

OUTLINE OF A THEORY OF
BRIGHTNESS, COLOR, AND FORM PERCEPTION

Stephen Grossberg
Center for Adaptive Systems
Boston University

This paper describes new concepts and mechanisms from a real-time visual processing theory that has been used to explain paradoxical data about brightness and form perception. These data include the Craik-O'Brien effect, the Land brightness and color demonstrations, the fading of stabilized images, neon color spreading, complementary color induction, completion of illusory contours, and binocular rivalry. Two functionally distinct contour processes interact to generate these brightness and form properties in the theory. A Boundary Contour process is sensitive to the amount of contrast but not to the direction of contrast in scenic edges. It includes a binocular matching stage that is sensitive to spatial scale, orientation, and binocular disparity, and whose outcome triggers a process of monocular contour completion. These completed contours form the boundaries of monocular perceptual domains. A Feature Contour process is sensitive to both the amount of contrast and to the direction of contrast in scenic edges. It triggers a diffusive filling-in reaction of featural quality within perceptual domains whose boundaries are dynamically defined by the completed boundary contours. The diffusive filling-in reactions take place within syncytia of cell compartments. These preprocessed monocular representations give rise to a percept via a process of binocular resonance. The percept takes the form of standing waves of patterned activity among multiple spatial scales. The Boundary Contour process is hypothesized to be analogous to interactions between the hypercolumns in area 17 of the visual cortex. The Feature Contour process is hypothesized to be analogous to interactions between the cytochrome oxydase staining blobs in area 17 and prestriate cortex in area 18.

1. INTRODUCTION : PARADOXICAL VISUAL DATA AS MANIFESTATIONS OF ADAPTIVE VISUAL MECHANISMS

This chapter outlines some of the new concepts and mechanisms that my colleagues and I have been using to explain paradoxical data about brightness and form perception (Cohen and Grossberg, 1983a, 1983b, 1984; Grossberg, 1981, 1983a, 1983b; Grossberg and Mingolla, 1984). That paradoxical data abound in the field of visual perception does not seem as surprising when one considers the manner in which visual information is acquired. Light passes through retinal veins before it reaches retinal photoreceptors, and light does not influence the retinal region corresponding to the blind spot. The percepts of human observers are not distorted, however, by their retinal veins or blind spots during normal viewing conditions. Thus some images that are retinally present are not perceived because our visual processes are adaptively designed to free our percepts from imperfections of visual scanning. These adaptive visual processes can also generate paradoxical percepts, as during the perception of stabilized images (Riggs, Ratliff, Cornsweet, and Cornsweet, 1953; Pritchard, Heron and Hebb, 1960; Pritchard, 1961; Yabus, 1967), and of filling-in reactions (Gerrits and de Haan, 1966; Gerrits and Timmerman, 1969; Gerrits and Vendrick, 1970).

Other adaptive visual mechanisms exist than those which compensate for retinal veins and the blind spot. The visual mechanisms that prevent perception of retinal veins and the blind spot are operative even when only one eye is open. Due to the action of binocular visual mechanisms, some retinal images that can be monocularly perceived may not be perceived during binocular viewing conditions. Binocular rivalry provides a classical example of this fact (Blake and Fox, 1964; Cogan, 1982; Kaufman, 1974; Kulikowski, 1978). The properties of binocular rivalry also seem less paradoxical when one considers how the suppression of some monocular data can achieve a more consistent depthful binocular representation of a scene (Kaufman, 1974; Grossberg, 1980, 1983a).

The above examples illustrate that retinally present images may be suppressed to achieve more consistent visual representations of the external world. For a similar reason, observers may see images that are not retinally present. The remarkable Land color demonstrations (Land, 1977) are, for example, often interpreted as consequences of the visual system's ability to process the reflectances of perceived objects despite the

variability of the lighting conditions under which the objects are observed. Our theory clarifies the adaptive mechanisms that give rise to the fading of stabilized images, filling-in, and the Land brightness color demonstrations. The same mechanisms also shed light on the perception of illusory figures (Gellatly, 1980; Kanizsa, 1974; Kennedy, 1978, 1979, 1981; Parks, 1980; Parks and Marks, 1983; Petry, Harbeck, Conway, and Levey, 1983), of neon color spreading (Redies and Spillman, 1981; van Tuijl, 1975; van Tuijl and de Weert, 1979; van Tuijl and Leeuwenberg, 1979), and of Julesz stereograms (Julesz, 1971; Kaufman, 1974).

Every theory of adaptive behavior is faced with the problem of specifying precisely the sense in which its targeted behavior is adaptive, without becoming circular or falling into an infinite regress. This difficult problem does not go away just by citing platitudes about Darwinian evolution. The principles of organization and real-time mechanisms that instantiate our view of adaptive behavioral processes are often far removed from teleological concepts that are easily expressed using consensual language. Some of our concepts analyse, for example, how the nervous system, or indeed any cellular machine, may respond to its input environment in a sensitive and stable fashion without losing its capability for regulated self-organization (Grossberg, 1978, 1980, 1982). This approach indicates how behavioral properties which seem to reflect an intelligent teleology may arise as the epiphenomena of general constraints on the possibility of doing accurate information processing at all.

The present chapter does not derive these principles and mechanisms from first principles. Rather it outlines how some of these general tools, once available, can be specialized for use in brightness and form perception. My focus will be on processes that are hypothesized to occur at a central level of neural processing, notably in the visual cortical areas 17 and 18. The neural interpretation of these processes is not, however, critical to understanding how they work. All of the processes may be interpreted as formal rules for real-time visual information processing. The paper with Professor Carpenter (Carpenter and Grossberg, (1981), by contrast, illustrates how an adaptive approach may be used to clarify data taken from the most peripheral photoreceptive cells of the vertebrate nervous system. This work may also be studied as a purely formal model, notably a model concerning how a receptive device can recalibrate its sensitivity in response to an input source whose intensity

can fluctuate over a very large range through time.

2. BOUNDARY CONTOURS AND FEATURE CONTOURS.

One of the most important new concepts of our theory is that several types of contour extracting processes are simultaneously operative during brightness and form perception (Grossberg, 1983a, 1983b). Two of these contour processes are said to generate boundary contours and feature contours, respectively. The very fact that more than one type of contour process is hypothesized to exist implies that care must be exerted in analysing what one means by a contour, or edge, and how it is computed. For example, neither of our contour processes is computed using zero crossings (Marr and Hildreth, 1980), which are wholly inadequate for our purposes (Grossberg, 1983a). This chapter will describe and use these new contour concepts in an intuitive and nontechnical way.

Before illustrating how boundary and feature contour processes help to explain paradoxical visual data, I will differentiate them using several of their more salient properties.

2.1. BOUNDARY CONTOURS AND CONTOUR COMPLETION

(i) Contrast : A boundary contour is sensitive to the amount of contrast, but not to the direction of contrast, at an edge of a visual scene. Thus a boundary contour can be activated by either a light-dark edge or a dark-light edge at a fixed scenic position.

(ii) Completion : Several boundary contour activations that correspond to contrast differences at different scenic positions may interact via a process of contour completion. The contour completion process is capable of synthesizing global visual contours from local contour fragments. The contours that are perceived in illusory figures are assumed to be generated by this contour completion process. Completed boundary contours are used to bound perceptually significant domains in the theory; hence the name boundary contour.

A simple demonstration of a contour process with these properties can be made as follows (Grossberg and Mingolla, 1984). Divide a square into two equal rectangles along an imaginary boundary. Color one rectangle a uniform shade of gray. Color the other rectangle in shades of gray that process from light to dark as one moves from end 1 of the rectangle to end 2 of the rectangle. Color end 1 a lighter shade than the uniform gray

of the other rectangle, and color end 2 a darker shade than the uniform gray of the other rectangle. Then as one moves from end 1 to end 2, an intermediate gray region is passed whose luminance approximately equals that of the uniform rectangle. At end 1, a light-dark edge exists from the nonuniform rectangle to the uniform rectangle. At end 2, a dark-light edge exists from the nonuniform rectangle to the uniform rectangle. Despite this reversal in the direction of contrast from end 1 to end 2, an observer can see an illusory edge that joins the two edges of opposite contrast and separates the intermediate rectangle regions of equal luminance.

This demonstration illustrates the existence of a process of completing contours between edges with opposite directions of contrast. This contour completion process is thus sensitive to amount of contrast, but not to direction of contrast.

(iii) Binocular Matching : A monocular boundary contour can be generated when a single eye views a visual scene. When two eyes view a visual scene, pairs of monocular boundary contours become engaged in a process of binocular matching. This binocular matching process is sensitive to such factors as spatial scale, binocular disparity, and edge orientation. Due to the binocular matching process, a fused binocular boundary contour may be generated from a pair of matched monocular boundary contours. Alternatively, one of the monocular boundary contours may be suppressed, or a double image of unfused monocular boundary contours may occasionally be generated. These outcomes provide the local substrate for such global processes as binocular rivalry and depth perception in the theory (Grossberg, 1980, 1983a). In particular, if monocular boundary contours corresponding to the left eye are momentarily suppressed by binocular mismatch, then these monocular boundary contours cannot contribute to the contour completion process corresponding to the left visual field. Visual percepts that can be seen during monocular viewing may thus be suppressed by this binocular matching process. These processes of binocular matching and contour completion are interpreted neurally to occur in the striate cortex (area 17).

The boundary contour processes are like frames without pictures. The pictorial data themselves are derived from the feature contour processes. This suggestive analogy with pictures and frames should not be taken too literally, however, since the same visual sources input to both the boundary contour process and the feature contour process, and the outputs of

both types of processes interact in a contex-sensitive way to generate visual representations.

2.2. FEATURE CONTOURS AND DIFFUSIVE FILLING-IN

(i) Contrast : The feature contour process is sensitive to direction of contrast as well as to amount of contrast, unlike the boundary contour process. For example, it is obviously important in computing the relative brightness across a scenic boundary to take account of which side of the scenic boundary has a larger reflectance. A similar remark holds at a red-green scenic boundary.

(ii) Completion : Boundary contours activate a contour completion process that generates the global contours which bound monocular perceptual domains. Feature contours activate a diffusive filling-in process that spreads featural qualities, such as brightness or color, across these monocular perceptual domains. Such a diffusive filling-in reaction is, for example, hypothesized to prevent perception of the blind spot.

This diffusive filling-in process is spatially limited by boundary contours. The binocular matching process that occurs between monocular boundary contours chooses among the possible spatial scales that will receive boundary contour signals. Thus, although a feature contour may be activated across all of these spatial scales, only those scales that receive boundary contour signals will generate barriers to the diffusion of featural quality. Only these scales will be capable of contributing to the final percept in the theory, as the subsequent discussion will show.

(iii) Monocular Filling-In : Because monocular boundary contours are binocularly matched before giving rise to completed contours, each visual field (left and right) may process different completed boundary contours, just as in binocular rivalry. These distinct completed boundary contours can differentially regulate the monocular processing of featural quality in their respective visual fields.

These concepts will now be refined using a macrocircuit diagram and applied to several types of perceptual phenomena.

3. MACROCIRCUIT OF PROCESSING STAGES.

Figure 1 describes a macrocircuit of processing stages that is capable of synthesizing global depth, brightness, and form information from monocularly

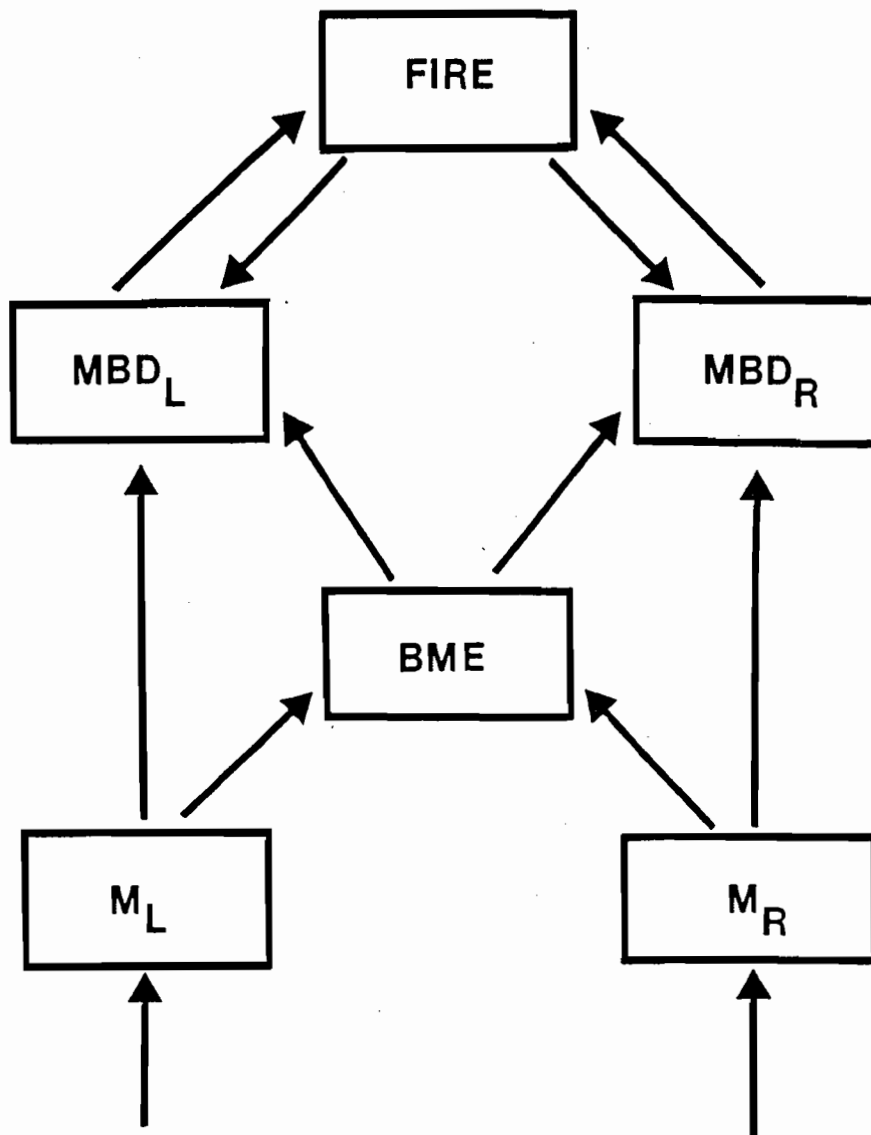


Figure 1

Network stages used to extract global depth, brightness, and form information from monocular and binocular images. Table I lists the names of these stages and their possible neural analogs.

and binocularly viewed patterns. Table I lists the names of the abbreviated processing stages, as well as the neural structures that presently seem most likely to execute analogous processes.

Table I

M_L	- Left Monocular Representation (Lateral geniculate nucleus)
M_R	- Right Monocular Representation (Lateral geniculate nucleus)
BME	- Binocular Matching of Edges (Hypercolumn interactions in striate cortex - Area 17 (Hubel and Wiesel, 1977))
MBD_L	- Left Monocular Brightness Domain (Cytochrome oxydase staining blobs - Area 17 - and/or prestriate cortex - Area 18 (Horton and Hubel, 1981; Hendrickson, Hunt, and Wu, 1981; Hubel and Livingstone, 1981; Livingstone and Hubel, 1982))
MBD_R	- Right Monocular Brightness Domain (Prestriate cortex - Area 18 - and/or cytochrome oxydase staining blobs in Area 17)
FIRE	- Filling-In Resonant Exchange (Prestriate cortex - Area 18)

The BME stage generates oriented monocular boundary contours in response to inputs from M_L and M_R , binocularly matches these monocular boundary contours, and monocularly completes the boundary contours that survive the binocular matching process. The completed left boundary contours generate scale-specific boundary signals to the MBD_L stage. The completed right boundary contours generate scale-specific boundary signals to the MBD_R stage.

Monocular feature contours are generated by $M_L \rightarrow MBD_L$ signals and $M_R \rightarrow MBD_R$ signals. Thus the monocular regions M_L and M_R give rise to both boundary contour signals and to feature contour signals. The boundary contours undergo binocular matching and monocular contour completion before rejoining the corresponding feature contours at MBD_L and MBD_R . This divergence of feature contour processing from boundary contour processing allows the boundary contours to benefit from scale-specific, orientation-

specific, and disparity-specific binocular selection within BME before the surviving boundary contours are passed among the spatial scales of MBD_L and MBD_R .

4. DIFFUSIVE FILLING-IN WITHIN MONOCULAR BRIGHTNESS AND COLOR DOMAINS.

The monocular boundary contour signals from stage BME to stage MBD_L define barriers to the lateral diffusion of featural activity that is triggered by feature boundary signals from M_L to MBD_L . This diffusive filling-in reaction is assumed to occur within a syncytium of cell compartments (Fig. 2). Within this cellular syncytium, the feature boundary signals trigger a rapid diffusion, or averaging, of electrical potential across compartment boundaries, except across those compartment boundaries that receive boundary contour signals from the BME stage. The feature contour signals from stage M_L to stage MBD_L are hereby smoothed within stage MBD_L between successive boundary contours that are considered perceptually significant by the BME stage (Cohen and Grossberg, 1983b, 1984; Grossberg, 1983a, 1983b).

Boundary contour signals from stage BME to stage MBD_L are assumed to prevent diffusion across their target compartments by increasing the electrical resistance of the target compartment membranes. It is assumed that the boundary contour signals achieve this increase in cell resistance by acting as inhibitory signals to their target compartments. At the same time that these inhibitory signals increase membrane resistances, they also help to decrease, or hyperpolarize, the target cell potentials. An analogous discussion shows how signals from the BME stage to the MBD_R stage define the boundaries that contain the filling-in reactions which occur in response to feature contour signals from the M_R stage to the MBD_R stage.

5. DOUBLE STEP BRIGHTNESS ILLUSION.

Using these interactions between boundary contour signals, feature contour signals, and diffusive filling-in, a large number of paradoxical brightness percepts have been simulated (Cohen and Grossberg, 1983b, 1984). One well-known phenomenon was studied by Arend, Buehler, and Lockhead (1971) during their investigations of the Craik-O'Brien illusion (O'Brien, 1958).

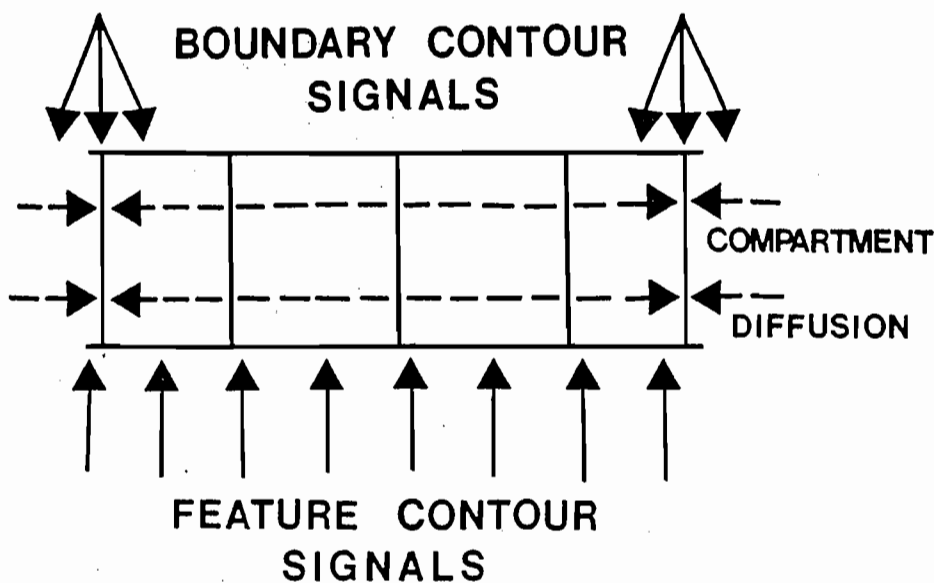


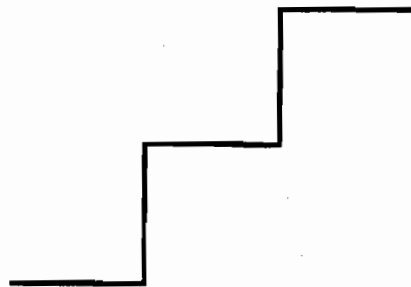
Figure 2

A monocular brightness domain in stage MBD_L or stage MBD_R . The feature contour signals excite cell compartments which permit rapid lateral diffusion, or filling-in, of activity across their membranes, except at those membranes which receive boundary contour signals from the BME stage. These boundary contour signals inhibit their target membranes and increase their resistance against the lateral flow of activation. The feature contour signals are thus averaged within domains whose boundaries are considered perceptually significant by the BME stage.

Figure 3a describes a luminance profile that was generated by rapidly rotating a disk covered with appropriately cut sectors of black and white papers. The luminances to the left and to the right of the luminance cusps are equal, and the average luminance across the cusps equals the background luminance. Figure 3b shows that this luminance profile is perceived as (approximate) steps of increasing brightness.



(a)



(b)

Figure 3

(a) A double Craik-O'Brien luminance profile in which the luminances to the left and right of the luminance cusps are equal, and the average luminance across the cusps equals the background luminance. (b) This luminance profile appears to an observer as two (approximate) steps of brightness.

A similar result is generated by our theory (Figure 4). In the theory, each of the two luminance cusps in the input pattern generates a boundary contour that restricts the lateral diffusion of featural activity. These two boundaries are depicted by the reduced diffusion coefficients in Figure 4. The feature contour process attenuates the background luminance of the input pattern and enhances the contrasts of the cusps. The boundary contour signals also contribute to the total input pattern that is received by the cell compartments. This total input pattern is called the filtered input pattern with boundaries in Figure 4. (The flanks of this total input pattern were artificially extended to the left and to the right to avoid contamination by boundary effects, because our computer could simulate only 2500 cells in a reasonable amount of time). When this total input is allowed to diffuse within the domains defined by the boundary contour signals, the steplike monocular preprocessed pattern of Figure 4 is the result. Appendix A describes the network equations that generate the patterns in Figure 4.

6. NEON COLOR SPREADING AND COMPLEMENTARY COLOR INDUCTION

The phenomenon of neon color spreading illustrates the existence of boundary contours and of feature in a particularly vivid way. Redies and Spillman (1981) have, for example, reported an experiment using a solid red cross and an Ehrenstein figure. When the solid red cross is perceived in isolation, it looks quite uninteresting (Figure 5a). When an Ehrenstein figure is perceived in isolation, it generates an illusory contour whose shape (e.g., circle or diamond) depends upon the viewing distance. When the red cross is placed inside the Ehrenstein figure, the red color flows out of its containing contours and tends to fill the illusory figure (Figure 5b).

In our theory, this percept is explained by showing how boundary contours of the Ehrenstein figure inhibit contiguous boundary contours of the red cross. This inhibitory process takes place within the BME stage as part of the contour completion process (Grossberg and Mingolla, 1984). This inhibitory action within the BME does not prevent the processing of feature contour signals from stage M_L to stage MBD_L and from stage M_R to stage MBD_R . Thus the red contour signals are received by MBD_L and MBD_R despite the fact that some of their corresponding boundary contour signals are inhibited within the BME stage.

STEP ILLUSION

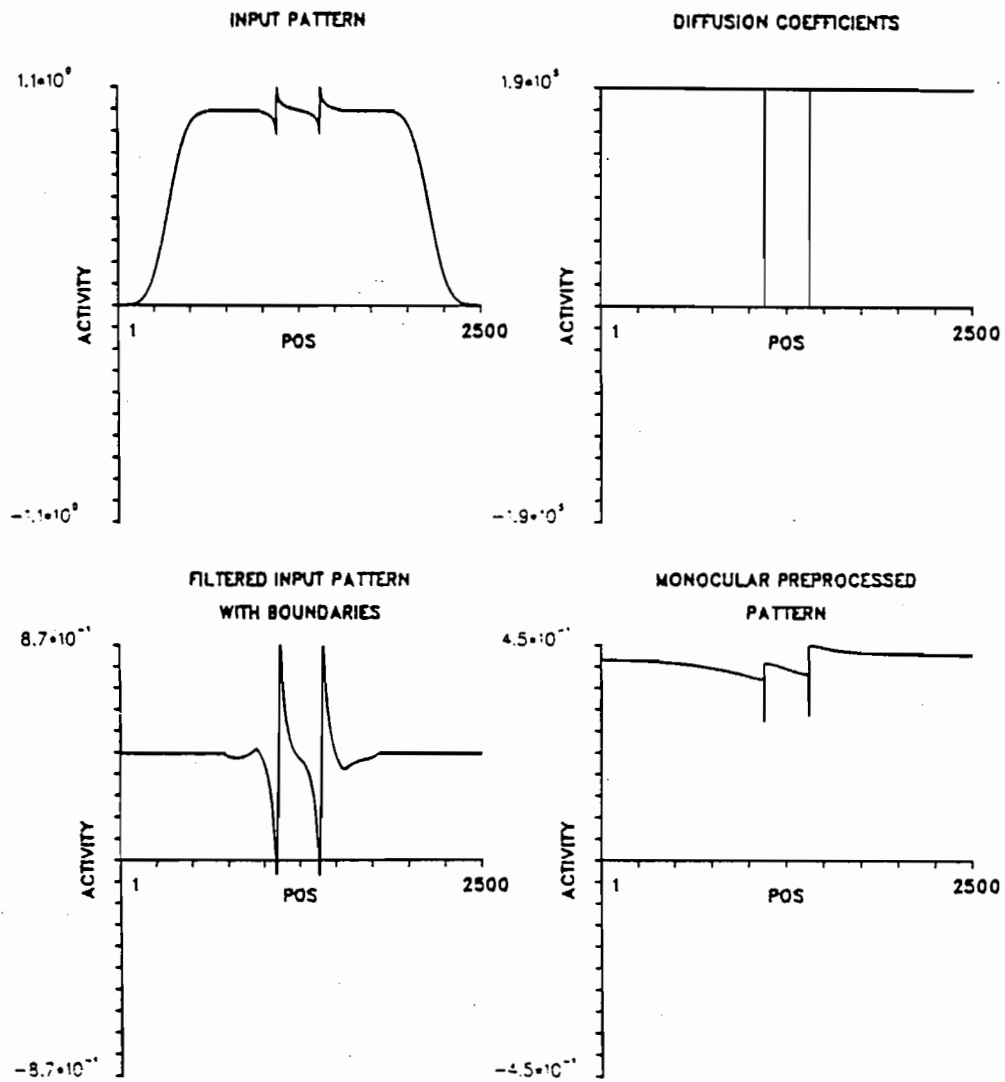
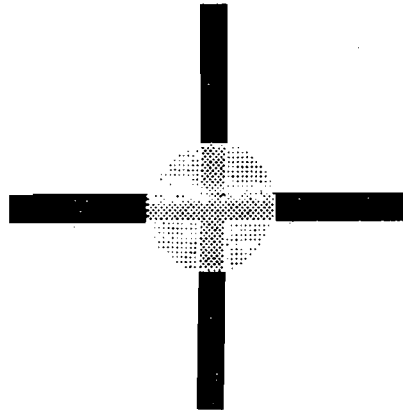


Figure 4

Simulation of the double step brightness illusion. The text describes how the double Craik-O'Brien input pattern generates boundary contours and feature contours that interact to form the step-like output pattern.



(a)



(b)

Figure 5

- (a) A red cross seems unremarkable when perceived in isolation.
(b) When perceived within an Ehrenstein figure, the red color flows outside of the cross and is contained by the illusory contour that is generated by the Ehrenstein figure.

The inhibition of these boundary contour signals within the BME stage allows the red featural activity to diffuse outside of the red cross. The illusory boundary contour that is induced by the Ehrenstein figure restricts the diffusion of this red-labelled activation. This example of neon color spreading thus provides a vivid example in which one can "see" the difference between boundary contours and feature contours, as well as the role of illusory boundary contours in restricting the diffusion of featural activity.

Redies and Spillman (1981) have systematically varied the distance of the red cross from the Ehrenstein contour, their relative orientations, their relative sizes, etc. to study how the strength of the spreading effect changes with scenic parameters. Of particular interest is the fact that "thin [red] flanks running alongside the red connecting lines" (p. 679) can occur if the Ehrenstein figure is slightly separated from the cross. In our theory, separating the Ehrenstein figure from the cross weakens the inhibitory effect of Ehrenstein boundary contours upon the boundary contours of the cross. When the boundary contours of the cross are partially weakened, they can still partially restrict the diffusion of red-labelled activation. Then the red color can only "bleed" outside the contours of the cross.

One might object to this explanation by questioning the ability of the Ehrenstein boundary contours to inhibit the boundary contours of the cross. Why do not the Ehrenstein boundary contours also inhibit contiguous Ehrenstein boundary contours? If they do, then how do any boundary contours survive this process of mutual inhibition? If they do not, then doesn't this explanation of neon color spreading collapse?

At this point, let us recall that the boundary contour process is sensitive to the amount of contrast, even though it is insensitive to the direction of contrast. I claim that the Ehrenstein boundary contours do mutually inhibit one another, but that this inhibition is a type of shunting lateral inhibition such that equally strong inhibitory contour signals can remain positive and balanced (Grossberg, 1973, 1983a). If, however, the Ehrenstein boundary signals are stronger than the boundary signals of the cross by a sufficient amount, then the latter signals can be significantly inhibited. This formal property of the contour completion process provides an explanation of the empirical fact that neon color spreading occurs only if the

contrast of a figure (e.g., the cross) relative to the background illumination is less than the contrast of the bounding contours (e.g., the Ehrenstein figure) relative to the background illumination (van Tuijl and de Weert, 1979).

A similar argument helps to explain the complementary color induction that van Tuijl (1975) reported in his original article about the neon effect. Draw on white paper a regular grid of horizontal and vertical black lines that form 5 mm. squares. Replace a subset of black lines by blue lines. Let this subset of lines be replaced, say, from the smallest imaginary diamond shape that includes complete vertical or horizontal line segments of the grid. When an observer inspects this pattern, the blue color of the lines appears to spread around the blue line elements until it reaches the subjective contours of the diamond shape. This percept has the same explanation as the percept in Figure 5b.

Next replace the black lines by blue lines and the blue lines by black lines. Then the illusory diamond looks yellow rather than blue. Let us suppose that the yellow color in the diamond is induced by the blue lines in the background matrix. Then why in the previous display is not a yellow color in the background induced by the blue lines in the diamond? Why is the complementary color yellow perceived when the background contains blue lines, whereas the original color blue is perceived when the diamond contains blue lines? What is the reason for this asymmetry?

I suggest that the reason is as follows. When the diamond is composed of blue lines, then opponent color processing enables the blue lines to induce contiguous yellow contour signals in the background. However these yellow contour signals are constrained by the boundary contour signals of the black lines to remain within the spatial domain of the feature contour of the black lines. Thus the yellow color is not seen in the background. By contrast, the boundary contour signals of the black lines in the background inhibit the contiguous boundary contour signals of the blue lines in the diamond. Hence the blue feature contour signals of the blue lines can flow within the diamond.

When blue lines form the background, they have two effects on the diamond. They induce yellow feature contours via opponent processing. They also inhibit the boundary contour signals of the contiguous black lines. Hence the yellow color can flow within the diamond.

To carry out this explanation quantitatively, we need to study how opponent color processes (red-green, yellow-blue) preprocess the feature contour signals from stage M_L to stage MBD_L and from stage M_R to stage MBD_R . Opponent color processes with the requisite properties are defined by gated dipole fields (Grossberg, 1980). We also need to quantitatively study how the contour completion process responds to complex spatial patterns such as grids and Ehrenstein figures (Grossberg and Mingolla, 1984).

7. FILLING-IN DUE TO STABILIZED IMAGES.

The distinction between boundary contours and feature contours also helps to explain percepts that are induced by stabilized images (Yarbus, 1967). Suppose, for example, that a large circle is surrounded by a uniform red background (Figure 6). Let the circle be divided down the middle by a vertical line. Within each hemicycle, place a small circle filled with the same red color as the background. Color the remainder of the left hemifield black and the remainder of the right hemifield white.

Now stabilize the retinal images of the large circular contour and its vertical edge. The background red is then seen to flow over the black and white hemifields. Since the background red color is the same color as the red color that fills the small circles, one might have supposed that the entire figure would look uniformly red. It does not. The red circle that lies within the previously black hemifield looks brighter, whereas the red circle that lies within the previously white hemifield looks darker.

To explain this percept, let us suppose that stabilizing a contour causes its boundary contour signals to decay (Grossberg, 1983b). Then the red feature contours that are exterior to the large circle can freely flow into the interior of the circle, even as the black and white featural contours of the interior of the circle intermix with one another and the red background. This filling-in process does not, however, enter the small red circles, because their scenic contours are not retinally stabilized.

The relative contrast of the red feature contour within the left red circle is greater than the relative contrast of the red feature contour within the right red circle. These red feature contours are contained by the uninhibited boundary contour signals around each small circle. When the

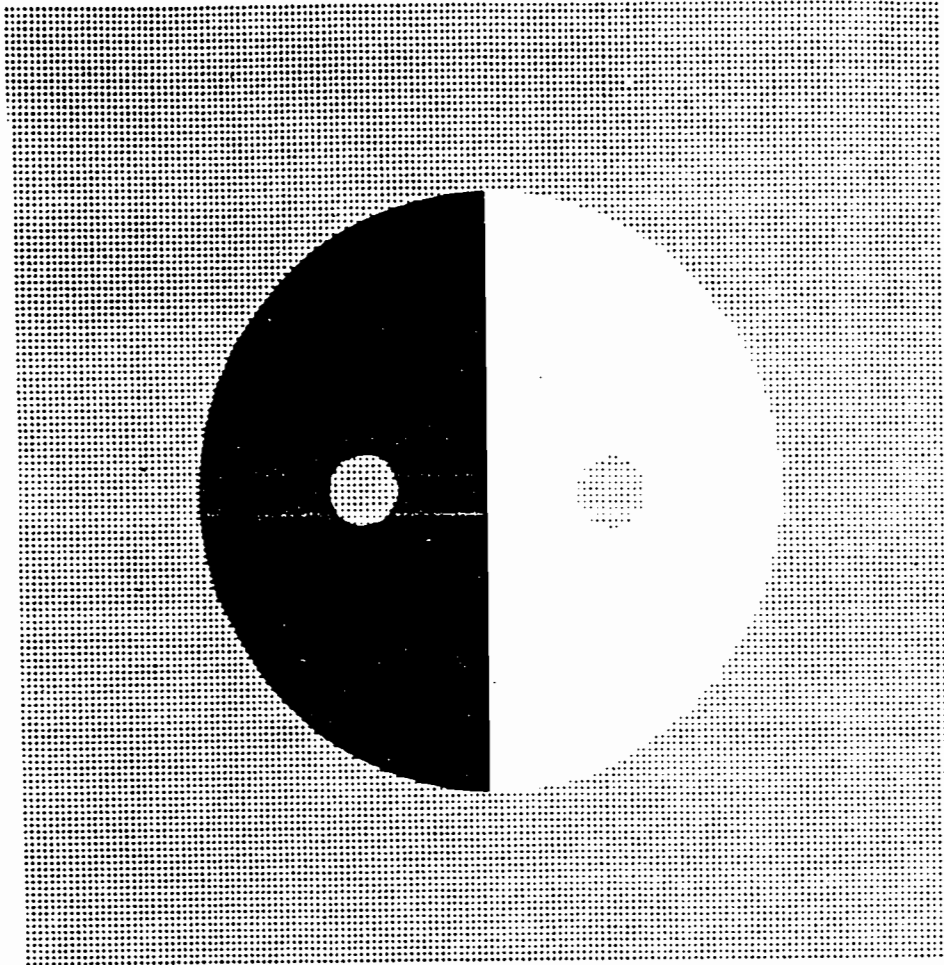


Figure 6

When the boundary of the large circle and the vertical line are stabilized on the retina, the red color outside the large circle fills-in the circle except at the small red circles. The left small circle appears as a brighter red, whereas the right small circle appears as a darker red than the filled-in red color, even though all the red regions are colored using the same red shade.

filling-in process takes place, these relative contrasts are therefore preserved as the total level of featural activity becomes more uniformly distributed around them. The net effect is the obtained percept.

8. THE LAND COLOR AXIOMS.

Land (1977) has developed a formal model to explain the remarkable brightness and color effects that he has demonstrated using McCann Mondrians. This model computes the cumulative relative contrast changes that occur when one serially traverses a path across a picture, and then averages the outcomes across many such paths. For brightness predictions, only shades of gray need to be considered. For color predictions, the computations are done with respect to more than one wavelength. This procedure leads to excellent predictions about the perceived brightnesses and colors of many visual displays. Despite these formal successes, it is difficult to physically interpret Land's serial scanning and serial averaging processes. Moreover, his formal procedure also fails to predict the correct result in a number of important cases. A simple example that illustrates the basic nature of these failures is shown in Figure 7.

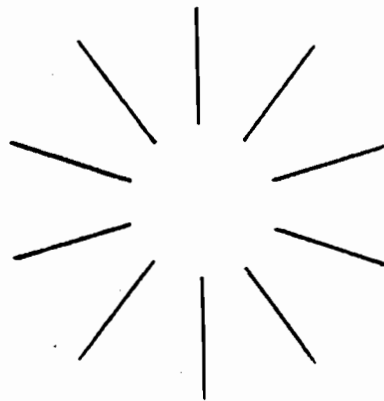


Figure 7

The fact that the illusory circle within the dark lines is brighter than its surround cannot be explained by the Land (1977) retinex theory. The present theory can explain illusory brightness differences and provides a physical interpretation of the formal operations of Land's retinex theory.

In Figure 7, an illusory circle is clearly seen at the inner tips of the radial lines. This illusory circle encloses a region of enhanced illusory brightness. No matter how many radially oriented serial scans are made between the radial lines, their average contrast change cannot predict this illusory brightness difference, because the contrast does not change at all during these scans. Including the dark radial lines within the scans does not improve this incorrect prediction, because the illusory brightness difference does not depend on the number, length, or contrast of these dark lines in a simple way. For example, the strength of the illusory brightness effect is reduced as the angle formed by the lines with the illusory circle decreases (Kennedy, 1978). This happens even if the number, length, and contrast of the lines remain unchanged. This empirical property cannot be explained by any theory that depends only on averages of local contrast changes. The effects are clearly due to the global configuration of the lines. In a similar fashion, the Land (1977) retinex theory cannot explain why rearranging the orientation of the lines in Figure 7 can change the shape of the region where illusory brightness occurs (Kennedy, 1975), or why a single line causes no illusory brightness change whereas even four such lines, if properly configured, can cause a significant illusory brightness change.

The present theory can deal with the illusory effects by showing how the contour completion process can generate illusory contours, and how feature contours can interact with illusory boundary contours to generate illusory brightness changes (Grossberg and Mingolla, 1984). Our theory also clarifies why Land's formal operations often work so well by giving these operations a physical interpretation. Serial computations of contrast changes along multiple paths are replaced by parallel computations of contrast-sensitive feature contours. Averages across serial scans are replaced by diffusive (hence, averaging) filling-in within domains defined by boundary contours. To quantitatively carry out this type of explanation for color percepts, we also need to preprocess the feature contour signals from stage M_L to stage MBD_L and from stage M_R to stage MBD_R using opponent color processes (red-green, yellow-blue) that are defined by gated dipole fields (Grossberg, 1980).

9. FILLING-IN RESONANT EXCHANGE.

All of the processing stages described above are assumed to be pre-perceptual. Only activity in the FIRE stage of Figure 1 is assumed to enter conscious perception. Signals from stage MBD_L and/or MBD_R that are capable of activating the FIRE stage are said to "lift" the preprocessed monocular patterns into the perceptual domain. I assume that only contours of the MBD_L and MBD_R activity patterns input to the FIRE stage. These contours should not be confused with the edges of a visual scene. They are non-linear statistical constructs of a rather high order by the time they generate signals from the MBD_L stage or the MBD_R stage to the FIRE stage. One needs only to remember that the contour completion process within the BME stage gives rise to the contours within MBD_L and MBD_R to appreciate this fact.

The scale-specific contour signals from the MBD_L and MBD_R stages undergo a binocular matching process in the FIRE stage that is analogous to the matching process whereby boundary contours from the M_L and M_R stages are matched at the BME stage. Only contours that survive this binocular matching process can send feedback signals back to the MBD_L and/or MBD_R stages. These local feedback signals initiate resonant exchanges between the FIRE stage and the MBD stages. These resonant exchanges selectively lift some, but not all, of the preprocessed monocular subdomains into conscious perception.

Once such a filling-in resonant exchange is initiated by a feedback exchange of matched contours, it can rapidly propagate across thousands of receptive fields within its spatial scale. Each such resonant exchange generates a standing wave of patterned activity across a subdomain of the FIRE stage. The ensemble of all such standing waves constitutes the network's percept.

The properties of these filling-in resonant exchanges lie beyond the scope of this chapter. The concept of a FIRE stage was introduced in Grossberg (1981, 1983a) and extensively analysed in Cohen and Grossberg (1983a, 1984).

10. BINOCULAR RIVALRY, STABILIZED IMAGES AND THE GANZFELD.

Three qualitative properties of these resonances are, however, easily noted. These properties further clarify how binocular rivalry and the

fading of stabilized images and of a ganzfeld occur within the network of Figure 1.

Suppose that, due to binocular matching, some left-monocular boundary contours are suppressed within the BME stage. Then these boundary contours cannot send boundary contour signals to the corresponding region of stage MBD_L . Featural activity then quickly diffuses across the network positions corresponding to these suppressed contours (Gerrits and Vendrick, 1970). Consequently, no contour signals can be emitted from these positions within the MBD_L stage to the FIRE stage. No edge matches within the FIRE stage can occur at these positions, and no feedback signals are returned from the FIRE stage to the MBD_L stage to lift the corresponding monocular subdomain into perception. Thus the subdomains whose boundary contours are suppressed within the BME stage are no longer perceived. As soon as the corresponding boundary contours win the BME binocular competition, however, these subdomains can again rapidly support the resonant lifting of their subdomain into perception at the FIRE stage.

A similar argument holds for the nonperception of a subdomain whether its boundary edges are suppressed by binocular rivalry within the BME stage, by image stabilization, or do not exist, as in a ganzfeld.

REFERENCES.

- [1] Arend, L.E., Buehler, J.N., and Lockhead, G.R., Difference information in brightness perception, *Perception and Psychophysics*, 9 (1971), 367-370.
- [2] Blake, R., and Fox, R., Binocular rivalry suppression: insensitive to spatial frequency and orientation change, *Vision Research*, 14 (1974) 687-692.
- [3] Carpenter, G.A., and Grossberg, S., Adaption and transmitter gating in vertebrate photoreceptors, *Journal of Theoretical Neurobiology*, 1 (1981), 1-42.
- [4] Cogan, A.I., Monocular sensitivity during binocular viewing, *Vision Research*, 22 (1982), 1-16.
- [5] Cohen, M.A., and Grossberg, S., Some global properties of binocular resonances: disparity matching, filling-in, and figure-ground synthesis, in P. Dodwell and T. Caelli (Eds.), *Figural synthesis*, Hillsdale, NJ., Erlbaum Associates (1983a).

- [6] Cohen, M.A., and Grossberg, S., Neural dynamics of binocular form perception, Society for Neuroscience Abstracts, 13th Annual Meeting, Boston, 353.8 (1983b).
- [7] Cohen, M.A., and Grossberg, S., The dynamics of brightness perception: a psychological and a neural theory, Submitted for publication (1984).
- [8] Gellatly, A.R.H., Perception of an illusory triangle with masked inducing figure, *Perception*, 9 (1980), 599-602.
- [9] Gerrits, H.J.M., de Haan, B., and Vendrick, A.J.H., Experiments with retinal stabilized images: relations between the observations and neural data, *Vision Research*, 6 (1966), 427-440.
- [10] Gerrits, H.J.M., and Timmerman, J.G.M.E.N., The filling-in process in patients with retinal scotomata, *Vision Research*, 9 (1969), 439-442.
- [11] Gerrits, H.J.M., and Vendrick, A.J.H., Simultaneous contrast, filling-in process and information processing in man's visual system, *Experimental Brain Research*, 11 (1970), 411-430.
- [12] Grossberg, S., Contour enhancement, short-term memory, and constancies in reverberating neural networks, *Studies in Applied Mathematics*, 52 (1973), 217-257.
- [13] Grossberg, S., How does a brain build a cognitive code?, *Psychological Review*, 87 (1980), 1-51.
- [14] Grossberg, S., Adaptive resonance in development, perception, and cognition, in S. Grossberg (Ed.), *Mathematical psychology and psychophysiology*, Providence, RI, American Mathematical Society (1981).
- [15] Grossberg, S., *Studies of mind and brain: neural principles of learning, perception, development, cognition, and motor control*, Boston, Reidel Press (1982).
- [16] Grossberg, S., The quantized geometry of visual space: the coherent computation of depth, form, and lightness, *The Behavioral and Brain Sciences*, in press (1983a).
- [17] Grossberg, S., Neural substrates of binocular form perception: filtering, matching, diffusion, and resonance, in E. Basar, H. Flohr, H. Haken, and A.J. Mandell (Eds.), *Synergetics of the brain*, New York, Springer Verlag (1983b).
- [18] Grossberg, S., and Mingolla, E., Neural dynamics of form perception: illusory figures and neon color spreading, submitted for publication (1984).
- [19] Hendrickson, A.E., Hunt, S.P., and Wu, J.-Y., Immunocytochemical localization of glutamic acid decarboxylase in monkey striate cortex. *Nature*.

- 292 (1981), 605-607.
- [20] Horton, J.C., and Hubel, D.H., Regular patchy distribution of cytochrome oxidase staining in primary visual cortex of macaque monkey, *Nature*, 292 (1981), 762-764.
- [21] Hubel, D.H., and Livingstone, M.S., Regions of poor orientation tuning coincide with patches of cytochrome oxidase staining in monkey striate cortex, *Society for Neuroscience Abstract*, 11th Annual Meeting, Los Angeles, 118.12 (1981).
- [22] Hubel, D.H., and Wiesel, T.N., Functional architecture of macaque monkey visual cortex, *Proceedings of the Royal Society of London (B)*, 198 (1977), 1-59.
- [23] Julesz, B., *Foundations of cyclonean perception*, Chicago, University of Chicago Press (1971).
- [24] Kanizsa, G., Contours without gradients or cognitive contours?, *Italian Journal of Psychology*, 1 (1974), 93-113.
- [25] Kaufman, L., *Sight and mind: an introduction to visual perception*, New York, Oxford University Press (1974).
- [26] Kennedy, J.M., Depth at an edge, coplanarity, slant depth, change in direction, and change in brightness in the production of subjective contours, *Italian Journal of Psychology*, 2 (1975), 107-123.
- [27] Kennedy, J.M., Illusory contours and the ends of lines, *Perception*, 7 (1978), 605-607.
- [28] Kennedy, J.M., Subjective contours, contrast, and assimilation, in C.F. Nodine, and D.F. Fisher (Eds.), *Perception and pictorial representation*, New York, Praeger (1979).
- [29] Kennedy, J.M., Illusory brightness and the ends of petals: change in brightness without aid of stratification or assimilation effects, *Perception*, 10 (1981), 583-585.
- [30] Kulikowski, J.J., Limit of single vision in stereopsis depends on contour sharpness, *Nature*, 275 (1978), 126-127.
- [31] Land, E.H., The retinex theory of color vision, *Scientific American*, 237 (1977), 108-128.
- [32] Livingstone, M.S., and Hubel, D.H., Thalamic innuts to cytochrome oxidase-rich regions in monkey visual cortex, *Proceedings of the National Academy of Sciences*, 79 (1982), 6098-6101.
- [33] Marr, D., and Hildreth, E., Theory of edge detection, *Proceedings of the Royal Society of London (B)*, 207 (1980), 187-217.

- [34] O'Brien, V., Contour perception, illusion, and reality, *Journal of the Optical Society of America*, 48 (1958), 112-119.
- [35] Parks, T.E., Subjective figures: some unusual concomitant brightness effects, *Perception*, 9 (1980), 239-241.
- [36] Parks, T.E., and Marks, W., Sharp-edges vs. diffuse illusory circles: the effects of varying luminance, *Perception and Psychophysics*, 33 (1983), 172-176.
- [37] Petry, S., Harbeck, A., Conway, J., and Levey, J., Stimulus determinants of brightness and distinctness of subjective contours, *Perception and Psychophysics*, 34 (1983), 169-174.
- [38] Pritchard, R.M., Stabilized images on the retina, *Scientific American*, 204 (1961), 72-78.
- [39] Pritchard, R.M., Heron, W., and Hebb, D.O., Visual perception approached by the method of stabilized images, *Canadian Journal of Psychology*, 14 (1960), 67-77.
- [40] Redies, C., and Spillman, L. The neon color effect in the Ehrenstein illusion, *Perception*, 10 (1981), 667-681.
- [41] Riggs, L.A., Ratliff, F., Cornsweet, J.C., and Cornsweet, T.N., The disappearance of steadily fixated visual test objects, *Journal of the Optical Society of America*, 43 (1953), 495-501.
- [42] van Tuijl, H.F.J.M., A new visual illusion: neonlike color spreading and complementary color induction between subjective contours, *Acta Psychologica*, 39 (1975), 441-445.
- [43] van Tuijl, H.F.J.M., and de Weert, C.M.M., Sensory conditions for the occurrence of the neon spreading illusion, *Perception*, 8 (1979), 211-215.
- [44] van Tuijl, H.F.J.M., and Leeuwenberg, E.L.J., Neon color spreading and structural information measures, *Perception and Psychophysics*, 25 (1979), 269-284.
- [45] Yarbus, A.L., *Eye movements and vision*, New York, Plenum Press (1967).

APPENDIX

This appendix describes the neural network that Cohen and Grossberg (1984) use to simulate brightness percepts. The following simulations were done on one-dimensional fields of cells. The input pattern (I_1, I_2, \dots, I_n) is transformed into the output pattern (z_1, z_2, \dots, z_n) via the following equations.

FEATURE CONTOURS

The input pattern (I_1, I_2, \dots, I_n) is transformed into feature contours via a feedforward on-center off-surround network of cells undergoing shunting, or membrane equation, interactions. The activity, or potential x_i , of the i^{th} cell in a feature contour pattern is

$$\frac{d}{dt} x_i = -Ax_i + (B - x_i) \sum_{k=1}^n I_k C_{ki} - (x_i + D) \sum_{k=1}^n I_k E_{ki}. \quad (1)$$

Both the on-center coefficients C_{ki} and the off-surround coefficients E_{ki} are Gaussian functions of intercellular distance; viz., $C_{ki} = C \exp[-\mu(i-k)^2]$ and $E_{ki} = E \exp[-\nu(i-k)^2]$, where $\mu > \nu$. System (1) is assumed to react much more quickly than the diffusive filling-in process. Hence, we assume that each x_i is in approximate equilibrium with respect to the input pattern. At equilibrium, $\frac{d}{dt} x_i = 0$ and

$$x_i = \frac{\sum_{k=1}^n (RC_{ki} - DE_{ki})I_k}{A + \sum_{k=1}^n (C_{ki} + E_{ki})I_k}. \quad (2)$$

The activity pattern (x_1, x_2, \dots, x_n) is sensitive to both the amount and the direction of contrast in edges of the input pattern (Grossberg (1983a)). These feature contour activities generate inputs of the form

$$F_i = \frac{x_i}{1 + \alpha S_i}$$

to the diffusive filling-in process. The inhibitory term S_i is defined by the boundary contour process in equation (7) below.

BOUNDARY CONTOURS

The input pattern (I_1, I_2, \dots, I_n) also activates the boundary contour process, which we represent as a feedforward on-center off-surround networking undergoing shunting interactions. This simplified view of the

boundary contour process is permissible in the present simulations because the simulations, being one-dimensional and monocular, do not need to account for orientational tuning or binocular matching. Since the simulations do not probe the dynamics of illusory contour formation, the contour completion process can also be ignored.

As in equation (2), the input pattern rapidly gives rise to an activity pattern

$$y_i = \frac{\sum_{k=1}^n (\tilde{B}\tilde{C}_{ki} - \tilde{D}\tilde{E}_{ki})I_k}{\tilde{A} + \sum_{k=1}^n (\tilde{C}_{ki} + \tilde{E}_{ki})I_k} \quad (4)$$

where $\tilde{C}_{ki} = \tilde{C} \exp[-\tilde{\mu}(i-k)^2]$ and $\tilde{E}_{ki} = \tilde{E} \exp[-\tilde{\nu}(i-k)^2]$. The Gaussian coefficients $\tilde{\mu}$ and $\tilde{\nu}$ are, however, larger than the Gaussian coefficients μ and ν , because boundary contours are narrower than feature contours.

The activity pattern (y_1, y_2, \dots, y_n) is sensitive both to the direction and amount of contrast in the input pattern (I_1, I_2, \dots, I_n) . The sensitivity to the direction of contrast is progressively eliminated by the following operations. Let the output signal from BME to MBD that is elicited by activity y_i equal $f(y_i)$, where $f(w)$ is a sigmoid signal of the rectified part of y_i ; viz.,

$$f(y_i) = \frac{\beta \{ [y_i]^+ \}^\gamma}{1 + \delta \{ [y_i]^+ \}^\gamma} \quad (6)$$

The notation $[w]^+ = \max(w, 0)$ and $\gamma > 1$. The output signals $f(y_i)$ are spatially distributed before influencing cell compartments of the cell syncytium. The total signal to the i^{th} cell compartment due to the activity pattern (y_1, y_2, \dots, y_n) is

$$S_i = \sum_{k=1}^n G_{ik} f(y_k), \quad (7)$$

where G_{ik} is a Gaussian function of intercellular distance; viz., $G_{ik} = G \exp[-\omega(i-k)^2]$. This Gaussian fall-off (ω) is less narrow than the boundary contour process ($\tilde{\mu}, \tilde{\nu}$), but more narrow than the feature contour process (μ, ν).

DIFFUSIVE FILLING-IN

The activity z_i of the i^{th} cellular compartment of the cellular syncytium

obeys the nonlinear diffusion equation

$$\frac{d}{dt} z_i = -Hz_i + J_{i+1,i}(z_{i+1} - z_i) + J_{i-1,i}(z_{i-1} - z_i) + F_i, \quad (8)$$

where the input F_i is defined by equation (3). The diffusion coefficients $J_{i+1,i}$ and $J_{i-1,i}$ are determined by boundary contour signals according to equations of the form

$$J_{i+1,i} = \frac{\lambda}{1 + \kappa [S_{i+1} - \Gamma]^+ + \kappa [S_i - \Gamma]^+} \quad (9)$$

and

$$J_{i-1,i} = \frac{\lambda}{1 + \kappa [S_{i-1} - \Gamma]^+ + \kappa [S_i - \Gamma]^+}, \quad (10)$$

where the threshold $\Gamma > 0$. Thus an increase in the boundary signal S_i decreases both diffusion coefficients $J_{i+1,i}$ and $J_{i-1,i}$. The feature contour signal F_i also decreases when the boundary signal S_i increases. In all the equations (3), (9), and (10), the inhibitory effects of boundary signals S_i on cell compartment membranes act via shunting inhibition. A positive threshold Γ occurs in (9) and (10), but not in (3), because we assume that the intercompartmental membranes that regulate diffusion of activity between compartments are less accessible to the signals S_i than are the exterior surface membranes that bound the cellular syncytium.

Supported in part by the Air Force Office of Scientific Research (AFOSR: 82-0148).