards the competitive level only if it receives large enough signals from both of its receptive fields. Thus a cooperative bipole cell can fire only if both of its receptive fields detect statistically significant trends in approximately colinear oriented contrasts.

Due to this colinearity constraint, bipole cells generate straight emergent boundaries in response to some displays, such as the Kanizsa figure (Fig. 2). On the other hand, each receptive field of a bipole cell performs a statistical average which enables it to fire in response to oriented contrasts which are not exactly colinear. Thus bipole cells can also fire in response to image curves as well as to spatially curved statistical contrasts that are generated by textures or surface gradients.

Once the concept of a bipole cell is defined, its possible role in detecting sharp corners is also raised. A bipole cell whose individual receptive fields are not colinear (Fig. 9b) would respond selectively to image corners within a range of prescribed angles. The competition between orientations within each bipole receptive field prevents both receptive fields from simultaneously inputting to their bipole cell body except when a corner of prescribed orientation is present. The possible role of noncolinear bipole receptive fields in boundary completion and visual pattern recognition will not be further analyzed here. It is mentioned primarily to emphasize that cooperative bipole cells can selectively respond to statistically non-colinear image contrast distributions.

10. FROM ANALOG CONTRASTS TO DIGITAL STRUCTURES

The CC loop, as a whole, behaves like an on-line statistical decision machine in response to its input patterns. It senses only those groupings of perceptual elements that possess enough "statistical inertia" to drive its cooperative–competitive feedback exchanges towards a non-zero stable equilibrium configuration. After a

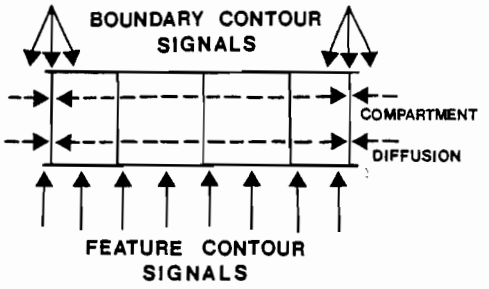

FIG. 10. A filling-in syncytium: Feature contour signals activate cell compartments that permit rapid lateral diffusion of activity, or potential, across their compartment boundaries, except at those compartment boundaries that receive boundary contour signals from the boundary contour system of Fig. 3. Thus boundary contour signals create barriers to filling-in within the feature contour system.

boundary structure emerges from the cooperative–competitive feedback exchange, it is stored in short-term memory by the feedback signals until it is actively reset by the next perceptual cycle. While the boundary structure is active, it possesses hysteretic and coherent properties due to the averaging properties of the receptive fields, the persistent suppression of alternative groupings by the competition, the persistent enhancement of the winning grouping by the cooperation, and the self-sustaining activation by the feedback. Thus, whereas inputs to the CC loop retain their "analog" sensitivity to amount-of-contrast in order to properly bias its operation to favor statistically important image groupings, once the CC loop responds to these inputs, it generates a more structural and "digital" representation of the form within the image.

All the rules for boundary completion just described, as well as the formation of the boundary webs to be described in Section 13, occur within the single box labelled BCS in Fig. 3. Completed BC system signals are fed to the FC system, in which the determination of perceived featural quality occurs through the spread, or filling-in, of feature contour signals within the form-sensitive compartments, or coordinate system, determined by boundary contour signals (Fig. 10). The cellular network in which such filling-in occurs is called a *syncytium*. The importance of the two system hypothesis for surface perception concerns the manner in which the rules for BC system and FC system interactions clarify how we simultaneously perceive a stable surface shape and labile surface qualities. In order to arrive at this insight, we first need to provide more information about how the CC loop can segment a scenic image.

11. TEXTURAL SEGMENTATION AND PREATTENTIVE GROUPING

Grossberg and Mingolla [37] describe how BC system and FC system properties, notably properties of the CC loop, help to explain much of the most difficult data of Beck and his colleagues on emergent features in textural segmentation and grouping [2, 3, 5, 7, 8, 11]. In these phenomena, extremely rapid, pre-attentive groupings of image data into highly context-sensitive emergent features promotes scene segmentation, as exemplified in Fig. 11. The crucial control in displays such as Figs. 11b and c is that overall image contrast is equated between the top and bottom regions, so no simple contrast filtering alone can account for the segmentation.

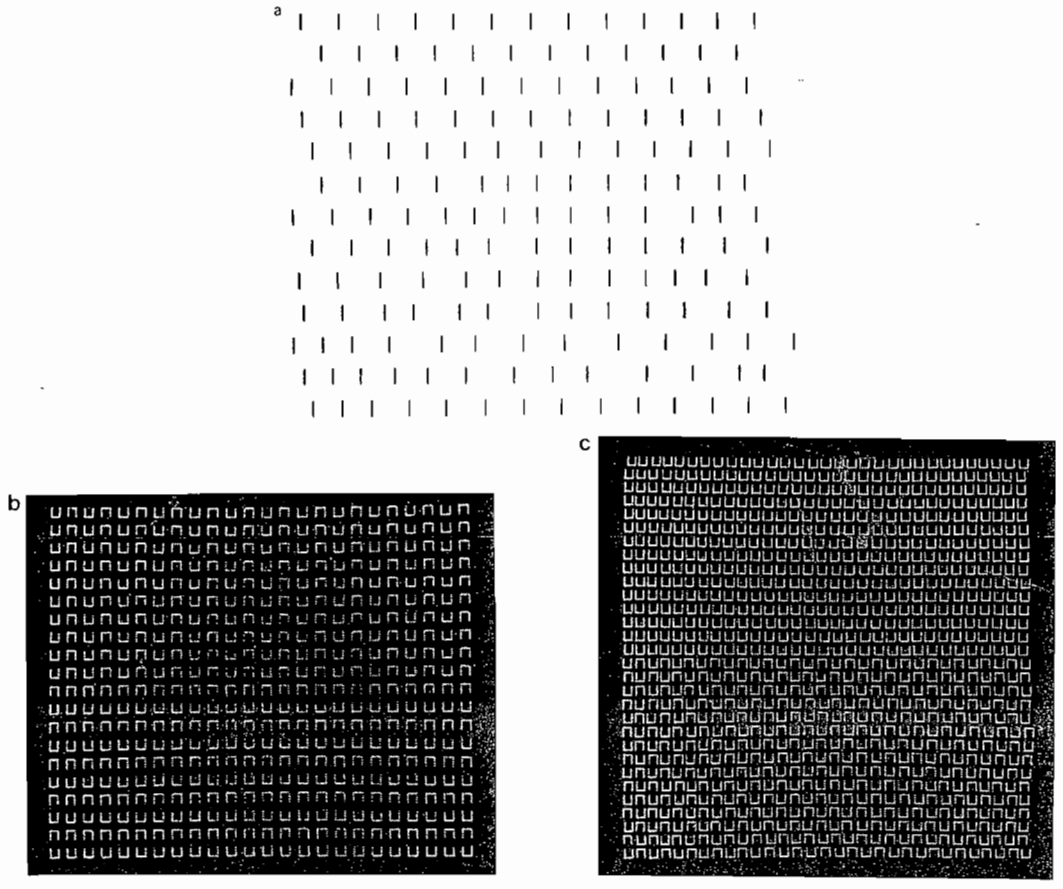

FIG. 11. Illustrations of emergent segmentation: (a) The colinear linking of short line segments into longer segments in response to one region of the image is an "emergent feature" that sustains perceptual grouping. (b) Grouping can also occur in a direction perpendicular to local image contrasts, as happens between the end cuts at the ends of the upright and inverted *U*'s in the bottom half of this figure. (c) The emergent grouping in the lower half of the figure is in diagonal directions, although the only local image contrasts are horizontal and vertical *U* segments. (Reprinted with permission from Beck, Prazdny, and Rosenfeld [11].)

Figure 12 displays some computer simulations illustrating the competence of the BC system to perform the kinds of grouping operations needed to explain the Beck data, as well as other data about textural grouping. Figure 12a depicts an array of nine vertically oriented input clusters. We call each cluster a *Line* because it represents a caricature of how a field of OC filter output cells respond to a vertical line. Figure 12b displays the equilibrium activities of the cells at the second competitive stage of the CC loop in response to these Lines. The length of an oriented line at each position is proportional to the equilibrium activity of a cell whose receptive field is centered at that position with that orientation. The input pattern in Fig. 12a possesses a manifest vertical symmetry: Triples of vertical Lines are colinear in the vertical direction, whereas they are spatially out-of-phase in the horizontal direction. The boundary contour system senses this vertical symmetry, and generates emergent vertical lines in Fig. 12b. The boundary contour system also generates horizontal end cuts at the ends of each Line, which can trap the featural contrasts of each Line within the FC system. Thus the segmentation simultaneously supports a vertical macrostructure and a horizontal microstructure among the Lines.

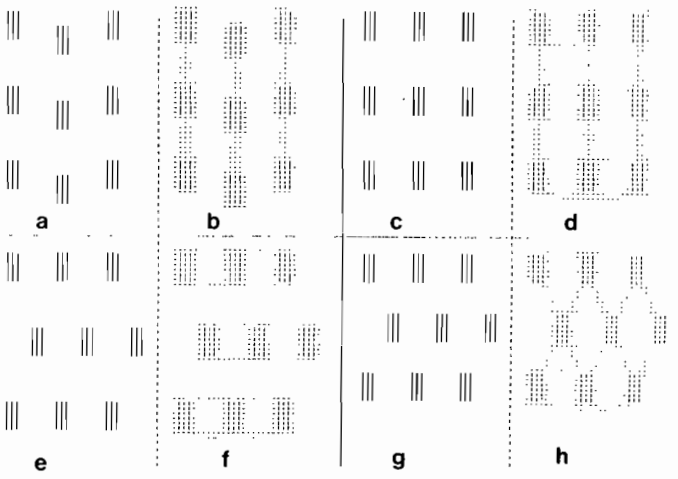

FIG. 12. Computer simulations of processes underlying textural grouping: The length of each line segment is proportional to the activation of a network node responsive to one of twelve possible orientations. Parts (a), (c), (e), and (g) display the activities of oriented cells which input to the CC loop. Parts (b), (d), (f), and (h) display equilibrium activities of oriented cells at the second competitive stage of the boundary contour system. A pairwise comparison of (a) with (b), (c) with (d), and so on indicates the major groupings sensed by the network.

In Fig. 12c the input Lines are moved so that triples of lines are colinear in the vertical direction and their Line ends are lined up in the horizontal direction. Now both vertical and horizontal groupings are generated in Fig. 12d. The segmentation distinguishes between Line ends and the small horizontal inductions that bound the sides of each Line. Only Line ends have enough statistical inertia to activate boundary completion via the CC loop.

In Fig. 12e the input Lines are shifted so that they become non-colinear in a vertical direction, but triples of their Line ends remain aligned. The vertical symmetry of Fig. 12c is hereby broken. Consequently, in Fig. 12f the boundary contour system groups the horizontal Line ends, but not the vertical Lines.

Figure 12h depicts the emergence of diagonal groupings where no diagonals exist in the input pattern. Figure 12g is generated by bringing the three horizontal rows of vertical Lines closer together until their ends lie within the spatial bandwidth of the cooperative interaction. Figure 12h shows that the boundary contour system senses diagonal groupings of the Lines. These diagonal groupings emerge on both microscopic and macroscopic scales. Thus diagonally oriented receptive fields are activated in the emergent boundaries, and these activations, as a whole, group into diagonal bands.

Figure 12 illustrates our claim that the mechanisms of textural segmentation are the same as those of boundary completion, whether of "real" or of "illusory" boundaries. We further claim that precisely the same mechanisms underly the generation of boundary webs for surface perception. Although the *results* of CC loop equilibration in our analysis of textural segmentation were always discrete boundary structures, the capacity for graded responses of units in the BC system have always been present, as indicated in Fig. 4 and Section 6. Even in the Beck displays of Fig. 11, the receptive fields of individual contrast sensitive cells can overlap two or more of the U's, resulting in gradations of activity at many orientations besides the horizontal and vertical orientations of the segments of the

U's themselves. It is precisely the ability of the CC loop to swiftly choose among all the possible groupings latent in these graded outputs from the OC filter that is its chief virtue. The ability of the CC loop to rapidly converge onto one of a very large number of possible stable configurations does not, however, imply that these configurations must necessarily be activated by line-like image contrasts only. In fact, without any change of equations or parameter values, the CC loop can generate dense activations of oriented cell activity in response to smoothly varying image luminance gradients.

Image processing and computer vision algorithms have for years been beset by tradeoffs of purported virtues between region-based algorithms with derived boundaries and edge-based algorithms with derived regions. In contrast, the BC system can simultaneously sustain both "discrete" and "continuous" modes for a single image. The BC system does not need external prompting about which style of processing to adopt for each image region. It does not need a preprocessor which is smart enough to segment regions containing boundaries, textures, and surfaces from one another before specialized algorithms can process these distinct types of image properties. Indeed, many natural images contain mixtures of boundaries, textures, and surface elements in each region of perceptual space, and human observers can benefit from, rather than be confused by, such combinations of image properties [24, 77, 78]. The BC system is also equally at home with several types of image information. Within the BC system, a boundary web can simultaneously encode boundaries, discrete textural elements, and smooth shading into a single form-sensitive network of boundary compartments.

12. RECOGNIZING FORM VS SEEING CONTRAST

How does a discrete boundary structure, no matter how form-sensitive, give rise to a percept of a smooth surface? This possibility also depends upon the identification of distinct BC and FC systems and upon an analysis of how interactions between these systems give rise to conscious percepts within the FC system, but not within the BC system. The key issue can be discerned through a consideration of how "emergent features" can influence *recognition* without necessarily generating a corresponding *visible contrast*.

For example, in the Beck textural segmentation displays of Fig. 11, local contrasts define short line segment features in the image. If certain of the segments are distributed in a regular manner, colinear groupings of these segments can become "emergent features," capable of setting one textural region apart from another. In such displays the colinear arrangement making up emergent features need not be in line with the directions of the local contrasts. They can, for example, be generated by colinearly arranged end cuts that are perpendicular to image contrasts (Fig. 11b).

A remarkable aspect of displays such as Fig. 11 is that we see a series of short lines despite the control of perceptual grouping by the long emergent features. We claim that within the BC system a boundary structure emerges corresponding to the long lines of Fig. 11, as in the computer simulation summarized in Fig. 12b. This structure includes long vertical structures as well as short horizontal components near the endpoints of the short scenic lines. Within the FC system (Fig. 10) these horizontal signals prevent featural filling-in of dark and light contrasts from crossing the boundaries corresponding to the short lines. On the other hand, the output from the BCS to an object recognition system (Fig. 3) reads out a long line

structure without regard to which subsets of this structure will be seen as dark or light.

Unless a connected boundary can be synthesized by the BC system, it cannot separate the activity of the FC system into domains capable of supporting different filled-in color or brightness signals. Thus if a later event can inhibit the cooperation before a connected boundary structure can emerge, no percept may be visible. This insight clarifies how later events can block the percept of earlier events during metacontrast [13, 21, 70]. Even if a connected boundary structure does emerge, boundary signals are *always* invisible within the BC system; visible percepts emerge within the FC system. Thus the contrast-sensitivity of cells within the BC system does not imply visibility of the final percept. Only the generation within the FC system of different filled-in featural contrasts on the opposite sides of a completed boundary (Fig. 10) can lead to a visible perceptual difference.Thus if a boundary separates two FC compartments with equal, or very similar, filled-in contrasts, the boundary may itself be invisible, even though it is necessary to support the contrasts that are perceived.

Thus a spatially dense set of boundary contours can separate the FC system into a set of compartments whose contiguous filled-in contrasts are so similar that the compartmental boundaries are invisible. In response to a continuous image gradient, the spatial density of boundary contour and feature contour signals covary in such a way that the filled-in contrasts of contiguous FC system compartments *are* very similar. In particular, a large flat region of an image may generate a single large compartment whose filled-in contrast is also approximately flat over the entire region. An image region in which a steep reflectance gradient exists may generate very small compartments whose individual filled-in contrasts are different, but due to the small size of the compartments, are very similar to the filled-in contrasts of contiguous compartments. Thus the fact that the BC system generates a set of compartments which is *form-sensitive* enables it to support a filled-in FC system percept, many of whose boundaries are invisible in response to continuously shaded image.

Further argument is needed to explain how visible qualities such as color and brightness seem to adhere to the structure of recognizable 3D form, giving substantial, continuous body to the skeleton of boundary structures. To achieve this insight, we will indicate how featural filling-in can be differentially controlled at several spatial scales in such a way as to give a coherent representation of form-and-color-in-depth.

13. COMPUTER SIMULATION: AN EXAMPLE OF A BOUNDARY WEB

Before introducing multiple spatial scales, we illustrate how image luminance gradients can give rise to an invisible boundary web in a BC system with only a single spatial scale. By a boundary web, we mean a spatially fine mesh of boundary contour activation within the CC loop (Sect. 9). This boundary structure is induced by the oriented receptive fields of the OC filter (Sect. 7). Accordingly, Fig. 13 shows a shaded image of a shiny ellipsoid under somewhat oblique illumination from a point source. Figure 14 shows the equilibrium pattern of activity that is induced within the OC filter by the ellipsoidal image in Fig. 13. All of the oriented masks, or receptive fields, are of the size indicated in Fig. 14. The spatial configuration of all such mask reactions is called an orientation field, or a mask field.

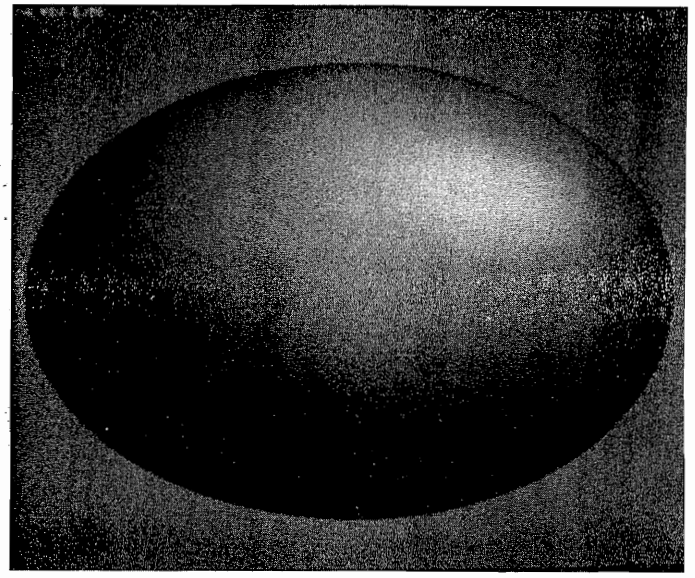

FIG. 13. A computer generated image depicts a shiny ellipsoid under oblique illumination against a gray background.

Boundary webs occur as a result of rapid CC loop equilibration in response to inputs from a mask field of the OC filter. Figures 15–18 show the equilibrium patterns of activity at four subsequent processing stages of the CC loop. Note that according to our theory, *none* of the CC loop activity patterns correspond as such to visible contrasts. Instead, the boundary webs merely describe those regions within which contrast can and cannot spread after BC system signals are fed into the FC

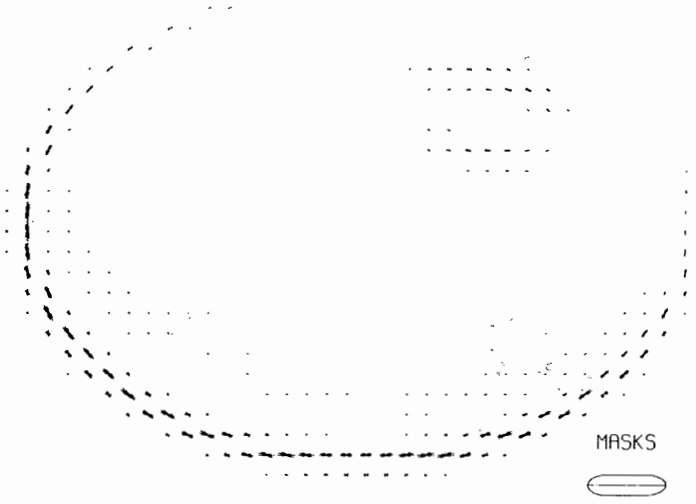

FIG. 14. An orientation field, or mask field, is shown in this lattice of oriented line segments. The length of each segment is proportional to the equilibrium response of an oriented contrast-sensitive mask of the size indicated in the lower right corner. The orientation of each line segment is that of the mask that it codes. This mask field was generated using a Gaussian blurred version of Fig. 13 as input. See Appendix, part A, for a detailed description of mask responses.

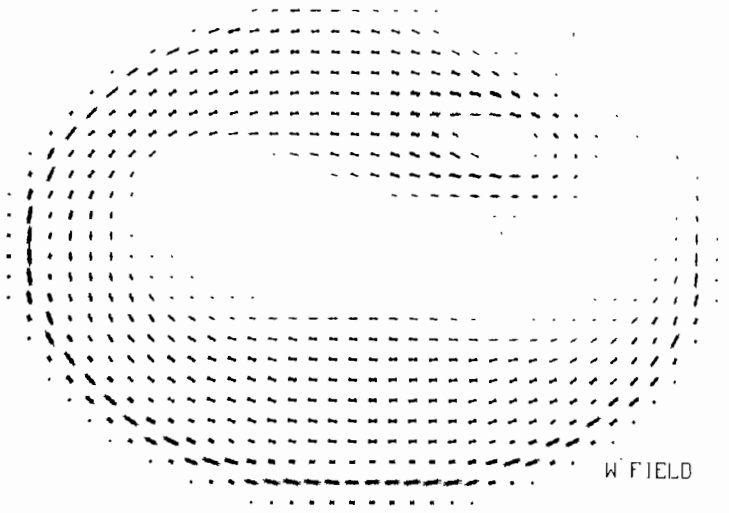

FIG. 15. Figures 15 through 18 depict the equilibrium activities of four subsequent processing stages in response to input from the mask field of Fig. 14. Line segment lengths and orientations code activities of the corresponding network nodes as in Fig. 14. This figure displays activities at the first competitive stage. See Appendix, part B, for a detailed description of this stage.

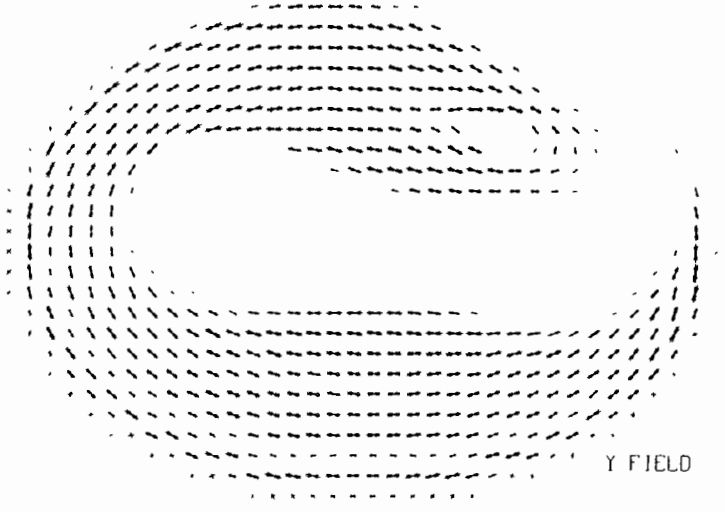

FIG. 16. Equilibrium activities of the second competitive stage in response to the mask field of Fig. 14. See Appendix, part C, for a detailed description of this stage.

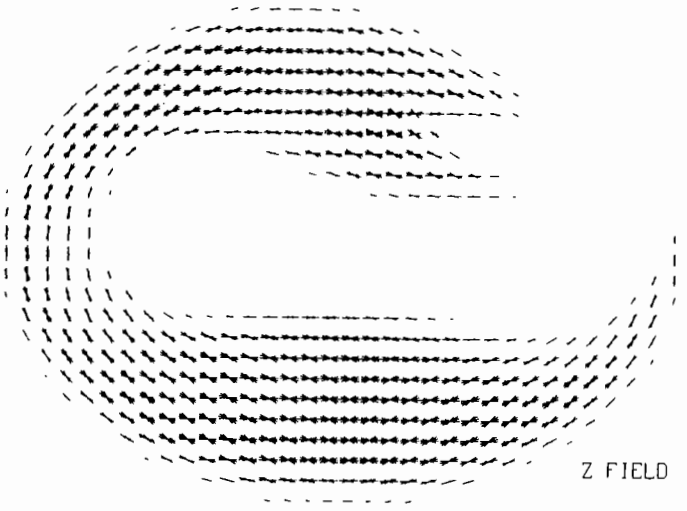

FIG. 17. Equilibrium activities of the cooperative stage in response to the mask field of Fig. 14. See Appendix, parts D and E, for a detailed description of this stage.

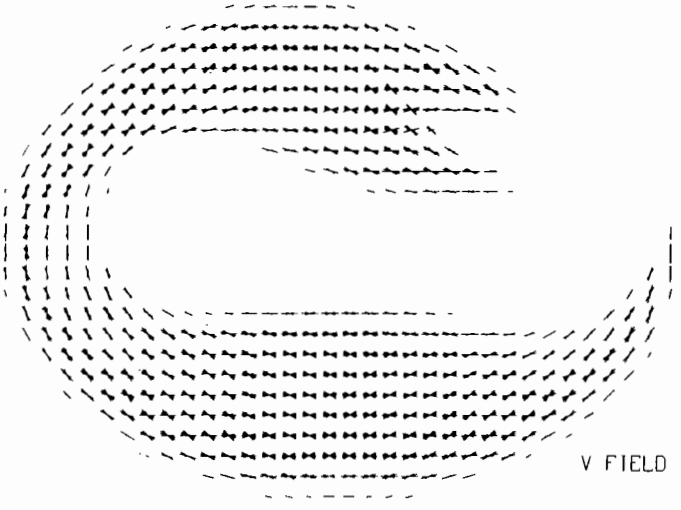

FIG. 18. Equilibrium activities of the feedback stage from the cooperative stage to the first competitive state in response to the mask field in Fig. 14. See appendix, part F, for a detailed description of this stage.

system. The novel ingredient in the boundary web simulations of Figs. 15–18, relative to the textural segmentation simulations of Fig. 12, is that the distributions of activity over nodes of the boundary webs have dense regions as well as line-like structures, while the textural segmentation displays generate only line-like structures. These completed boundary structures, whether discrete or continuous, map topographically into the FC system. Consequently, within the FC system, the boundary structures of the textural simulations would trap brightness discretely on either side of thin boundary contours, whereas the dense regions of boundary webs would trap brightness in local areas over large regions. Note that the signals of the FC system can have different strengths, accounting for the perception of different degrees of brightness even within regions containing essentially identical levels of BC system activation.

14. MULTIPLE SPATIAL SCALES: DISTINGUISHING SIZE FROM DEPTH

Experiments demonstrating the existence of simultaneous binocular fusion and rivalry at different spatial scales [48, 55] underscore the subtlety of issues about spatial scale that are addressed by our theory. Many of these issues pertain to binocular matching of boundaries, and in particular to the perceptual suppression of binocularly mismatched boundaries. While the main goal of this paper does not include discussion of stereopsis, many of the theory's concepts about monocular perception of curved surfaces in depth can best be understood in juxtaposition with related ideas about binocular viewing. To motivate the use of multiple spatial scales in our theory, consider the following gedanken experiment [34, 35]. This gedanken experiment suggests the need for multiple spatial scales, such that only those scales capable of supporting a match can be allowed to generate a visible percept. This experiment can also be phrased in terms of the fixation process.

As a rigid object approaches an observer, the binocular disparities between its nonfixated features increase proportionally. In order to maintain the fixation process and to achieve a percept of object permanence, mechanisms capable of correlating these progressively larger disparities are needed. Other things being equal, the largest disparities will lie at the most peripheral points on the retina. The cortical magnification factor, whereby cortical regions of fixed size process larger retinal regions as a function of retinal eccentricity, is one mechanism whereby this is accomplished [42, 73].

It is not sufficient, however, for a single spatial scale to exist at each retinal position, such that scale size increases with retinal eccentricity. This is because objects of different size can approach an observer. The observer can confuse object size with object depth unless multiple scales exist corresponding to each retinal position. In particular, objects of different sizes can generate the same monocular retinal image if they lie at different distances from an observer, with larger objects further away. If these objects possess spatially uniform interiors, then the boundary disparities of their paired retinal images carry information about their depth. Because all the objects are at different depths, they all generate different disparities. Within the range of accurate depth perception, all of these disparities need to be computable with respect to the fixed retinal positions in one eye that are excited by all the objects' boundaries. Multiple spatial scales corresponding to each retinal position can carry out these multiple disparity computations.

This gedanken experiment suggests the functional utility of suppressing percepts corresponding to binocularly mismatched boundaries. Each monocular image can excite more spatial scales corresponding to each retinal position than can binocularly match. Only the binocularly matched boundaries provide correct information concerning form-in-depth. Consequently, the mismatched boundaries must be prevented from generating a conscious percept.

The gedanken experiment also clarifies the utility of allowing certain lower spatial frequencies to match and be fused, whereas mismatched higher spatial frequencies are suppressed or rivalrous, when two images are binocularly viewed at a fixed disparity. As an object approaches an observer, the sizes of its monocular retinal images and their binocular disparities increase together. Other things being equal, larger spatial scales should therefore be able to match pairs of images with larger disparities.

15. MONOCULAR SELF-MATCHES: GRADIENT DEPTH AND MOTION DEPTH

In order to clarify the relationship of binocular perception to multiple scale monocular perception, it is useful to recall that when an observer closes one eye, vivid perception is still possible. Moreover, monocular percepts can retain a significant impression of depth. Thus a binocular match within the boundary contour system is not necessary to generate a conscious percept, let alone a depthful percept. What is needed is the absence of a binocular mismatch. Because the visual world can vividly be perceived through a single eye, certain boundary contour system cells that project to the cell syncytium must be capable of being monocularly activated. We call such activation a monocular self-match to distinguish it from a binocular match [34].

In the absence of binocular mismatches, more monocular self-matches can occur than under binocular viewing conditions. This property helps to explain why, when viewed under reduction conditions (one eye looks through a small aperture in dim light), depth percepts can be ambiguous [29–31]. On the other hand, the existence of more monocular self-matches raises the question of why depth is ever perceived under monocular viewing conditions? One factor is the correlation between scale size and disparity that was mentioned in Section 14. Larger scales can, other things being equal, preferentially respond to larger image elements. Under binocular viewing conditions, larger images are often closer and generate larger disparities. Due to the preferential response of larger scales to large image elements, a monocularly viewed image which contains spatial gradients (Fig. 13) can be parsed among multiple spatial scales in a manner similar to its parsing during binocular viewing conditions. Gibson [22] is notable among classical perceptual theorists for his many illustrations of how spatial gradients can influence depth perception, whether the spatial gradients are continuous as in Fig. 13 or discrete as in Fig. 19.

A monocularly viewed moving object can activate a succession of monocular self-matches which are capable of matching or mismatching previous self-matches before they can decay. Temporally staggered pairs of monocular self-matches can hereby generate "binocular" matches or mismatches across the multiple spatial scales, thereby strengthening the percept of depth.

The present theory suggests, more generally, that any visual operations which cause equivalent activations across the multiple scale boundary contour computations that input to the cell syncytium will generate equivalent depth percepts,

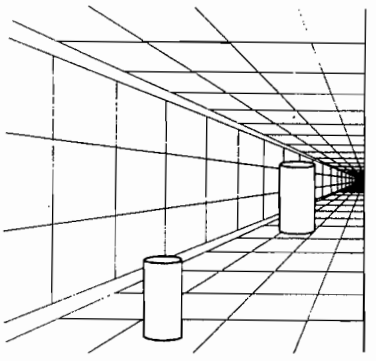

FIG. 19. The corridor illusion: The phenomenally nearer cylinder appears smaller than the phenomenally further cylinder due to the influence of the convergent gradient of lines. (Reprinted with permission from Kaufman [48].)

whether they be due to monocular spatial gradients, monocular motion cues, binocular disparities, or top-down "cognitive contours." This tenet of the theory is of such importance that we call it the *principle of scale equivalence* [34]. The principle of scale equivalence clarifies why conclusions about multiple scale boundary processing that are derived from an analysis of binocular perception can be used to analyze monocular processing of a surface image.

16. CONTINUOUS MODULATION OF MULTIPLE SCALE ACTIVITY

As an observer moves within a scene, the scenic forms and their depths seem to change continuously. By contrast, at most a finite number of spatial scales can exist in the brain. In many contemporary models of binocular depth perception, depth jumps discretely between a few values as an observer moves about. Moreover, these binocular models do not explain how the computation of disparity values leads to a percept of form-and-color-in-depth. Grossberg [34] summarizes and analyzes a number of these models.

The present theory suggests that multiple spatial scales exist within the boundary contour system and that these scales can be simultaneously activated by a monocularly or binocularly viewed scene, albeit by different amounts corresponding to different scenic positions. As an observer moves about a scene, the relative and absolute degree to which each of these multiple boundary scales is activated changes. These changes in the energy balance across multiple boundary scales alters the ability of the corresponding feature contour signals to generate a visible percept. We suggest that multiple feature contour syncytia exist corresponding to the multiple boundary contour system spatial scales, each syncytium capable of contributing to a visible percept, but to different degrees (Fig. 10). In the limit wherein no boundary contour signal within a given spatial scale excites the corresponding feature contour syncytium, that feature contour syncytium cannot contribute to a visible percept, as approximately occurs during binocular rivalry.

This type of multiple scale concept supports the strong kernel of truth that exists within the Fourier theory of spatial perception [32, 33], but also replaces the Fourier theory by one with a greater explanatory range.

17. SURFACE CURVATURE AND MULTIPLE SYNCYTIAL SCALES: FILLING-IN AMBIGUOUS REGIONS

Another basic role for multiple scale featural filling-in can be appreciated from the following gedanken experiment. When both eyes focus on a single-point within a patterned planar surface viewed in depth, the fixation point is a point of zero disparity. Points increasingly far from the fixation point have increasingly large binocular disparities. Why does such 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? If the relative sizes of boundary disparities contribute to relative depth percepts, then how do we ever perceive planar surfaces? How do we even perceive rigid surfaces?

The severity of this problem is further indicated by the fact that perceived depth can, under certain circumstances, depend upon the choice of fixation point. Starting at one point in a Julesz stereogram can result in a gradual loss of depth [48]. Also, in a stereogram composed of three vertical lines to the left eye and just the two outmost lines to the right eye, the depth of the middle line depends upon whether the left line or the right line is fixated [48]. If depth can depend on the fixation point when discrete lines are viewed, then why do not observers perceive planar surfaces as being highly curved? What is the crucial difference between the way we perceive the depths of curves and of surfaces?

These examples raise the fundamental issue of how an observer knows that a planar surface is being viewed, not just whether the observer can estimate the depths of some parts of the surface. Moreover, when a homogeneous planar surface is being viewed, it is not possible to compute any unambiguous disparity computation within the interior of the plane. Determining that such a surface is planar thus cannot just be a matter of showing that the same disparity can be computed at all interior points of the surface. Somehow locations where unambiguous depth computations can be performed can generate a filling-in reaction which enables nearby regions to inherit these depth values. Such a filling-in reaction is not, moreover, just a matter of filling-in disparity values, as a number of authors have proposed [17, 44, 59, 74], because the featural qualities, such as brightness and color, of the percept inherit these depth values.

A similar point is made through the study of stereograms. Julesz [44, p. 336] has, for example, constructed a stereogram in which a "figure" and its "ground' are both indicated in the left image of the stereogram by a 5% density of randomly placed black dots on a white background. The "figure" is indicated in the right image of the stereogram by shifting the positions of the black dots that fall within the figure relative to the positions of the corresponding black dots in the left image, while placing the dots indicating the "ground" in the same positions as those in the left image. When this pair of images is stereoscopically viewed, the whole figure, including the entire 95% of white image between its black dots, seems to hover at a different depth than the ground. How does the white region of the "figure" inherit the depth quality arising from the disparities of its meagerly distributed black dots, and the white region of the "ground" inherit the distinct depth quality due to its black dots? What mechanism organizes the locally ambiguous white patches that dominate 95% of the pictorial area into two unambiguous white regions hovering at distinct perceived depths?

The present theory suggests that when a particular boundary contour scale is strongly activated by a given scene, this boundary contour activation can trigger a strong filling-in reaction within the corresponding syncytium of the feature contour system. By definition, featural filling-in within a syncytium is restricted to the spatial scale in which this syncytium resides. We suggest that a surface percept will appear flat if it is generated by a pair of bounding boundary contours within a single spatial scale, because featural filling-in is restricted to the single syncytial scale that these boundary contours activate. We trace the perceived curvature of a nonplanar surface to a multiple scale boundary contour reaction which causes the distribution of filled-in featural activity to be "curved" among several syncytial scales as perceptual space is traversed.

This explanation of perceived surface flatness and curvature suggests that feature contour signals corresponding to a fixed retinal position send inputs to the filling-in syncytia of *all* the multiple spatial scales. Only those scales that also receive boundary contour signals can, however, convert these feature contour signals into visible percepts. One of the fundamental tasks of our binocular theory is to explain how such an interaction between boundary contour signals and feature contour signals can convert some, but not all, of the feature contour signals into filled-in percepts. Grossberg [35] discusses model processes whereby this can be accomplished in greater detail. Although such an analysis goes beyond the scope of this article, a number of general conclusions can be drawn.

18. THE $2\frac{1}{2}$D SKETCH DOES NOT EXIST

For example, the above considerations suggest that the hypothesis that a $2\frac{1}{2}$D sketch exists, distinct from a full 3D representation, is invalid [58]. Such a $2\frac{1}{2}$D sketch is an "orientation and depth map of the visible surfaces around a viewer" [60, p. 306]. In contrast, the above considerations suggest that "ambiguous" regions of a scene, whose positions do not possess their own boundary contours, derive a relative depth value from the energy balance of their filled-in featural activities across all the spatial scales at that position. In other words, a depth map is completed by the featural filling-in process which generates a full 3D representation of form-and-color-in-depth.

19. INTERACTIONS BETWEEN BRIGHTNESS AND DEPTH INFORMATION

This conclusion implies, in particular, that brightness and depth information can mutually influence one another. Grossberg [34] reviews classical data that support this assertion. Three types of data are summarized in this section for illustrative purposes.

Kaufman, Bacon, and Barroso [49] studied stereograms built up from the two monocular pictures shown in Fig. 20a. When these pictures are fused through a stereogram, the two lines are perceived at a different depths due to the disparity between the two monocular views. If the stereogram is changed so that the left eye sees the same picture as before, whereas the right eye sees the two pictures superimposed (Fig. 20b), then depth is still perceived. If both eyes view the same superimposed pictures, then no depth is seen. However, if one eye sees the pictures superimposed with equal luminance, whereas the other eye sees the two pictures superimposed each with different luminance, then depth is again seen. In this last example, there is no disparity between the two figures, although there is a

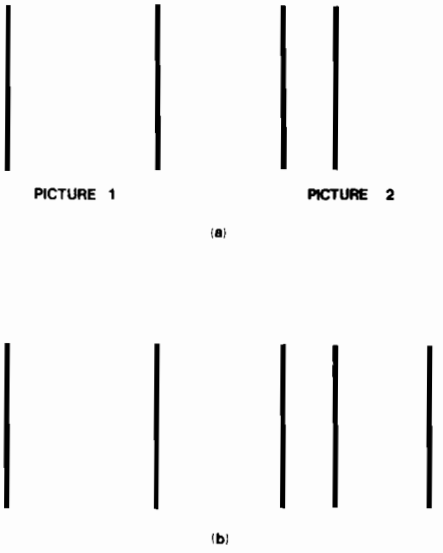

FIG. 20. Combinations of the two pictures in (a), such as in (b), yield a depth percept when each picture is viewed through a separate eye. Depth can be seen even if the two pictures are combined to yield brightness differences but no disparity differences. (Reprinted with permission from Kaufman [48].)

luminance difference. Thus a luminance difference *per se* can induce a depth percept.

Schwartz and Sperling [72] and Dosher, Sperling, and Wurst [18] have further analyzed influences of luminance differences upon both perceived depth and perceived rigidity of form. In their studies they consider *proximity-luminance covariance* (PLC) as a factor influencing percepts of depth and form. To manipulate PLC, the luminance of each line in a 2D projection of an object was made to depend on the 3D depth of that line. A larger luminance was used to signal a closer object projection. PLC's that confirm and that conflict with the 3D depth were analyzed. The interaction of PLC with stereo information was also analyzed. Both studies concluded that PLC is a powerful factor in determining a depth percept and that, moreover, stereo and PLC information combine in a way that can be summarized by a weighted linear model.

Egusa [19] has studied the effects of luminance differences on perceived depth by constructing stimuli consisting of two hemifields of different colors, and asking the subject to state which appeared nearer and to judge the perceived depth between them. When both hemifields were achromatic, the perceived depth increased with increasing brightness difference. With chromatic–chromatic combinations, the perceived depth depended upon the hue combination. In terms of decreasing frequency of "nearer" judgments, the hue order was red, green, and blue.

Thus there seems to exist a well-replicated effect of brightness and color on perceived depth, as is consistent with our conception that relative depth information is carried by the energetic loading of filled-in featural contrasts across the FC system syncytia corresponding to multiple BC system spatial scales.

20. TRANSPARENCY

Transparency phenomena [10, 61–63] provide another type of data which support the concept that multiple syncytia exist corresponding to different spatial scales, and

that filling-in of feature contour signals within some of these syncytia but not others can generate a percept of color-and-form-in-depth. In percepts of transparency, a phenomenal scission occurs which replaces the percept of a single color at a fixed perceptual location with the simultaneous perception of two colors: the color of the object seen through the transparency and the color of the transparent layer. Within our theory, such a scission is analyzed by considering how one color elicits filling-in within a syncytium of one spatial scale, whereas the other color elicits filling-in within a syncytium of a different spatial scale. Such an analysis is possible due to the hypothesis that each FC signal is topographically broadcast to the syncytia of all spatial scales, and that the spatial distribution of BC signals among the several scales determines which of the syncytia will react to such an FC signal by triggering a filling-in reaction.

The formal rules which Metelli [61] is articulated for predicting the occurrence of transparency are similar to the conditions under which the BC system triggers neon color spreading [68, 79, 80] within the FC System. Meyer and Senecal [63] studied a variant of the Kanizsa [45, 46] subjective contour configuration. Unlike Fig. 2, some of the pac-man figures which they used to induce a rectangular subjective contour were completed in the image using faintly colored wedge shaped regions. Meyer and Senecal [63] showed that a percept of transparency covaries with the percept of a chromatically filled-in rectangle surrounded by a strong rectangular subjective contour. In our explanation of neon color spreading [36], the strength of such a subjective contour also regulates the strength of the chromatic filling-in reactions by inhibiting boundaries that would otherwise prevent filling-in from escaping from the colored inducing regions. Thus our theory suggests that certain instances of transparency are due to featural filling-in reactions across some, but not all, of the syncytia corresponding to each fixed perceptual location, and that these differential filling-in reactions are associated with a perception of a relative difference in depth.

21. COMPUTER SIMULATION: NESTING OF BOUNDARY WEBS ACROSS SPATIAL SCALES

With this background in mind of how BC signals can interact with FC signals to generate a completed percept of form, we now illustrate several key properties of boundary webs through computer simulations.

Figures 14–18 showed how a single spatial scale of the OC filter and CC loop react to the shiny ellipsoid in Fig. 13. Figures 21 and 22 show how a smaller spatial scale reacts to the same image. Figure 21 depicts the equilibrium activity pattern of the orientation field, as in Fig. 14. Figure 22 depicts the equilibrium activity pattern of the second competitive stage, as in Fig. 16. (See the Appendix, part B, for a mathematical definition of this stage.)

Figures 23 and 24 show how a larger spatial scale than that used in Figs. 14–18 reacts to the shiny ellipsoid in Fig. 13. Figure 23 describes the equilibrium activity pattern of the orientation field. Figure 24 describes the equilibrium activity pattern of the second competitive stage.

In these simulations, "spatial scale" can be defined as the size of the oriented masks used by the OC filter. The primary fact of interest is seen by comparing the orientation fields in Figs. 14, 21, and 23 and the second competitive stage in Figs. 16, 22, and 24. This comparison shows that, in all scales, there exists a dense band of boundary web activity bounded by the exterior contour of the ellipsoid. The

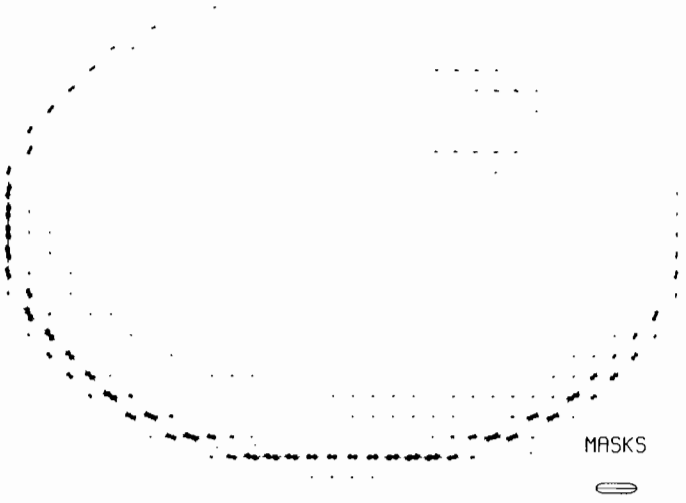

FIG. 21. This mask field is also based on the shaded image in Fig. 13, but for a smaller set of masks (indicated on the lower right) than was used for Fig. 14.

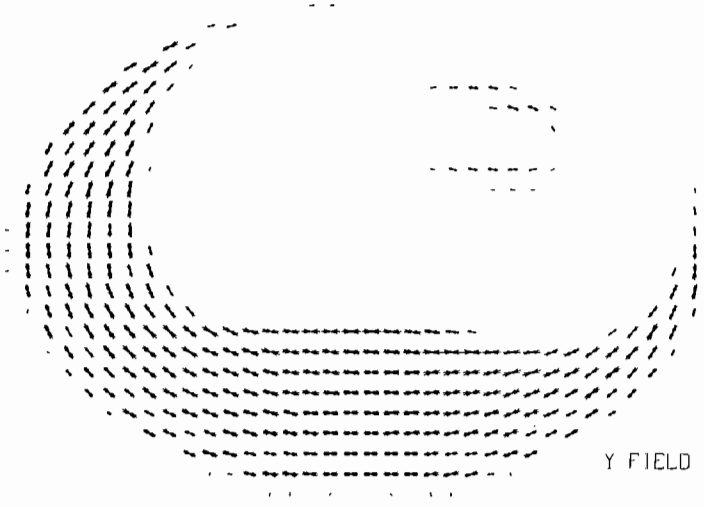

FIG. 22. Equilibrium activities of the second competitive state in response to the input of the masks field shown in Fig. 21. Compare this with Fig. 16.

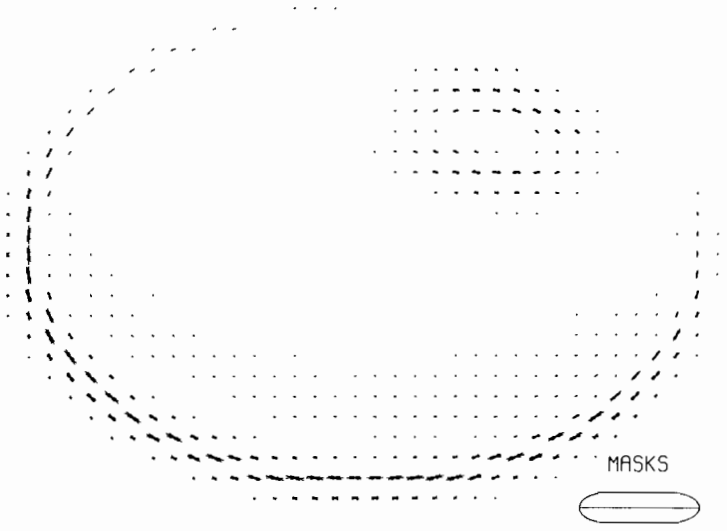

FIG. 23. This mask field is also based on Fig. 13, but for a larger set of masks (indicated on the lower right) than was used for Fig. 14.

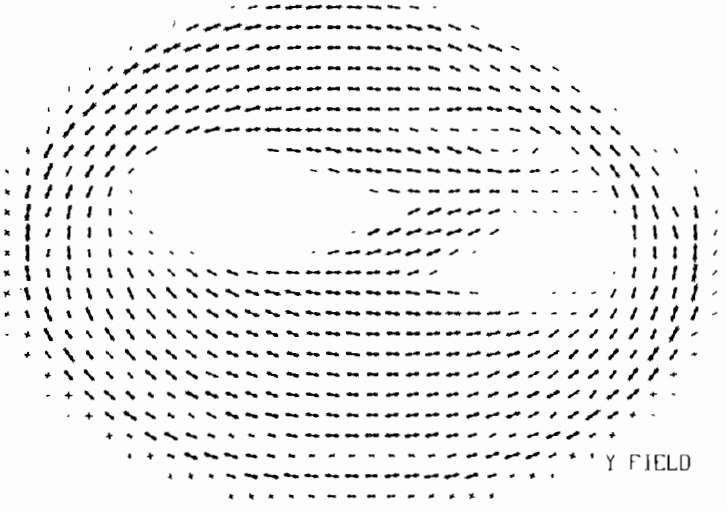

FIG. 24. Equilibrium activities of the second competitive stage in response to the orientation field of Fig. 23. Compare with Figs. 16 and 22.

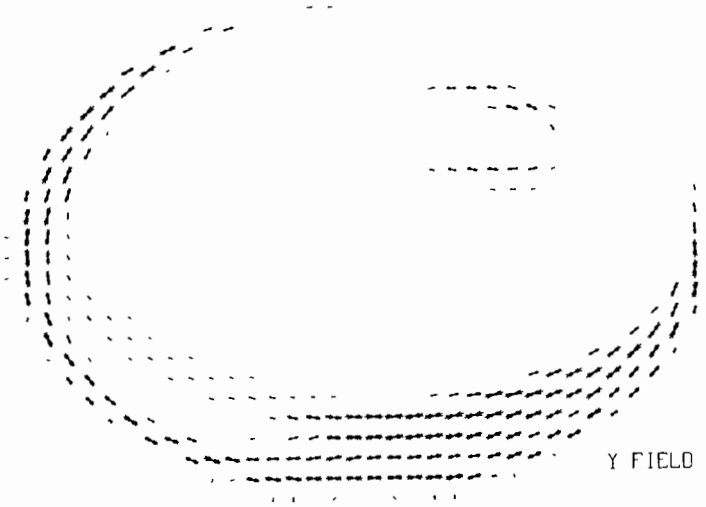

FIG. 25. Equilibrium activities of the second competitive stage in response to the mask field of Fig. 21. This figure differs from Fig. 22 in that the spatial scale at which cooperative feedback occurs is smaller than for Fig. 22. The spatial gradients that are synthesized again follow the isophotes of the image.

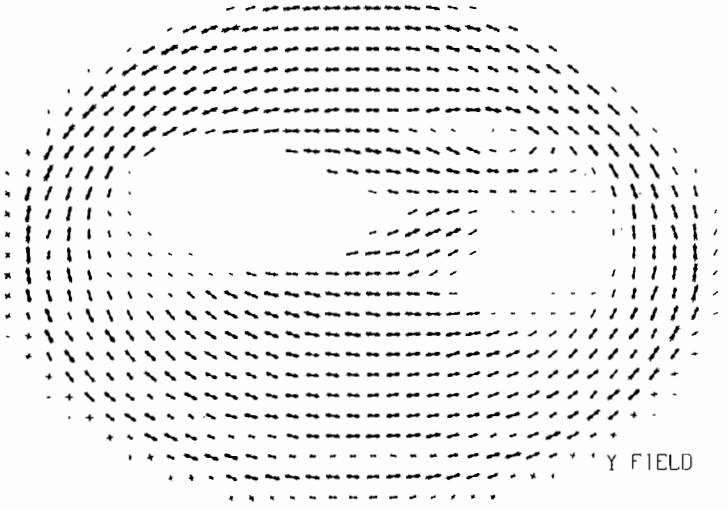

FIG. 26. Equilibrium activities of the second competitive stage in response to the mask field of Fig. 23. This figure differs from Fig. 24 in that the spatial scale at which cooperative feedback occurs is larger than for Fig. 24. Unlike Figs. 22 and 25, the larger cooperative bandwidth does not cause a qualitative change in grouping because the larger masks can detect isophotes at a dense set of image positions.

orientations of all the boundary webs also tend to be parallel to the isophotes of the ellipsoidal image. On the other hand, larger input masks generate broader bands of boundary web activity than smaller input masks. This *nesting property* of boundary web structures across multiple spatial scales is, we claim, on of the key sources of information for deriving 3D shape from monocularly viewed luminance gradients of a 2D image.

Figures 25 and 26 illustrate a more sophisticated concept of spatial scale. In these simulations, the smaller and larger input masks of Figs. 21 and 23 were again employed. In addition, the spatial bandwidths of the cooperative interactions within the CC loop were also altered in a proportional, or self-similar, manner. That is, the cooperative feedback in Fig. 25 operated at a shorter spatial range than in Fig. 22, whereas the cooperative feedback in Fig. 26 operated at a longer spatial range than in Fig. 24. This covariation of input and cooperative bandwidths again reproduces the nesting of boundary webs across spatial scales. Such a self-similar covariation of cooperative bandwidth with input bandwidth realizes the plausible idea that inputs which are averaged by input masks over a larger image domain can afford to group over a larger image domain.

22. COMPUTER SIMULATION: COHERENT COMPLETION OF BOUNDARY WEBS ACROSS INCOMPLETE BOUNDARY DATA

One of the basic properties of the CC loop is that its nonlinear feedback mechanisms transform analog inputs from the OC filter into coherent structures with hysteretic properties. In addition, the long-range cooperative interactions of the CC loop can complete boundaries over perceptual regions whose image contrasts have been attenuated, as across the retinal veins or the retinal blind spot [50].

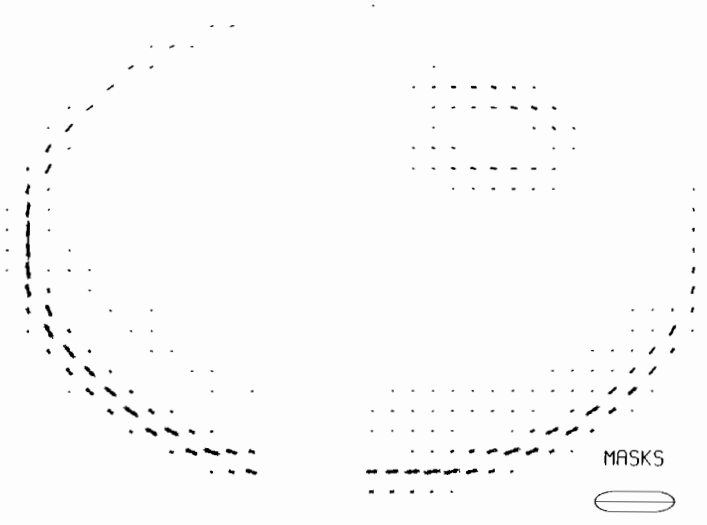

FIG. 27. The mask field shown in Fig. 21 has been altered by the deletion of activities in five lattice columns as an idealization of disruption of visual input by a retinal vein.

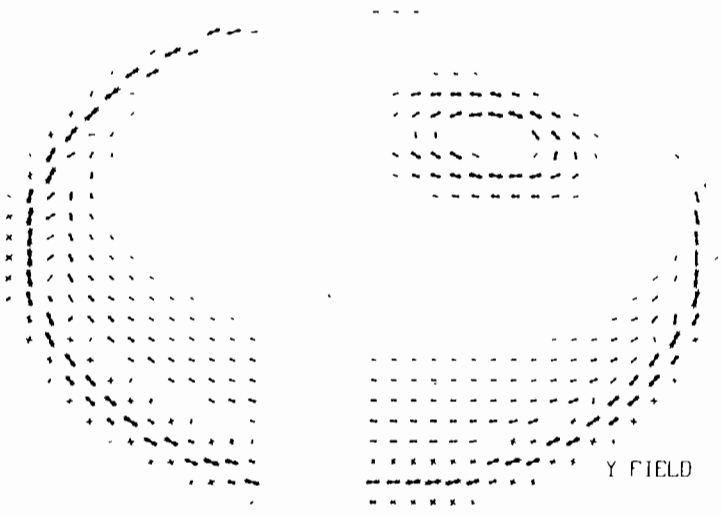

FIG. 28. Equilibrium activities of the second competitive stage in an "open loop" simulation in response to the inputs shown in Fig. 27. Unlike simulations shown in previous figures, cooperative feedback was omitted in this simulation. As a result, the gap in the mask field of Fig. 27 is reflected in the second competitive stage.

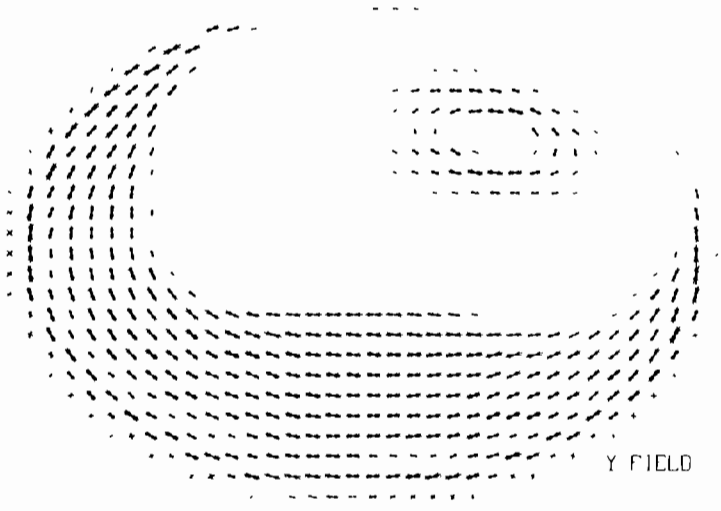

FIG. 29. Equilibrium activities of the second competitive stage in response to the mask field of Fig. 27 in the normal, "closed loop" condition. Unlike Fig. 28, the equilibrium activity pattern in Fig. 29 completes a boundary web over the gap in the mask field of Fig. 27.

We illustrate how such coherent completion can occur in a boundary web by deleting a vertical band of activations at the orientation field that is generated by the shiny ellipsoid in Fig. 13. Such an excised orientation field is depicted in Fig. 27. In order to describe the effects of cooperative feedback upon boundary web completion, we feed this orientation field into both an *open loop* and a *closed loop* version of the CC loop. In the open loop case (Fig. 28), the gap is not closed because cooperative feedback cannot be used to bridge the gap in the input mask contrasts. In the closed loop, or normal, case (Fig. 29), the gap is closed nicely, as would be desired if the image were obstructed by a retinal vein or blind spot. In fact, the completed boundary web is almost indistinguishable from the boundary web that is due to direct activation by a complete orientation field (Fig. 16). Comparison of Figs. 28 and 29 shows, moreover, how the nonlinear feedback within the CC loop smoothes and renders more symmetrical the global configuration of the boundary web, a process consistent with the data of Mingolla and Todd [64].

23. HIGHLIGHTS AND THE PERCEPTION OF GLOSSINESS

The completion properties described in Fig. 29 illustrate how long-range statistical properties of image contrasts can interpolate over weaker short-range properties of image contrasts. Indeed, a colinear grouping due to long-range cooperation can feed back to second competitive stage to inhibit weaker short range effects via orientational competition (Sect. 7). Using such mechanisms, one can imagine examples in which the CC loop obliterates the OC filter activations due to a highlight in certain spatial scales, notably the large spatial scales whose cooperative cells can average contrast gradients over a large portion of the image. Remarkably, the CC loop remains sensitive to structural properties, such as isophotes, despite the existence of interactions which are broad enough to overcome local image distortions. The active dynamical behavior of the CC loop thus stands in stark contract to that of a low spatial frequency filter or other passive smoothing operator.

The postulated ability of a boundary web of the CC loop to fracture activations of the OC filter by a surface highlight is supported by data of Beck and Prazdny [9]. In this investigation, artificial highlights were attached to digitized images of vases. Beck [5] had earlier published pictures of a vase with and without highlights. The vase with highlights attached looked glossy. A sharply bounded white dot attached to a vase would not make it look glossy. A highlight, in contrast, decreases continuously in intensity from its center to its borders.

The data of Beck and Prazdny [9] are consistent with the hypothesis that, in some spatial scales, a boundary web can break through the boundary contours induced by the highlight and can trigger a filling-in reaction that spreads within the boundary web compartments of these spatial scales. Thus, Beck and Prazdny [9] showed that each highlight makes its surrounding area look shiny, but not the whole vase. A highlight attached to a uniform gray image of a vase does not produce the perception of glossiness. Correspondingly, such an image would not generate a boundary web through the highlight. The perception of glossiness depends not only upon the presence of highlights, but also on the intensity gradient on the surface. We claim that such an intensity gradient is needed to induce a boundary web that can penetrate the highlight. The exact form of the gradient is not important. What is important is that the intensity gradient produces the perception of a curved surface.

In our theory, an intensity gradient is needed to generate a boundary web that can coherently span the interior regions of the vase where the highlight is placed.

The luminance profile of a highlight points to another possible source of glossiness in some spatial scales. Certain highlights may generate more boundary contour activation near their centers than near their borders due to the continuous decrease in intensity from center to border. Consequently, in these spatial scales, the feature contour signals induced by these highlights can fall *exterior* to these boundaries, in the sense that they can fill-in their syncytia until they hit the nearest boundaries formed by the figure to which the highlight is attached.

Whether a spread of glossiness is due to fracture of a highlight's local structure by a more global grouping or due to spread from a highlight's center boundary, the resulting spread conforms to the constraints imposed by the boundary structure within that scale. Consequently, the highlight tends to look like it is attached to the surface.

24. SHAPE FROM TEXTURE GRADIENTS

In many natural scenes, smooth shading and textural variations occur concurrently. This fact raises the issue of whether mechanisms competent to represent smooth gradients are also able to deal with gradients built out of discrete textural elements. The experiments of Todd and Akerstrom [77] investigate how a percept of 3D form is generated by texture gradients in a 2D image. Todd and Akerstrom [77] and Todd [76] have analyzed how our concept of a multiple scale boundary web accounts for their data. Some of their main observations are now summarized.

A striking aspect of displays depicting surfaces using gradients of discrete texture elements is how the phenomenal impression of shape is smoothly distributed over and even between the discrete elements [22]. We have already demonstrated how a dense form-sensitive boundary web can be generated by a smooth luminance gradient. This fact raises the basic question: how smooth is smooth? The question can be answered by analyzing the response of the OC filter to changes in the spatial grain of the image. As emphasized in Section 6, the input masks of the OC filter are oriented local contrast detectors, not edge detectors. In particular, such oriented masks can respond to spatially nonuniform densities of textural elements, such as dots, rectangles, etc.

The basic issue is whether or not the densities and orientations of these elements can generate statistically coherent form-sensitive groupings within the CC loop. Clearly, a sufficiently small perturbation off a smooth luminance gradient, say by replacing a smooth image with densely packed small squares whose luminance equals the average luminance of the area which they replace, will not appreciably change the CC loop response. Thus 2D images which generate a percept of 3D form using texture gradients provide a powerful probe of the spatial scale for averaging and grouping within the BC system.

Figure 30 displays a textured ellipsoid surface, viewed along its longest axis. Note that the outer portion of this figure give the impression of circular bands, much like those that occur in the Glass pattern of Fig. 31. In a series of experiments on such figures, Todd and Akerstrom [77] perturbed various aspects of texture—including projected element area, amount of compression (the foreshortening of obliquely viewed elements), and element orientation in the picture plane—in an effort to isolate the conditions affecting perceived curvature or depth. An example of such a

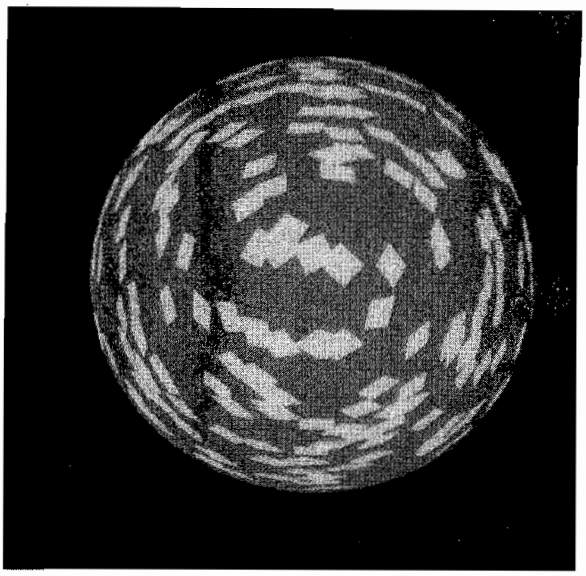

FIG. 30. An ellipsoid viewed along its longest axis of symmetry is depicted through a distribution of texture computed in accordance with the constraints of perspective. (Adapted with permission from Todd and Akerstrom [77].)

perturbation is shown in Fig. 32, in which the elements have appropriately computed areas but where compression has been held constant. This figure gives little or no impression of depth. After conversation with the present authors, Todd and Akerstrom reasoned that according to our theory, this effect could be traced to the absence of preferred orientations in the periphery of the ellipsoid. That is, the constant compression prevents the texture elements from appearing elongated and aligned with the curved shape of the ellipsoid.

In terms of the model, the distribution of activated input masks in response to Fig. 32 does not favor any given orientation. This does not prevent input masks that are activated by edges of individual textural elements from generating large CC loop

FIG. 31. In a Glass pattern, circular groupings emerge from the spatial distribution of small black dots. Unlike Fig. 30, this figure does not induce an impression of surface curvature or depth. (Reprinted with permission from Glass [28].)

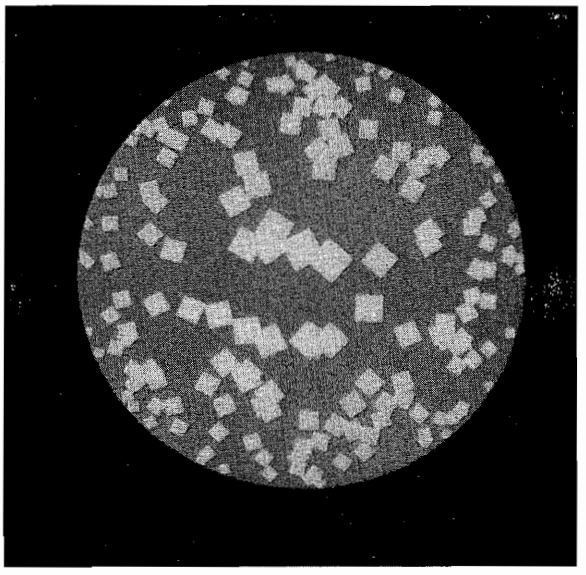

FIG. 32. The ellipsoid in this figure is the same shape as that in Fig. 30 but the elongation due to compression of texture elements has been suppressed, resulting in a much diminished impression of surface curvature or depth. (Adapted with permission from Todd and Akerstrom [77].)

responses. On the other hand, cooperative groupings across individual textural elements tend to get inhibited at the second competitive stage by orientational competition due to other cooperative groupings. Thus the absence of preferred orientations in the ellipsoid enables orientational competition to attenuate cooperative activity between the individual textural elements.

Accordingly, Todd and Akerstrom reasoned that a strong prediction follows from the model: An image generated equivalently to that of Figure 32 (i.e., constant compression) but having elongated texture elements ought to produce a vivid impression of depth, provided that the direction of elongation tends to parallel the circular symmetry of the figure. More technically, the direction of elongation for each element is the same as would have occurred if appropriate compression values had been computed, as for Fig. 30. In the model, the favored directions of elongation cause enhanced activations of the corresponding orientations at the second competitive stage, and thereby enable the CC loop to form strong oriented bands of cooperative activity between individual textural elements.

To test this hypothesis, Todd and Akerstrom generated displays such as shown in Fig. 33. Their psychophysical procedure strikingly confirmed this prediction, which is consistent with no other proposed model perception of curvature from texture. In fact, Todd and Akerstrom [77] used the model to devise a measure of the distribution of element elongation over several spatial scales. The lowest correlation of this measure with subjects' impression of depth in the various conditions of their study was 0.985.

Figure 34 depicts a simulation of an orientation field in response to the image in Fig. 30. The size of an individual oriented mask relative to the scale of the image is depicted in the figure. Note that activity is irregularly distributed across several orientations. Figure 35 depicts the response at the second competitive stage of the

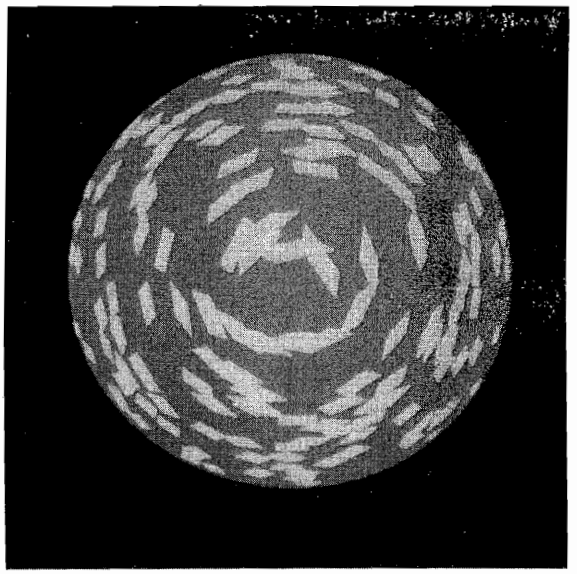

FIG. 33. The same ellipsoid shape used in Figs. 30 and 32 is depicted. This time, although compression of elements has been suppressed as in Fig. 32, each element is elongated and aligned with the local orientation of the underlying depicted surfaces resulting in an impression of curvature or depth close to that of Fig. 30. (Adapted with permission from Todd and Akerstrom [77].)

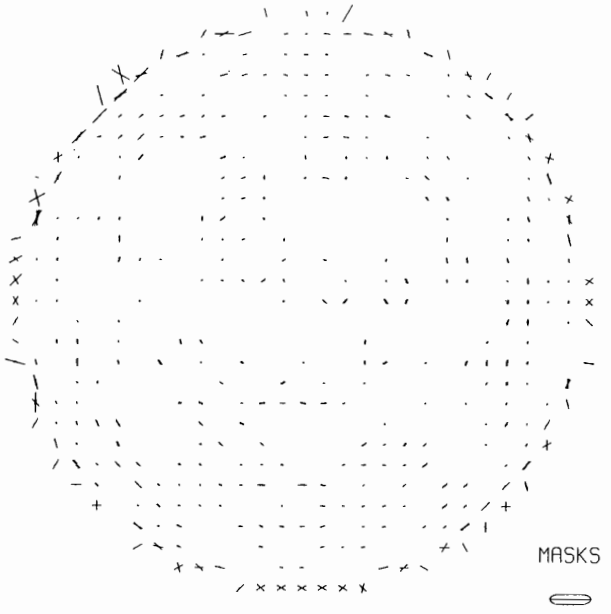

FIG. 34. Equilibrium activities of a mask field computed directly from Fig. 30 are shown here.

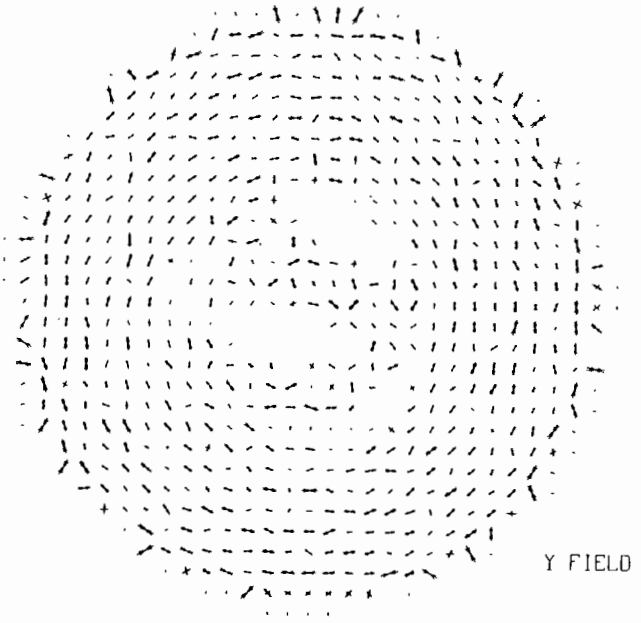

FIG. 35. Equilibrium activities of the second competitive stage in response to the mask field of Fig. 34.

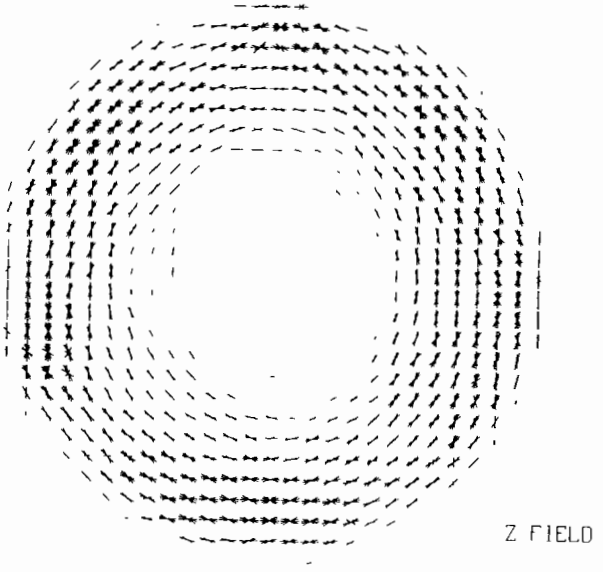

FIG. 36. Equilibrium activities of the cooperative stage in response to the mask field of Fig. 34. A smooth tracking of underlying statistical groupings of orientations is accomplished, even though the texture elements themselves are discrete patches and the mask field responses are weak and locally ambiguous.

CC loop within this spatial scale. Here a dense boundary web forms that completes a coherent representation of the globally consistent statistical groupings of oriented activity in Fig. 34. Note that while near the center of Fig. 35 most activity is in direct response to the luminance contrasts of the texture elements with the background, near the edges of the figure cooperative activity over large regions emerges. The tendency of this cooperative activity to be strongest in the periphery of the figure can be appreciated in Fig. 36, which depicts the output of the cooperative stage in response to the input of Fig. 34.

25. CONCLUDING REMARKS: THE DYNAMIC GEOMETRY OF SURFACE FORM AND APPEARANCE

Classical mathematical theories of surfaces are predicated on the possibility of decomposing a surface into infinitesimal regions within which concepts such as surface normal and Jacobian can be unambiguously defined. In contrast, our analysis suggests that the concept of infinitesimal regions is not consonant with emergent percepts of pre-attentive surface form, and that attendant concepts such as surface normal are irrelevant in perceptual processes. Instead, we have indicated how self-scaling form-sensitive coordinate systems are induced dynamically by combinations of image contrasts, and that perceived surface shape emerges not from any single measure taken on explicitly defined surface elements, but rather from the collective action of multiple completion and filling-in events distributed across several spatial scales.

This nonclassical approach to the perception of 3D form is equally at home with discrete and continuous sources of scenic information, and in fact can utilize multiple sources of information cooperatively to generate a less ambiguous form percept than any single source of information could sustain. While local computational units play a central role in classical geometrical theories, they are fundamentally insufficient to explain the stability and coherence of a percept of form. In the present theory, whereas relatively local analog signals are needed to initiate the processing of form information, the grouping mechanisms which feed upon these signals transform them into structural representations using nonlinear feedback interactions that are sensitive to large image domains. The design of the CC loop illustrates how such long-range interactions can override imperfections of local analog signalling without becoming insensitive to global structural properties. Thus our model suggests that an analysis of form perception which begins to address the phenomenological richness of perceptual data in this field mandates a radical break from the classical concepts of geometry which have been the source of most models of surface perception until the present time.

APPENDIX: BOUNDARY CONTOUR SYSTEM EQUATIONS

The network which we used to define the boundary contour system (BCS) is defined in stages below. This network further develops the BCS system that was described in Grossberg and Mingolla [37].

A. *Oriented Masks*

To define a mask, or oriented receptive field, centered at position (i, j) with orientation k, divide the elongated receptive field of the mask into a left half L_{ijk} and a right half R_{ijk}. Let all the masks sample a field of preprocessed inputs. If S_{pq}

equals the preprocessed input to position (p, q) of this field, then the output J_{ijk} from the mask at position (i, j) with orientation k is

$$J_{ijk} = \frac{\left[U_{ijk} - \alpha V_{ijk}\right]^+ + \left[V_{ijk} - \alpha U_{ijk}\right]^+}{1 + \beta\left(U_{ijk} + V_{ijk}\right)}, \tag{A1}$$

where

$$U_{ijk} = \sum_{(p,q) \in L_{ijk}} S_{pq}, \tag{A2}$$

$$V_{ijk} = \sum_{(p,q) \in R_{ijk}} S_{pq}, \tag{A3}$$

and the notation $[p]^+ = \max(p, 0)$. The sum of the two terms in the numerator of (A1) says that J_{ijk} is sensitive to the orientation and amount-of-contrast, but not to the direction-of-contrast, received by L_{ijk} and R_{ijk}. The denominator term in (A1) enables J_{ijk} to compute a ratio scale in the limit where $\beta(U_{ijk} + V_{ijk})$ is much greater than 1. In all of our simulations, we have chosen $\beta = 0$.

B. *On-Center Off-Surround Interaction Within Each Orientation (Competition I)*

Inputs J_{ijk} with a fixed orientation k activate potentials w_{ijk} at the first competitive stage via on-center off-surround interactions: each J_{ijk} excites w_{ijk} and inhibits w_{pqk} if $|p - i|^2 + |q - j|^2$ is sufficiently small. All the potentials w_{ijk} are also excited by the same tonic input I, which supports disinhibitory activations at the next competitive stage. Thus

$$\frac{d}{dt} w_{ijk} = -w_{ijk} + I + f\left(J_{ijk}\right) - w_{ijk} \sum_{(p,q)} f\left(J_{pqk}\right) A_{pqij}, \tag{A4}$$

where A_{pqij} is the inhibitory interaction strength between positions (p, q) and (i, j) and $f(J_{ijk})$ is the input signal generated by J_{ijk}. In our runs, we chose $f(J_{ijk})$ proportional to J_{ijk}. Section C defines the on-cell subfield of the dipole field described in Fig. 8.

C. *Normalization and Push-Pull Opponent Process between Orientations at Each Position (Competition II)*

For simplicity, we assume that the output signals from the w_{ijk} potentials equal the potentials themselves, which are always nonnegative. Each w_{ijk} influences potentials y_{ijm} at the second competitive stage via a shunting on-center off-surround interaction across orientations m at each position (i, j),

$$\frac{d}{dt} y_{ijk} = -y_{ijk} + \left(B - y_{ijk}\right) \sum_m w_{ijm} C_{mk} - \left(y_{ijk} + D\right) \sum_m w_{ijm} E_{mk} \tag{A5}$$

where

$$C_{mk} = C\exp\left[-\mu(m-k)^2\right] \tag{A6}$$

and

$$E_{mk} = E\exp\left[-\nu(m-k)^2\right]. \tag{A7}$$

Equation (A5) says that input w_{ijk} excites orientation k maximally and nearby orientations m with an exponentially decreasing strength C_{mk} at each position (i, j). Input w_{ijk} also inhibits orientations m with an exponentially decreasing strength E_{mk} at each position (i, j). Since $\mu < \nu$, the definitions (A6) and (A7) permit w_{ijk} to generate maximal inhibition at, or near, the orientation K that is perpendicular to k. In addition, the responses y_{ijk} tend to normalize the inputs w_{ijk} since, at equilibrium,

$$y_{ijk} = \frac{\sum_m w_{ijm}(BC_{mk} - DE_{mk})}{1 + \sum_m w_{ijm}(C_{mk} + E_{mk})}. \tag{A8}$$

By (A8), y_{ijk} is sensitive to ratios of weighted w_{ijm} values if the coefficients C_{mk} and E_{mk} in (A6) and (A7) are sufficiently large.

D. *Opponent Inputs to the Cooperative Stage*

The w_{ijk}, x_{ijk}, and y_{ijk} potentials are all assumed to be part of the on-cell subfield of a dipole field. If y_{ijk} is excited, an excitatory signal $f(y_{ijk})$ is generated at the cooperative stage. When potential y_{ijk} is excited, the potential y_{ijK} corresponding to the perpendicular orientation is inhibited. Both of these potentials form part of the on-cell subfield of a dipole field. Inhibition of an on-cell potential y_{ijK} disinhibits the corresponding off-cell potential $\bar{y}_{ijK}$, which sends an inhibitory signal $-f(\bar{y}_{ijK})$ to the cooperative level. The signals $f(y_{ijk})$ and $-f(\bar{y}_{ijK})$ thus occur together. In order to instantiate these properties, we made the simplest hypothesis, namely that

$$\bar{y}_{ijK} = y_{ijk}. \tag{A9}$$

E. *Oriented Cooperation: Statistical Gates*

The cooperative potential z_{ijk} can be supraliminally activated only if both of its cooperative input branches receive enough net positive excitation from similarly aligned competitive potentials. Thus

$$\frac{d}{dt}z_{ijk} = -z_{ijk} + g\left(\sum_{(p,q,r)}\left[f(y_{pqr}) - f(\bar{y}_{pqr})\right]F_{pqij}^{(r,k)}\right) + g\left(\sum_{(p,q,r)}\left[f(y_{pqr}) - f(\bar{y}_{pqr})\right]G_{pqij}^{(r,k)}\right). \tag{A10}$$

In (A10), $g(s)$ is a signal function that becomes positive only when s is positive, and has a finite maximum value. A slower-than-linear function

$$g(s) = \frac{H[s]^+}{K + [s]^+} \tag{A11}$$

was used in our simulations. A sum of two sufficiently positive $g(s)$ terms in (A10) is needed to activate z_{ijk} above the firing threshold of its output signal $h(z_{ijk})$. A threshold-linear signal function

$$h(z) = L[z - M]^+ \tag{A12}$$

was used. Each sum such as

$$\sum_{(p,q,r)} f(y_{pqr}) F_{pqij}^{(r,k)} \tag{A13}$$

and

$$\sum_{(p,q,r)} f(y_{pqr}) G_{pqij}^{(r,k)} \tag{A14}$$

is a spatial cross-correlation that adds up inputs from a strip with orientation (approximately equal to) k that lies to one side or the other of position (i, j), as in Fig. 9. The orientations r that contribute to the spatial kernels $F_{pqij}^{(r,k)}$ and $G_{pqij}^{(r,k)}$ also approximately equal k. The kernels $F_{pq}^{(r,k)}$ and $G_{pq}^{(r,k)}$ are defined by

$$F_{pqij}^{(r,k)} = \left[\exp\left[-2\left(\frac{N_{pqij}}{P} - 1\right)^2\right]\left[\left|\cos(Q_{pqij} - r)\right|\right]^R\left[\cos(Q_{pqij} - k)\right]^T\right]^+ \tag{A15}$$

and

$$G_{pqij}^{(r,k)} = \left[-\exp\left[-2\left(\frac{N_{pqij}}{P} - 1\right)^2\right]\left[\left|\cos(Q_{pqij} - r)\right|\right]^R\left[\cos(Q_{pqij} - k)\right]^T\right]^+. \tag{A16}$$

where

$$N_{pqij} = \sqrt{(p - i)^2 + (q - j)^2}, \tag{A17}$$

$$Q_{pqij} = \arctan\left(\frac{q - j}{p - i}\right), \tag{A18}$$

and P, R, and T are positive constants. In particular, R and T are odd integers.

Kernels F and G differ only by a minus sign under the $[\cdots]^+$ sign. This minus sign determines the polarity of the kernel; namely, whether it collects inputs for z_{ijk} from one side or the other of position (i, j). Term $\exp[-2(N_{pqij}/P - 1)^2]$ determines the optimal distance P from (i, j) at which each kernel collects its inputs. The kernel decays in a Gaussian fashion as a function of N_{pqij}/P, where N_{pqij} in (A17) is the distance between (p, q) and (i, j). The cosine terms in (A15) and (A16) determine the orientational tuning of the kernels. By (A18), Q_{pqij} is the direction of position (p, q) with respect to the position of the cooperative cell (i, j) in (A10). Term $|\cos(Q_{pqij} - r)|$ in (A15) and (A16) computes how parallel Q_{pqij} is to the receptive field orientation r at position (p, q). By (A18), term $|\cos(Q_{pqij} - r)|$ is maximal when the orientation r equals the orientation of (p, q) with respect to (i, j). The absolute value sign around this term prevents it from becoming negative. Term $\cos(Q_{pqij} - k)$ in (A15) and (A16) computes how parallel Q_{pqij} is to the orientation k of the receptive field of the cooperative cell (i, j) in (A10). By (A18), term $\cos(Q_{pqij} - k)$ is maximal when the orientation k equals the orientation of (p,q) with respect to (i, j). Positions (p, q) such that $\cos(Q_{pqij} - k) < 0$ do not input to z_{ijk} via kernel F because the $[\cdots]^+$ of a negative number equals zero. On the other hand, such positions (p, q) may input to z_{ijk} via kernal G due to the extra minus sign in the definition of kernel G. The extra minus sign in (A16) flips the preferred axis of orientation of kernel $G_{pqij}^{(r,k)}$ with respect to the kernel $F_{pqij}^{(r,k)}$ in order to define the two input-collecting branches of each cooperative cell, as in Fig. 9. The product terms $\pm|\cos(Q_{pqji} - r)|^R \cos(Q_{pqij} - k)^T$ in (A15) and (A16) thus determine larger path weights from dipole field on-cells whose positions and orientations are nearly parallel to the preferred orientation k of the cooperative cell (i, j), and larger path weights from dipole field off-cells whose positions and orientations are nearly perpendicular to the preferred orientation k of the cooperative cell (i, j). The powers R and T determine the sharpness of orientational tuning: Higher powers enforce sharper tuning.

F. *On-Center Off-Surround Feedback within Each Orientation*

We assume that each z_{ijk} activates a shunting on-center off-surround interaction within each orientation k across position (i, j). The target potentials v_{ijk} therefore obey an equation of the form

$$\frac{d}{dt}v_{ijk} = -v_{ijk} + h(z_{ijk}) - v_{ijk}\sum_{(p,q)} h(z_{pqk})W_{pqij}. \tag{A19}$$

The bottom-up transformation $J_{ijk} \to w_{ijk}$ in (A4) is thus similar to the top-down transformation $z_{ijk} \to v_{ijk}$ in (A19). Functionally, the $z_{ijk} \to v_{ijk}$ transformation enables the most favored cooperations to enhance their preferred positions and orientation as they suppress nearby positions with the same orientation. The signals v_{ijk} take effect by inputting to the w_{ijk} opponent process. Equation (A4) is thus changed to

$$\frac{d}{dt}w_{ijk} = -w_{ijk} + I + f(J_{ijk}) + v_{ijk} - w_{ijk}\sum_{(p,q)} f(J_{pqk})A_{pqij}. \tag{A20}$$

At equilibrium, the computational logic of the BCS is determined, up to parameter choices, by the equations

$$J_{ijk} = \frac{\left[U_{ijk} - \alpha V_{ijk}\right]^+ + \left[V_{ijk} - \alpha U_{ijk}\right]^+}{1 + \beta\left(U_{ijk} + V_{ijk}\right)}, \tag{A1}$$

$$w_{ijk} = \frac{I + BJ_{ijk} + v_{ijk}}{1 + B\sum_{(p,q)} J_{pqk} A_{pqij}}, \tag{A21}$$

$$y_{ijk} = \frac{\sum_m w_{ijm}\left(BC_{mk} - DE_{mk}^{'}\right)}{1 + \sum_m w_{ijm}\left(C_{mk} + E_{mk}\right)}. \tag{A8}$$

$$z_{ijk} = g\left(\sum_{(p,q,r)} \left[f(y_{pqr}) - f(y_{pqR})\right] F_{pqij}^{(r,k)}\right) + g\left(\sum_{(p,q,r)} \left[f(y_{pqr}) - f(y_{pqR})\right] G_{pqij}^{(r,k)}\right), \tag{A22}$$

and

$$v_{ijk} = \frac{h(z_{ijk})}{1 + \sum_{(p,q)} h(z_{pqk}) W_{pqij}}. \tag{A23}$$

Wherever possible, simple spatial kernels were used. For example the kernels W_{pqij} in (A19) and A_{pqij} in (A20) were both chosen to be constant within a circular receptive field:

$$A_{pqij} = \begin{cases} A & \text{if } (p-i)^2 + (q-j)^2 \le A_0 \\ 0 & \text{otherwise} \end{cases} \tag{A24}$$

and

$$W_{pqij} = \begin{cases} W & \text{if } (p-i)^2 + (q-j)^2 \le W_0. \\ 0 & \text{otherwise.} \end{cases} \tag{A25}$$

The oriented receptive fields $L_{ijk} \cup R_{ijk}$ in (A2) and (A3) were chosen to have parallel linear sides with hemicircular ends.

ACKNOWLEDGMENTS

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

REFERENCES

1. J. Beck, The effect of gloss on perceived lightness, *Amer. J. Psychol.* **77**, 1964, 54–63.
2. J. Beck, Perceptual grouping produced by changes in orientation and shape, *Science* **154**, 1966, 538–540.
3. J. Beck, Effect of orientation and of shape similarity on perceptual grouping, *Percept. Psychophys.* **1**, 1966, 300–302.
4. J. Beck, Surface lightness and cues for the illumination, *Amer. J. Psychol.* **84**, No. 1, 1971, 1–11.

5. J. Beck, Similarity grouping and peripheral discriminability under uncertainty, *Amer. J. Psychol.* **85**, 1972, 1–19.
6. J. Beck, *Surface Color Perception*, Cornell Univ. Press, Ithaca, N.Y., 1972.
7. J. Beck, Textural segmentation, in *Organization and Representation in Perception* (J. Beck, Ed.), Erlbaum, Hillsdale, N.J., 1982.
8. J. Beck, Textural segmentation, second-order statistics, and textural elements, *Biol. Cybern.* **48**, 1983, 125–130.
9. J. Beck, and S. Prazdny, Highlights and the perception of glossiness, *Percept. Psychophys.*, No. 4, 1981, 407–410.
10. J. Beck, K. Prazdny, and R. Ivry, The perception of transparency with achromatic colors, *Percept. Psychophys.* **35**, 1984, 407–422.
11. J. Beck, K. Prazdny, and A. Rosenfeld, A theory of textural segmentation, in *Human and Machine Vision* (J. Beck, B. Hope, and A. Rosenfeld, Eds.), Academic Press, New York, 1983.
12. S. S. Bergström, Common and relative components of reflected light as information about the illuminations, colour and three-dimensional form of objects, *Scand. J. Psychol.* **18**, 1977, 180–186.
13. B. G. Breitmeyer, Unmasking visual masking: A look at the "why" behind the veil of the "how," *Psychol. Rev.* **87**, 1980, 52–69.
14. D. Brewster, On the conversion of relief by inverted vision, *Edinburgh Philos. Trans.* **15**, 1847, 657.
15. M. A. Cohen and S. Grossberg, Some global properties of binocular resonances: Disparity matching, filling-in, and figure-ground synthesis, in *Figural Synthesis* (P. Dodwell and T. Caelli, Eds.), Erlbaum, Hillsdale, N.J., 1984.
16. M. A. Cohen and S. Grossberg, Neural dynamics of brightness perception: Features, boundaries, diffusion, and resonance, *Percept. Psychophys.* **36**, 1984, 428–456.
17. P. Dev, Perception of depth surfaces in random-dot stereograms: A neural model, *Int. J. Man-Mach. Stud.* **7**, 1975, 511–528.
18. B. A. Dosher, G. Sperling, and S. Wurst, Tradeoffs between stereopsis and proximity luminance covariance as determinants of perceived 3D structure, *Vision Res.* in press.
19. H. Egusa, Effects of brightness, hue, and saturation on perceived depth between adjacent regions in the visual field, *Perception* **12**, 1983, 167–175.
20. H. R. Flock, Illumination: Inferred or observed? *Percept. Psychophys.* **35**, 1984, 293.
21. A. R. H. Gellatly, Perception of an illusory triangle with masked inducing figure, *Perception* **9**, 1980, 599–602.
22. J. J. Gibson, *Perception of the Visual World*, Houghton-Mifflin, Boston, 1950.
23. J. J. Gibson, *The Senses Considered as Perceptual Systems*, Houghton-Mifflin, Boston, 1966.
24. J. J. Gibson, *The Ecological Approach to Visual Perception*, Houghton-Mifflin, Boston, 1979.
25. A. L. Gilchrist, Perceived lightness depends on perceived spatial arrangement, *Science*, **195**, 1977, 185–197.
26. A. L. Gilchrist, When does perceived lightness depend on perceived spatial arrangement? *Percept. Psychophys.* **28**, No. 6, 1980, 527–538.
27. A. L. Gilchrist, The classification and integration of edges is critical to the perception of reflectance and illumination, *Percept. Psychophys.* **33**, No. 5, 1983, 425–436.
28. L. Glass, Perception of random dot interference patterns, *Nature* **246**, 1973, 360–362.
29. W. C. Gogel, The tendency to see objects as equidistant and its reverse relations to lateral separation, *Psychol. Monograph* **70**, 1956.
30. W. C. Gogel, Equidistance tendency and its consequences, *Psychol. Bull.* **64**, 1965, 153–163.
31. W. C. Gogel, The adjacency principle and three-dimensional visual illusions, *Psychonom. Monograph Suppl.* **3**, No. 45, 1970, 153–169.
32. N. Graham, The visual system does a crude Fourier analysis of patterns, in *Mathematical Psychology and Psychophysiology* (S. Grossberg, Ed.), Amer. Math. Soc., Providence, R.I., 1981.
33. N. Graham and J. Nachmias, Detection of grating patterns containing two spatial frequencies: A test of single-channel and multiple channel models, *Vision Res.* **11**, 1971, 251–259.
34. S. Grossberg, The quantized geometry of visual space: The coherent computation of depth, form, and lightness, *Behav. Brain Sci.* **6**, 1983, 625–692.
35. S. Grossberg, Cortical dynamics of three-dimensional form, color, and brightness perception: Parts I and II. *Percept. Psychophys.*, in press.
36. S. Grossberg and E. Mingolla, Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading, *Psychol. Rev.* **92**, 1985, 173–211.

37. S. Grossberg and E. Mingolla, Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations, *Percept. Psychophys.* **38**, 1985, 141–171.
38. S. Grossberg and E. Mingolla, The role of illusory contours in visual segmentation, in *Proceedings of the International Conference on Illusory Contours* (G. Meyer and S. Petry, Eds.), Pergamon, Elmsford, N.Y., 1986.
39. B. K. P. Horn, Obtaining shape from shading information, *The Psychology of Computer Vision* (P. H. Winston, Ed.), McGraw–Hill, New York, 1975.
40. B. K. P. Horn, Understanding image intensities, *Artif. Intell.* **8**, 1977, 201–231.
41. B. K. P. Horn, Hill-shading and the reflectance map, *Proc. IEEE* **19**, 1981, 14–47.
42. D. H. Hubel and T. N. Wiesel, Functional architecture of macaque monkey visual cortex, *Proc. R. Soc. London B* **198**, 1977, 1–59.
43. K. Ikeuchi and B. K. P. Horn, Numerical shape from shading and occluding images, *Artif. Intell.* **17**, 1981, 141–184.
44. B. Julesz, *Foundations of Cyclopean Perception*, Univ. of Chicago Press, Chicago, 1971.
45. G. Kanizsa, Subjective contours, *Sci. Amer.* **234**, 1976, 48–52.
46. G. Kanizsa, *Organization in Vision*, Praeger, New York, 1979.
47. D. Katz, *The World of Colour* (transl. by R. B. MacLeod and C. W. Fox), Kegan Paul, Trench, Trübner, London, 1935.
48. L. Kaufman, *Sight and Mind: An Introduction to Visual Perception*, Oxford Univ. Press, New York, 1974.
49. L. Kaufman, J. Bacon, and F. Barroso, Stereopsis without image segregation, *Vision Res.* **13**, 1973, 137–147.
50. N. Kawabata, Perception at the blind spot and similarity grouping, *Percept. Psychophys.* **36**, 1984, 151–158.
51. J. J. Koenderink and A. J. van Doorn, The structure of two-dimensional scalar fields with applications to vision, *Biol. Cybern.* **33**, 1979, 151–158.
52. J. J. Koenderink and A. J. van Doorn, Photometric invariants related to solid shape, *Opt. Acta* **27**, 1980, 981–996.
53. J. J. Koenderink and A. J. van Doorn, Perception of solid shape and spatial lay-out through photometric invariants, in *Cybernetics and Systems Research* (R. Trappl, Ed.), North-Holland, Amsterdam, 1982.
54. J. Krauskopf, Effect of retinal image stabilization on the appearance of heterochromatic targets, *J. Opt. Soc. Amer.* **53**, 1963, 741–744.
55. J. J. Kulikowski, Limit of single vision in stereopsis depends on contour sharpness, *Nature* **275**, 1978, 126–127.
56. E. H. Land, The retinex theory of color vision, *Sci. Amer.* **237**, 1977, 108–128.
57. D. Marr, *Vision*, Freeman, San Francisco, 1982.
58. D. Marr and H. K. Nishihara, Representation and recognition of the spatial organization of three-dimensional shapes, *Proc. R. Soc. London B* **200**, 1978, 269–294.
59. D. Marr and T. Poggio, Cooperative computation of stereo disparity, *Science* **194**, 1976, 283–287.
60. D. Marr and T. Poggio, A computational theory of human stereo vision, *Proc. R. Soc. London B* **204**, 1979, 301–328.
61. F. Metelli, The perception of transparency, *Sci. Amer.* **230**, No. 4, 1974, 90–98.
62. F. Metelli, O. Da Pos, and A. Cavedon, Balanced and unbalanced, complete and partial transparency, *Percept. Psychophys.* **38**, No. 4, 1985, 354–366.
63. G. G. Meyer and M. Senecal, The illusion of transparency and chromatic subjective contours, *Percept. Psychophys.* **34**, No. 1, 1983, 58–64.
64. E. Mingolla and J. T. Todd, Perception of solid shape from shading, *Biol. Cybern.* **53**, 1986, 137–151.
65. A. P. Pentland, The visual inference of shape: Computation from local features, unpublished doctoral dissertation, MIT, 1982.
66. K. Prazdny, On the nature of inducing forms generating perception of illusory contours, *Percept. Psychophys.* **37**, 1985, 237–242.
67. W. Preyer, On certain optical phenomena: Letter to Professor E. C. Sanford, *Amer. J. Psychol.* **9**, 1897/98, 42–44.
68. C. Redies and L. Spillmann, The neon color effect in the Ehrenstein illusion, *Perception* **10**, 1981, 667–681.
69. E. Reed and R. Jones, *Reasons for Realism: Selected Essays of James J. Gibson*, Erlbaum, Hillsdale, N.J., 1982.

70. R. I. Reynolds, Perception of an illusory contour as a function of processing time, *Perception* **10**, 1981, 107–115.
71. R. Shapley and J. Gordon, Nonlinearity in the perception of form, *Percept. Psychophys.* **37**, 1985, 84–88.
72. B. J. Schwartz and G. Sperling, Luminance controls the perceived 3-D structure of dynamic 2-D displays, *Bull. Psychonom. Soc.* **21**, No. 6, 1983, 456–458.
73. E. L. Schwartz, Computational anatomy and functional architecture of striate cortex: A spatial mapping approach to perceptual coding, *Vision Res.* **20**, 1980, 645–669.
74. G. Sperling, Binocular vision: A physical and a neural theory, *Amer. J. Psychol.* **83**, 1970, 461–534.
75. L. G. Thorell, R. L. De Valois, and D. G. Albrecht, Spatial mapping of monkey V1 cells with pure color and luminance stimuli, *Vision Res.* **24**, 1984, 751–769.
76. J. T. Todd, perception of three dimensional form from patterns of optical texture, presented at the *Third Human and Machine Vision Workshop*, Boston, 1985.
77. J. T. Todd and R. Akerstrom, Perception of three-dimensional form from patterns of optical texture, *J. Exp. Psychol. Human Percept. Perform.*, in press.
78. J. T. Todd and E. Mingolla, Perception of surface curvature and direction of illumination from patterns of shading, *J. Exp. Psychol. Human Percept. Perform.* **9**, No. 4, 1983, 583–595.
79. H. F. J. M. van Tuijl, A new visual illusion: Neonlike color spreading and complementary color induction between subjective contours, *Acta Psychol.* **39**, 1975, 441–445.
80. H. F. J. M. van Tuijl and C. M. M. de Weert, Sensory conditions for the occurrence of the neon spreading illusion, *Perception* **8**, 1979, 211–215.
81. R. J. Woodham, Analyzing images of curved surfaces, *Artif. Intell.* **17**, 1981, 117–140.
82. A. L. Yarbus, *Eye Movements and Vision*, Plenum, New York, 1967.