Neural Networks, Vol. 4, pp. 723–742, 1991
Printed in the USA.

0893-6080/91 $3.00 + .00

ORIGINAL CONTRIBUTION

A Neural Network Architecture for Figure-Ground Separation of Connected Scenic Figures

STEPHEN GROSSBERG[1] AND LONCE WYSE[2]

Boston University

(*Received:* 11 *December* 1990; *revised and accepted* 15 *May* 1991)

Abstract—*A neural network model, called an* FBF *network, is proposed for automatic parallel separation of multiple image figures from each other and their backgrounds in noisy gray-scale or multicolored images. The figures can then be processed in parallel by an array of self-organizing Adaptive Resonance Theory* (ART) *neural networks for automatic target recognition. An* FBF *network can automatically separate the disconnected but interleaved spirals that Minsky and Papert introduced in their book* ***Perceptrons****. The network's design also clarifies why humans cannot rapidly separate interleaved spirals, yet can rapidly detect conjunctions of disparity and color, or of disparity and motion, that distinguish target figures from surrounding distractors. Figure-ground separation is accomplished by iterating operations of a Feature Contour System* (FCS) *and a Boundary Contour System* (BCS) *in the order* FCS-BCS-FCS, *hence the term* FBF. *The* FCS *operations include the use of nonlinear shunting networks to compensate for variable illumination and nonlinear diffusion networks to control filling-in. A key new feature of an* FBF *network is the use of filling-in for figure-ground separation. The* BCS *operations include oriented filters joined to competitive and cooperative interactions designed to detect, regularize, and complete boundaries in up to 50% noise, while suppressing the noise. A modified* CORT-X *filter is described, which uses both on-cells and off-cells to generate a boundary segmentation from a noisy image.*

Keywords—Vision, Sensor fusion, Figure-ground separation, Segmentation, Neural network, Pattern recognition, Filling-in, Visual cortex.

1. INTRODUCTION: AUTOMATIC FIGURE-GROUND SEPARATION AND GENERAL-PURPOSE VISION

An important stage in the perception and recognition of objects is the process whereby a figure, or object, in a scene is separated from other figures and background clutter. This is called the stage of *figure-ground separation*. Whereas knowledge about a figure may facilitate its separation, such knowledge is clearly not *necessary* for biological vision systems to carry out figure-ground separation. Experiences abound of unfamiliar figures that "pop out" from their backgrounds before they ever enter our corpus of learned knowledge about the world. The fact that figure-ground separation can occur even for unfamiliar figures contributes to the general-purpose nature of biological vision, which can process both unfamiliar and familiar scenes, and does not require prior instruction about an environment in order to operate effectively.

The present article describes a new type of system that is capable of automatic figure-ground separation. This process separates scenic figures whose emergent boundary segmentations (defined below) surround a connected region. As a result of this property, such a system can automatically distinguish between connected and disconnected spirals (Figure 1), a benchmark that gained fame through its emphasis in the book by Minsky and Papert (1969, 1988) on perceptrons. Why the present biologically motivated algorithm can distinguish these figures in a way that humans cannot is discussed in Section 9.

[1] Supported in part by the Air Force Office of Scientific Research (AFOSR 90-0175), the Army Research Office (ARO DAAL-03-88-K0088), DARPA (AFOSR 90-0083), and Hughes Research Laboratories (S1-804481-D and S1-903136).

[2] Supported in part by the American Society for Engineering Education and Hughes Research Laboratories (S1-804481-D).

Acknowledgements: The authors wish to thank Cynthia E. Bradford, Carol Yanakakis Jefferson, and Diana Meyers for their valuable assistance in the preparation of the manuscript.

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

$2S$, $3S$, . . . , MS of speeds, within some tolerance ΔS; and an $N \times M$ matrix of intersection images can be generated which extract the figure at each combination of distance iD and speed jS within this tolerance. Then $N \times M$ copies of the architecture illustrated in Figure 2 can work simultaneously in parallel to recognize these figures independent of their position, orientation, and size, by preprocessing each of the $N \times M$ laser radar images with CORT-X and invariant filters before they activate a parallel array of ART 2 architectures.

4. FIGURE-GROUND SEPARATION IN FACADE THEORY

In biological vision, the retinal detectors do not, in themselves, separate figure from ground. One task of neural network research is to suggest how subsequent network processes which are activated by the retinal detectors may generate this competence. Grossberg (1987; reprinted in Grossberg, 1988) has, for example, introduced a neural theory of binocular vision in which figural components of an image are separated from one another into distinct network levels, or slabs. A macrocircuit of this theory is shown in Figure 3, where the vertically hatched boxes form part of the Boundary Contour System (BCS) and the dotted boxes form part of the Feature Contour System (FCS).

The theory describes how parallel and hierarchical interactions between the BCS and FCS generate a multiplexed, multiple-scale representation, called a FACADE representation, of the scene's Form-And-Color-And-DEpth. Within this representation, figural components which encode distinctive combinations of features, such as prescribed combinations of color, depth, and size, are segregated from one another into different network levels. These levels, in turn, activate subsequent stages of network processing that are designed for visual object recognition (Figure 4).

For present purposes, the main insight that may be derived from FACADE theory is that a properly designed sequence of FCS-BCS-FCS operations can separate figure from ground. Henceforth all networks that use this strategy will be called *FBF networks*. To arrive at this insight, we consider two different competences of FACADE theory: discounting variable illumination and multidimensional fusion.

5. DISCOUNTING VARIABLE ILLUMINATION AND FILLING-IN

The theory provides an explanation of how variable illumination conditions are automatically discounted and used to trigger a filling-in process that completes a surface representation over image regions which

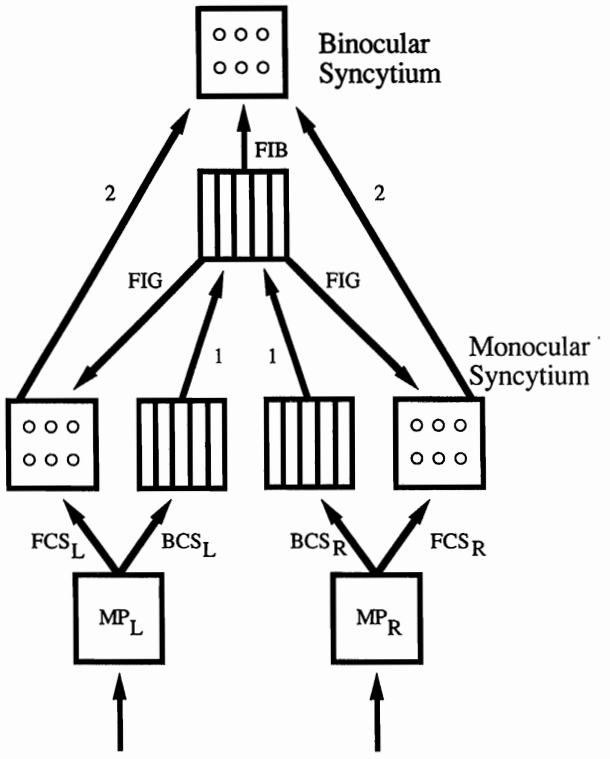

FIGURE 3. Macrocircuit of monocular and binocular interactions within the Boundary Contour System (BCS) and the Feature Contour System (FCS): Left and right monocular preprocessing stages (MP_L and MP_R) send parallel monocular inputs to the BCS (boxes with vertical lines) and the FCS (boxes with three pairs of circles). The monocular BCS_L and BCS_R interact via bottom-up pathways labelled 1 to generate a coherent binocular boundary segmentation. This segmentation generates output signals called filling-in generators (FIGs) and filling-in barriers (FIBs). The FIGs input to the monocular filling-in domains, or syncytia, of the FCS. The FIBs input to the binocular filling-in domains, or syncytia, of the FCS. Inputs from the MP stages interact with FIGs at the monocular syncytia where they select those monocular FC signals that are binocularly consistent. The selected FC signals are carried by the pathways labelled 2 to the binocular syncytia, where they interact with FIB signals from the BCS to generate a multiple-scale representation of Form-And-Color-And-Depth within the binocular syncytia.

are suppressed by the discounting process. A monocular version of this process was modelled by Cohen and Grossberg (1984) and Grossberg and Todorović (1988) to explain data about monocular brightness perception. This monocular model is schematized in Figure 5.

In this model, variable illumination conditions are discounted by a shunting on-center off-surround network (Level 2), which constitutes the first FCS stage. Image regions of high relative contrast are amplified and regions of low relative contrast are attenuated as a consequence of the discounting process. The shunting network, in turn, topographically activates a filling-in network (Level 6) which constitutes the

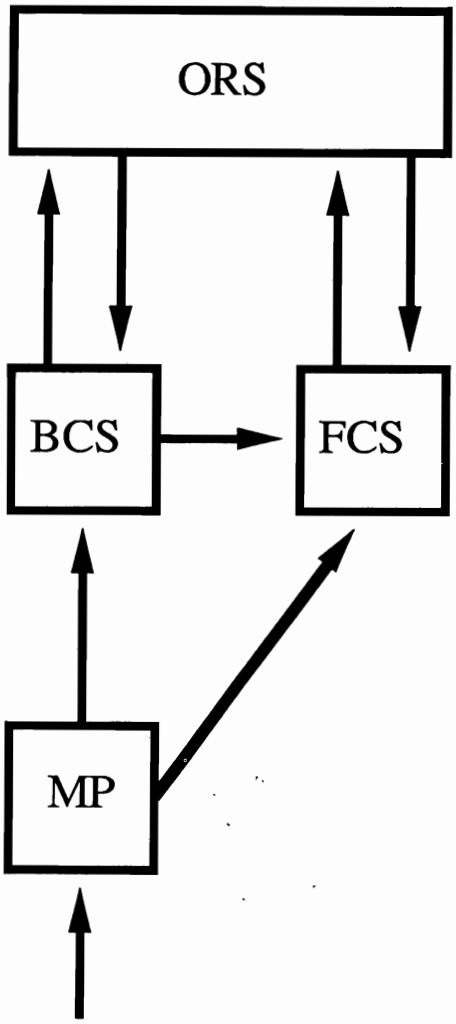

FIGURE 4. A miacrocircuit of processing stages: Monocular preprocessed signals (MP) are sent independently to both the BCS and the FCS. The BCS preattentively generates coherent boundary structures from these MP signals. These structures send outputs to both the FCS and the Object Recognition System (ORS). The ORS, in turn, rapidly sends top-down learned template signals, or expectations, to the BCS. These template signals can modify the preattentively completed boundary structures using learned, attentive 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. The completed FCS representation, in turn, also interacts with the ORS.

second FCS state. This filling-in network uses a nonlinear diffusion process to complete a brightness representation over both the amplified and attentuated image regions.

Filling-in is restricted to compartments whose boundaries are defined by topographic signals from the BCS (Levels 3–5). The BCS converts signals from the first FCS stage (Level 2) into a boundary segmentation one of whose functions is to trigger a BCS-FCS interaction that contains the filling-in process at the second FCS stage. The result of this FBF interaction is a surface representation of featural quality, such as brightness or color, that is relatively uncontaminated by illumination conditions.

Figures 6 and 7 summarize computer simulations of Grossberg and Todorović (1988) that illustrate how the illuminant is discounted in FCS Level 2, and how the subsequent BCS-FCS interaction at Level 6 controls the filling-in process that completes the brightness representation. The image schematized in Figure 6a is called a McCann–Mondrian (Land, 1977). It is a patchwork of rectangular regions each with a different luminance. The image is uniformly illuminated. In Figure 6a, each circle's radius is proportional to the luminance registered by a network node located at the center of the circle.

Figure 6b represents the activation pattern of the shunting on-center off-surround network at Level 2, Figure 6c represents the boundary representation at Level 5, and Figure 6d represents the filled-in representation at Level 6. The diffusion spatially averages the activation differences of Figure 6b within the compartments defined in Figure 6c.

In Figure 7a, the same image depicted in Figure 6a is illuminated from the lower right corner. Because the shunting on-center off-surround network at Level 2 effectively discounts the illuminant, the Level 2 activation patterns in Figs. 7b and 6b are essentially identical. Hence, the subsequent boundary patterns (Figs. 7c and 6c) and filled-in patterns (Figs. 7d and 6d) are also essentially identical.

Note in Figure 7d that the brightness, or activation level, of the square region in the upper left corner is larger than that of the square region in the lower right corner. In contrast, in Figure 7a, the luminance, or activation level, of the upper left corner is smaller than that of the square region in the lower right corner. This luminance-to-brightness reversal compensates for the larger intensities of illumination in the lower right region.

6. MULTIDIMENSIONAL FUSION

An FBF interaction may also be used to represent scenic form, notably scenic surface properties, and to separate figure from ground. This type of Form-and-Color fusion is achieved by suitably embedding an FBF interaction into a binocular version of the theory. In the binocular theory, depth is also encoded into the representation; hence the mnemonic Form-And-Color-And-DEpth, or FACADE.

The binocular version of the theory suggests how monocular image data from both eyes can be selectively processed so that only the binocularly consistent monocular data from each eye is allowed to

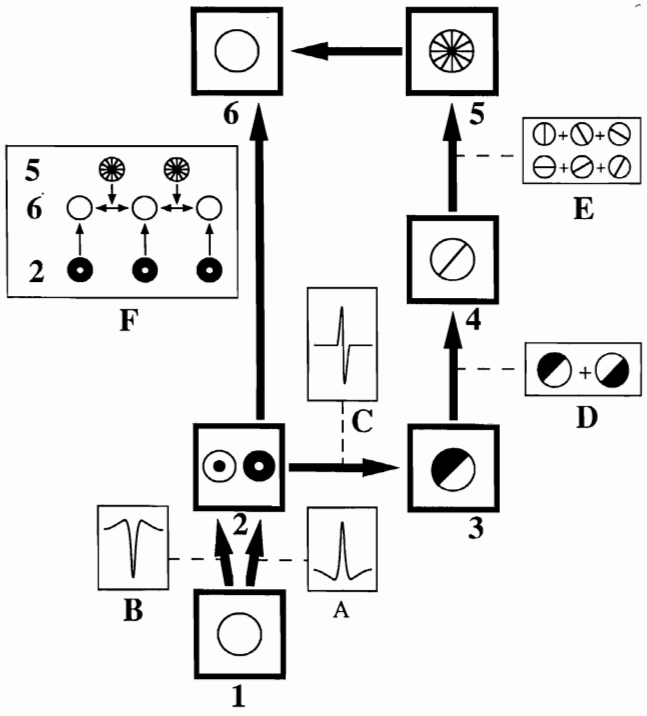

FIGURE 5. Model of how the Feature Contour System discounts variable illuminants and regulates featural filling-in: The thick-bordered rectangles numbered from 1 to 6 correspond to the levels of the system. The symbols inside the rectangles are graphical mnemonics for the types of computational units residing at the corresponding model level. The arrows depict the interconnections between the levels. The thin-bordered rectangles coded by letters A through E represent the type of processing between pairs of levels. Inset F illustrates how the activity at Level 6 is modulated by outputs from Level 2 and Level 5. This simplified model directly extracts boundaries from image contrasts, rather than generating emergent segmentations from image contrasts. The model's key elements concern how the Level 2 network of shunting on-center off-surround interactions discounts variable illuminants while extracting Feature Contour signals, and how Level 5 fills-in these signals via a nonlinear diffusion process within the compartments defined by Boundary Contour System output signals.

influence the FACADE representation. Figure 3 schematizes the network that is used. In it, pairs of monocular BCS signals are derived from monocular FCS patterns that discount the illuminant. These monocular BCS signals interact to form the binocular boundary segmentation along the pathways labelled 1. The boundary segmentation regularizes and completes all the boundary data, across multiple spatial scales, that are capable of being binocularly fused. This binocular boundary segmentation sends topographic signals, called *filling-in barriers* (FIBS), to the monocular filling-in networks within the FCS that are generated by the left eye (FCS_L) and the right eye (FCS_R). This BCS-FCS interaction allows only those monocular featural data from (FCS_L) and (FCS_R) that are consistent with the binocular boundary segmentation to fill-in and generate topographic output signals, labelled 2, to the binocular FCS stage. This BCS-FCS interaction carries out a type of figure-ground separation, since only those FCS regions can generate output signals that are surrounded by binocular FIGS from the BCS.

The binocular FCS stage is called the binocular syncytium. In the binocular syncytium, the selected monocular FCS signals from both eyes interact once again with the binocular BCS signals. Here the FCS signals again activate a filling-in process within the compartments that are defined by the BCS signals. These BCS signals are thus called *filling-in barriers* (FIBS). The FACADE representation that is generated within the binocular syncytium groups distinctive combinations of features into figures within separate network levels, or slabs. These slabs thereupon send adaptively filtered signals to subsequent processing levels for purposes of visual object recognition.

Within such a biological theory of vision, the process of separating figures into different slabs exploits the fact that the retina contains photodetectors with different spectral sensivities; for example, three types of retinal cones and one type of retinal rod. The theory suggests how figures may be spatially parsed into separate slabs based, in part, upon the distinctive colors that are derived from these detectors. In addition, there exist multiple spatial scales and multiple binocular disparity computations within the theory that further parse figural components into separate slabs based upon different size–disparity correlations (Grossberg, 1987; Grossberg & Marshall, 1989).

Thus, although FACADE theory uses a FBF network—actually an FBFBF network—to achieve figure-ground separation, this network exploits the existence of multiple detector types and multiple-scale reactions to these detector types to carry out the separation. It remains to consider how well figural components may be separated into separate slabs even if only a single general-purpose detector is used that, in itself, cannot separate figure from ground. For example, how can individual figures be separated from the cluttered ground of a picture taken with a camera that uses monochromatic film? We now show how a suitably designed FBF network can accomplish this task for at least certain classes of images. Section 7 provides an intuitive description of network stages and their effects. Section 8 describes network equations and parameters.

7. FIGURE-GROUND SEPARATION BY A MONOCHROMATIC FBF NETWORK: THE DYE-INJECTED FBF

Because multiple detectors are not available in the monochromatic case, we assume that the filling-in process is activated by *internally* generated input sources. In particular, the network "paints" each connected figure of the image by using an internally generated "dye" that triggers the filling-in of that figure. This heuristic is realized by using the following procedure, which was briefly reported in Grossberg and Wyse (1990):

Step 1 (Discount the illuminant). At the first FCS stage, variable illumination conditions are discounted both by a shunting on-center/off-surround network ("ON-C") and an off-center/on-surround network ("OFF-C"), operating in parallel. The ON-C network has a zero baseline activity (decays to zero if there is no signal within its entire receptive field) while the OFF-C network has a positive baseline activity. Because the OFF-C filter has a positive baseline activity and is inhibited by positive signal values, the network performs an image inversion which is normalized by the shunting interactions (Figure 8).

The ON-C and OFF-C networks operate in a complementary fashion. Along a straight boundary between a region of strong signal and one of no signal,

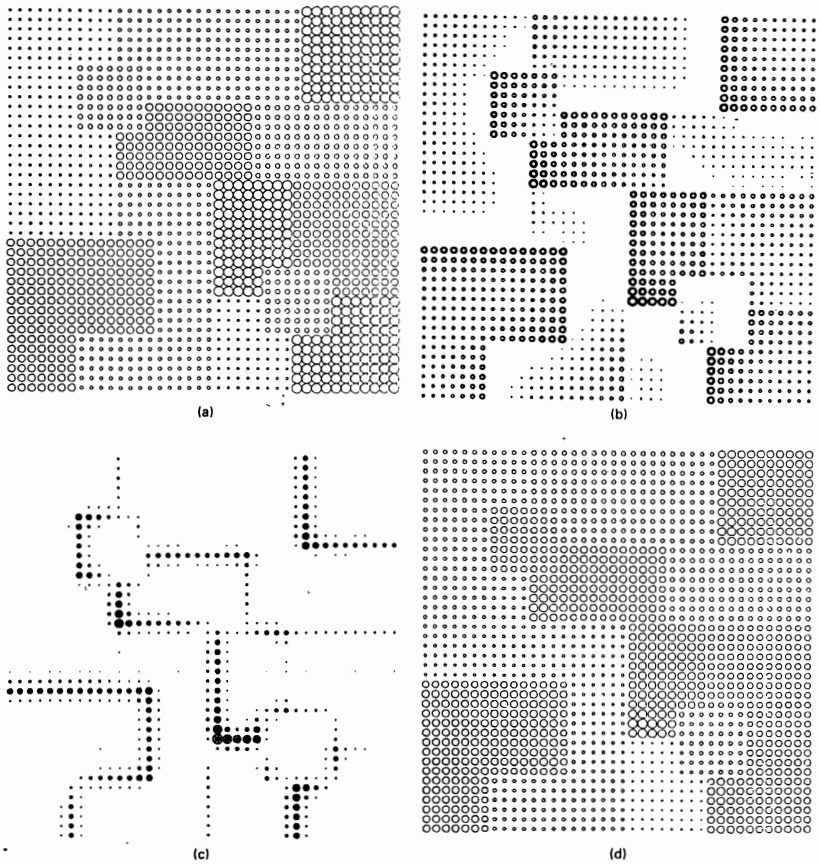

FIGURE 6. The evenly illuminated Mondrian. (a) The stimulus distribution consists of 13 homogeneous polygons with 4 luminance levels. Note that the square in the upper left portion of the stimulus has the same luminance as the square in the lower right portion. However, the average luminance of the regions surrounding the lower square is higher than the corresponding average luminance for the upper square; (b) The on-cell distribution. The amount of on-cell activity within the upper square is higher than within the lower square; (c) The Boundary Contour output; and (d) The filled-in syncytium. The upper square is correctly predicted to look brighter than the lower square.

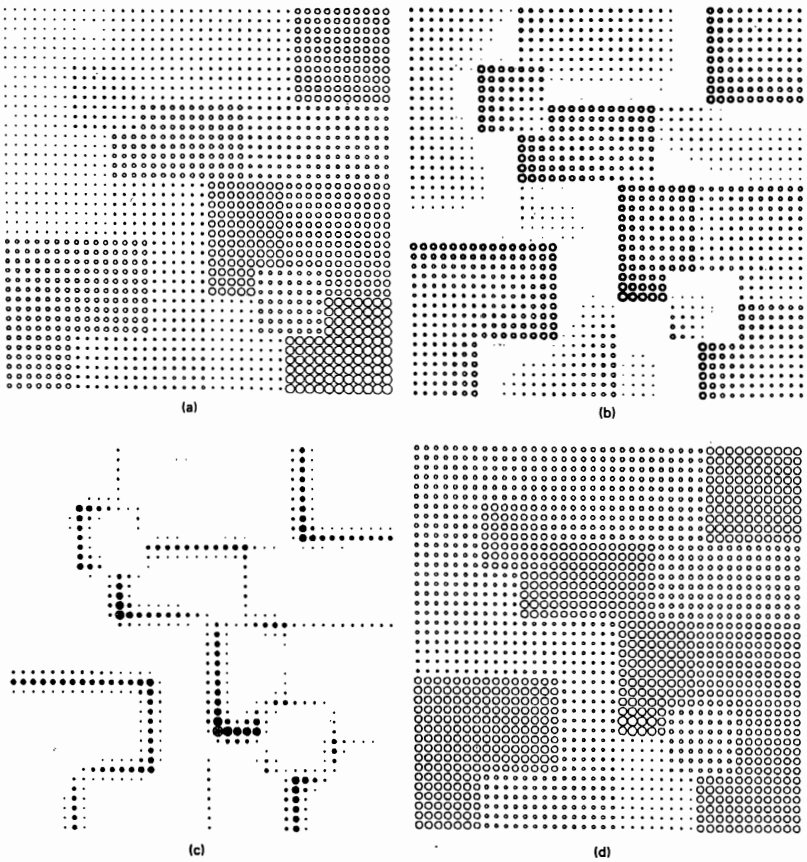

FIGURE 7. The unevenly illuminated Mondrian. (a) The stimulus distribution simulates the transformation of Figure 6a caused by the presence of a light source whose intensity decreases linearly from the lower right corner toward the upper left corner of the stimulus. The lower square is now more luminant than the upper square; (b) The on-cell distribution; (c) The Boundary Contour output; and (d) The filled-in syncytium. Figures 6b, 6c, and 6d are very similar to the corresponding figures for the evenly illuminated Mondrian (Figure 6). This illustrates the model's discounting of the illuminant. In addition, the upper square is still predicted to appear brighter than the lower square.

both types of networks respond similarly by enhancing the contrast. At a concave corner of a high signal region, the ON-C network responds more strongly than the OFF-C network, while at a convex corner the converse is true (Grossberg & Todorović, 1988, Figure 6). This ON-OFF Complementarity Property also plays an important role in noise suppression when it interacts with the CORT-X filter, as the next section explains.

Step 2 (CORT-X Filter). The ON-C and OFF-C shunted images are transformed by a CORT-X Filter into a boundary representation. The CORT-X boundary filter as developed for binary monochromatic images has been described in detail elsewhere (Carpenter, Grossberg & Mehanian, 1989). Some modifications have been made herein so that analogue monochromatic images can be processed (Section 8). This modified filter is called CORT-X 2. The rho-space model of Walters (1987) is an alternative oriented filter that shares a number of properties with the CORT-X filter. A brief description and motivation for the filter stages follows.

Each processing layer has the same number of cells as pixels in the image. The architecture is completely feedforward. Cells at a given layer have input fields ("IFs") that integrate over an area in the previous layer local to its position in the field. Two separate scales (input field sizes) are used in parallel in the early stages of processing and are subsequently combined to take advantage of the best of their respective processing capabilities. The term *input field*, or in-field, is used instead of *receptive field* because the latter term from neurophysiology typically refers to the region at the first processing layer that influences the activity of a cell at any subsequent layer. Our layer-by-layer analysis of scale sizes requires a more microscopic analysis of network geometry.

The model's first stage, called the *simple cell layer*, is an oriented contrast detector that is sensitive to the orientation, amount, direction, and spatial scale of image contrast at a given image location. The orientation sensitivity is the result of an elliptically

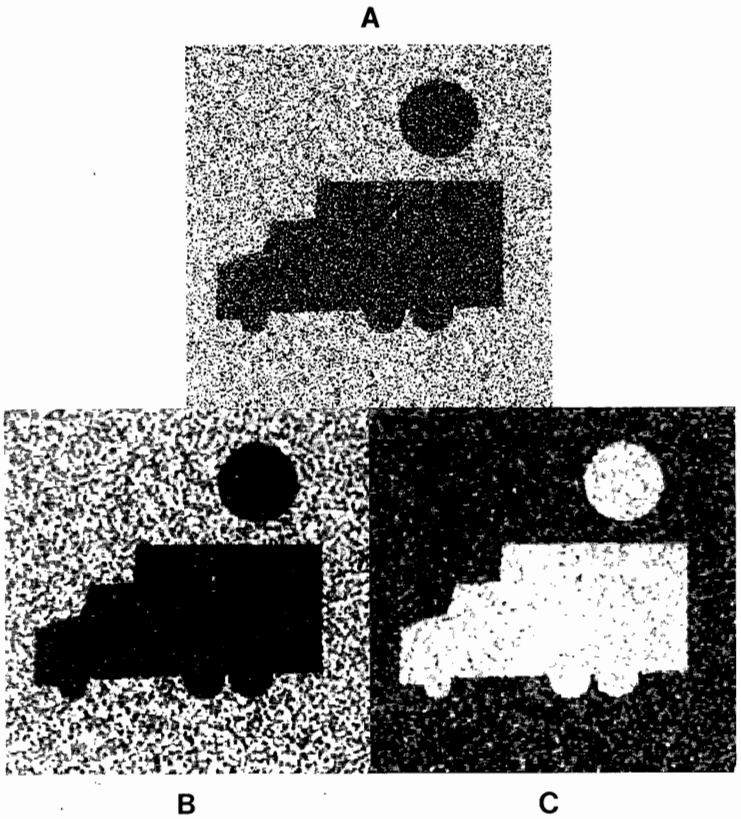

FIGURE 8. (a) The original figure in 50% noise (half the pixels are random); (b) The result of the ON-C shunting filter; and (c) The result of the OFF-C shunting filter. The Appendix explains the way in which image intensities were normalized.

shaped IF. Several fields of different orientations each operate in parallel at each position. The sensitivity to amount and direction of contrast is produced by exciting the cell with signal present in one half of its IF, inhibiting its output with some proportion of the signal in the other half, and then thresholding the result (Figure 9). The result is a half-wave rectification of the signal.

The ON-C shunting network and the OFF-C shunting network each input to separate networks of simple cells at each receptive field size. Thus the ON-C and OFF-C networks together activate four networks of simple cells.

The outputs of these parallel simple cell networks are next combined at each position (Figure 9). This second stage of the CORT-X filter, called the *complex cell layer*, is sensitive to the orientation, amount, and spatial scale of the contrast of the image at a given point, but not to the direction-of-contrast. This is achieved by summing the outputs of all like-oriented simple cells at each position, including cells that are sensitive to opposite direction-of-contrast and that receive inputs from either the ON-C or OFF-C shunting networks. Two networks of complex cells sensitive to the two scale sizes are generated in this way. Adding the half-wave rectified outputs from

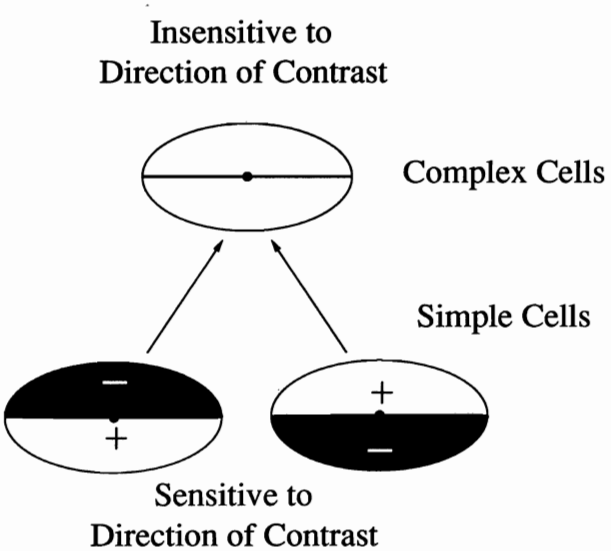

FIGURE 9. The simple cell and complex cell layers that process each of the two shunting network images in parallel. A horizontally oriented set of cells is shown. In Figs. 10 and 11, only the outputs of horizontally oriented cells are displayed.

FIGURE 10. (a) Small scale horizontal complex cell output from the ON-C image of Figure 8b; (b) Small scale horizontal complex cell output from the OFF-C image of Figure 8c; (c) Large-scale horizontal complex cell output from the ON-C image of Figure 8b; and (d) Large-scale horizontal complex cell output from the OFF-C image of Figure 8c.

pairs of simple cells that are sensitive to opposite direction-of-contrast has the same net effect as full-wave rectification.

The complementarity property helps to suppress noise when the ON-C and OFF-C outputs are processed separately by their own parallel networks of simple cells and complex cells. To a first approximation, the contrast detectors respond to the ratio of signal between each half of an oriented receptive field. As such, a small amount of noise signal against a background of no signal would effect the contrast detectors more than the same amount of "drop out" noise against a background of strong signal. The inversion of the image performed by the OFF-C filter changes the direction of contrast between any noise and its signal background. Thus, the noise will be disruptive in only one of the two parallel networks of simple cells and complex cells, while actual region boundaries will be strongly detected in both (Figure 10).

The output from both networks of contrast detectors are then summed, an operation which takes advantage of the ON-OFF Complementarity Property. First it roughly yields equal responses to both

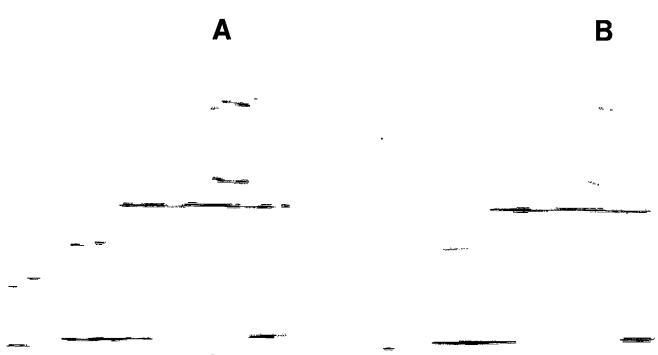

FIGURE 11. (a) ON-OFF small-scale horizontal complex cell rersponse derived by adding Figs. 10a and 10b; and (b) ON-OFF large-scale horizontal complex cell response derived by adding Figs. 10c and 10d.

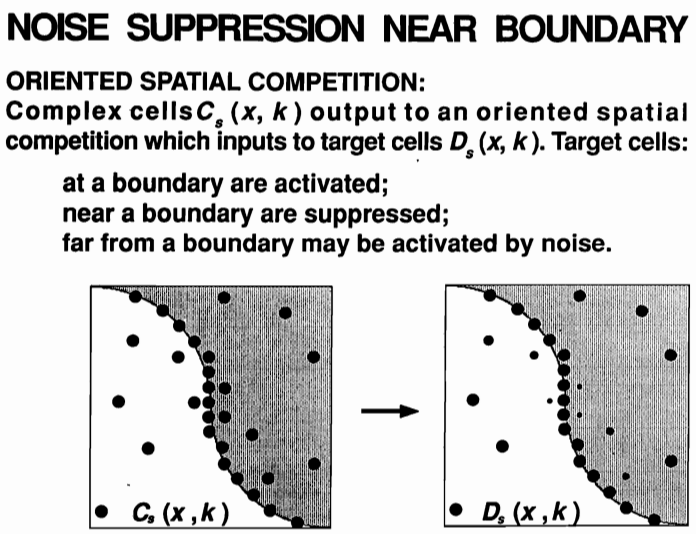

FIGURE 12. Oriented spatial competition inhibits noise pixels near the boundary. Reprinted from Carpenter, Grossberg, & Mehanian (1989) with permission.

concave and convex curvatures. Second, since noise in a given region is suppressed in one or the other contrast-filtered shunting outputs, while boundaries one are strong in both, the summation strengthens the boundary signals relative to the noise (Figure 11).

Figure 11 demonstrates that the two scales also exhibit another type of complementary processing capabilities (Carpenter, Grossberg, & Mehanian, 1989). The smaller scale filter does a better job of boundary localization than the larger scale filter, especially at positions of high boundary curvature, whereas the larger filter does a better job of noise suppression and boundary completion. Canny (1986) has suggested how a single spatial scale can trade off between these virtues, but notes that "we cannot improve both simultaneously" (p. 684). The CORT-X family of models suggests a strategy whereby two or more scales can be combined to realize the best features of each. In particular, the large-scale filter achieves good noise suppression far from the image boundaries, but not within the radius of large scale IFs near these boundaries. The small scale filter is relatively poor at noise suppression anywhere. The next stage of filtering is designed to control this noise near boundaries. Each complex cell excites the hypercomplex cell at its position and orientation in the next level, called the first competitive stage, while inhibiting hypercomplex cells at nearby positions that are not colinear with its axis of symmetry (Figure 12).

The next level, the second competitive stage, sharpens the activation pattern across orientations at each position and scale. In particular, at each position and scale, that cell is chosen whose orientation receives the maximal input from the first competitive stage (Figure 13).

The final operations include cooperative interactions between both filter sizes that select their desirable properties and eliminate their undesirable ones. The small scale's ability to localize boundaries and the large scale's ability to suppress noise and complete gaps in the boundaries are maintained by these cooperative interactions (Figure 14a).

Gaps in the boundary become more likely as the

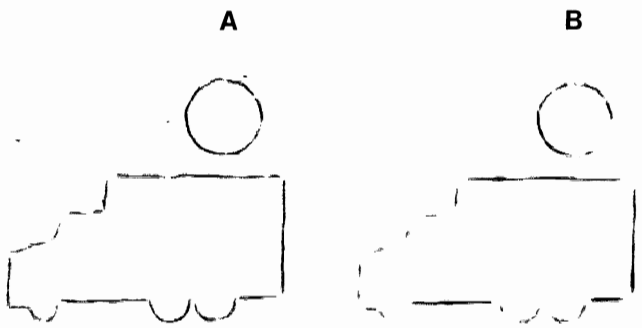

FIGURE 13. Output from the second competitive stage due to competition between orientations at each location. (a) Small scale; and (b) Large scale.

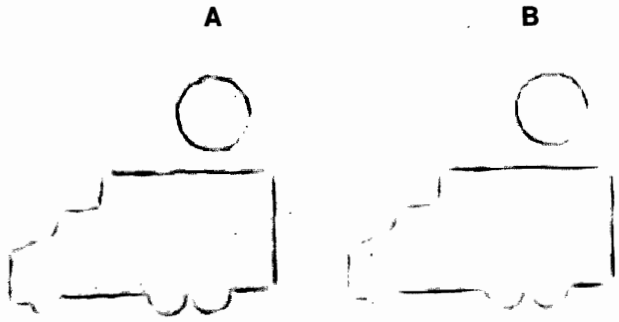

FIGURE 14. (a) Unoriented cooperation between both scales; and (b) Oriented cooperation within the large scale.

noise level of the original image increases. To overcome this problem, the cooperative level also includes oriented long-range interactions among the selected maximal responses in Figure 13. Because of these long-range interactions, an inactive cell can become active if enough cells of a particular orientation at the previous level are active on both sides of the cell along the oriented axis of this cell (Figure 14b).

The final output of the CORT-X filter is the sum of the combined-scales image and the completed-gap image (Figure 15).

Step 3 (Filling-In). The output from the CORT-X filter is topographically mapped into M filling-in networks F_m, $i = 1, 2, \ldots, M$. In the Grossberg and Todorović (1988) article, the signals that trigger filling-in are generated by the image (Section 5). In the present application, they are generated by input sources that lie within the network. Moreover, each internally generated input is delivered to its filling-in network at a different position.

Imagine, for definiteness, an $n \times n$ grid of $M = n^2$ nodes laid out over the boundary image generated by the CORT-X Filter. Each filling-in network F_m is associated with a different grid point where it will receive a featural "dye" injection into its copy of the boundary image. The injection then spreads unimpeded where there is no boundary signal, but does not spread through points where a boundary signal exists (Section 8). Thus, each injection fills-in the connected figure that surrounds the injection point (Figure 16). If the grid is dense enough, then all connected figural components will receive an injection within its boundary in at least one filling-in network F_m. This process is easy to replicate in large numbers because all the networks are identical except for the different, but regular, locations of the injected inputs, and all can operate independently and asychronously in parallel.

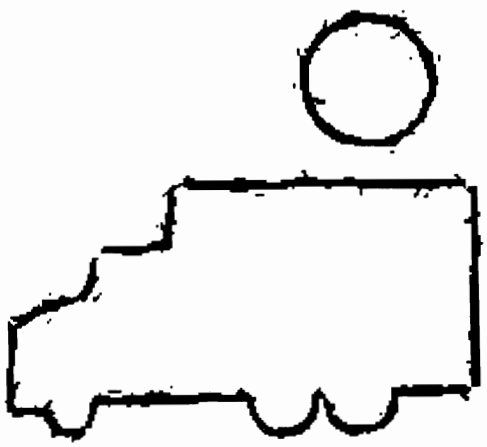

FIGURE 15. Final CORT-X filter output.

Step 4 (Figure-Ground Separation). Each filling-in network feeds its activation pattern in parallel to another pair of shunting networks, one on-center/off-surround (ON-C) and one off-center/on surround (OFF-C). This completes the second "F" operation in the FBF model. Because of their contrast enhancing and ratio-processing properties, these filters amplify the filled-in activity near figural boundaries while tending to suppress the low-contrast regions generated by the spreading of activation across the interiors of figures and background regions.

Consider first the case where there are no holes in a boundary at which the spreading activation could leak through. If the filling-in process has been given enough time to reach all of the enclosing boundary signals, then the ON-C filter produces an output only inside the enclosing boundaries, due to the injected activity (Figure 17E). The OFF-C filter produces signal only outside the enclosing boundaries, due to the spontaneous baseline activity which has not been quenched by the input injection (Figure 17H). In this hole-free case, the ON-C filtered images effectively separate each connected region from all others in the original image. However, because of the Gaussian shapes of the kernels used in these filters, the edges in the ON-C patterns also exhibit a Gaussian spatial spread. Also, smooth gradients, such as those gen-

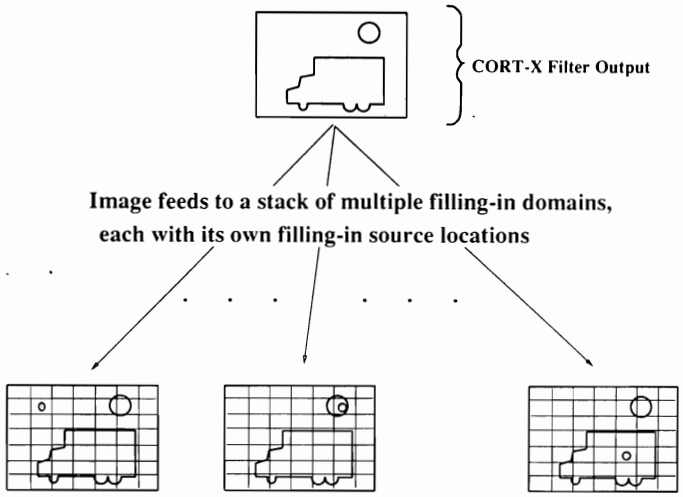

FIGURE 16. Copies of the CORT-X filter output are sent to *M* filling-in networks (three of which are pictured here). Activity is injected into a different place in each (gray disks) and begins to spread.

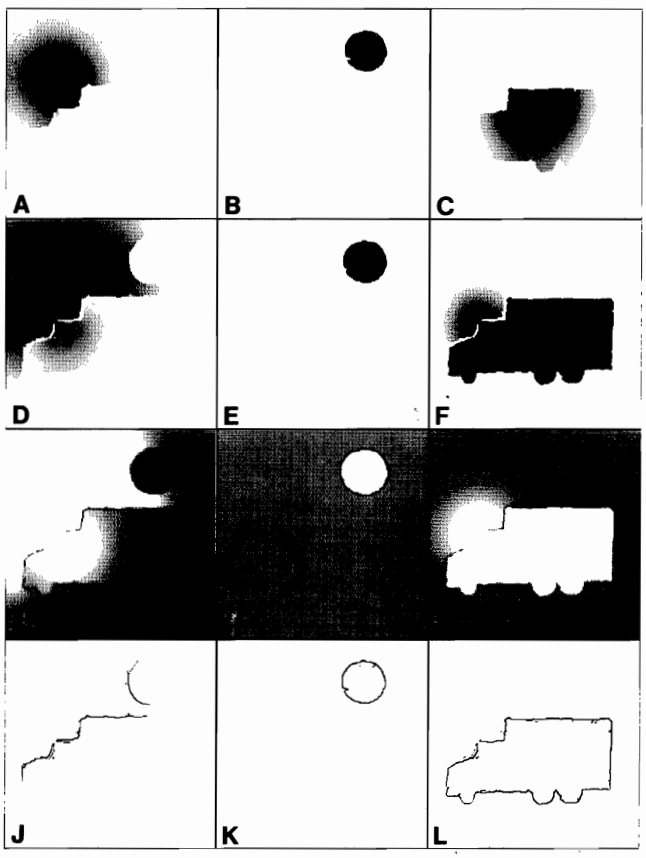

FIGURE 17. Row 1: Filling-in of activity is initiated at three different places in parallel filling-in networks. Due to the quantization of the gray scale, small filled-in activations do not print, even though they are detected by the shunting networks, as noted in rows 2 through 4. Row 2: The ON-C filter output of the respective filled-in regions. Row 3: The OFF-C filter output of the respective filled-in regions. Row 4: The final "separated figure" outputs to be passed along to a pattern recognizer such as an ART network.

erated by activation spreading across regions where there are no boundary signals, may not be quenched to zero if the center and surround kernels are of unequal area. If the pattern recognition stage is based on image boundary signals, then better boundaries must be reconstructed. One way is to subject each image to another pass through the CORT-X filter. Such a step would be useful if the second CORT-X filter used a longer range cooperative kernel to complete boundaries across any remaining gaps. An alternative approach is to multiplicatively gate each ON-C pattern with the original CORT-X output. The resulting M images are then ready to be passed along and contain among them all the separated figural boundaries of the original image (Figure 17J–L). This method can be run in real-time until the output from each copy generates a recognition event at its ART network, or the input image is removed.

If the original image contains so much noise that the CORT-X filter is unable to produce boundaries without weak points or small gaps, then significant activity could leak out of figural components during the filling-in process (Figure 17A). On the other hand, the ART recognition process, operating in real-time, could recognize the figure before the equilibrium state of equal activation on both sides of the boundary is reached.

Leakage of diffusing activity causes no problem if there are no other nearby object boundaries in the original scene. The spread of activation would produce a smooth gradient through the gap, and the shunting operation would not detect any contrast at the point of leakage or outside the object until well after the recognition event occurred. If, however, another object's boundary were near a boundary gap, then leaking activation followed by the ON-C shunting operation could detect the spurious boundary if the injection site were closer to the spurious boundary than to other boundaries of the figure.

In this case, the OFF-C filtered image is helpful. The ON-C signal at boundary regions exterior to the desired object is not as strong as the OFF-C filter signal at these points unless the boundary and the injection site are both proximal to the boundary gap. This property is due to the fact that very little injected activity would have spread there to excite the ON-C field, and the OFF-C field is tonically active, as shown in eqn (16) below. Subtracting the OFF-C image from the ON-C image before CORT-X gating therefore helps to correct this pathological situation. The OFF-C subtraction does not, moreover, distort the desired object boundary when the ON-C signal is stronger than the OFF-C signal. In Figure 18, for example, a hole was made in the boundary of the truck near the boundary of another figure before filling-in occurs. The ON-C network clearly detects the outer boundary of the moon (Figure 18A). The OFF-C filtered signal is, however, stronger at these

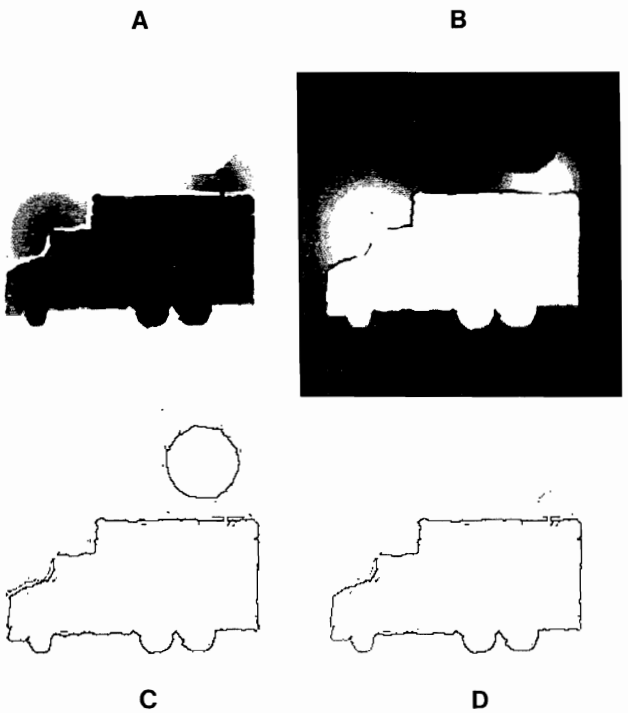

FIGURE 18. (a) The ON-C filter output of the filled in region with a hole in the boundary proximal to another figure; (b) The OFF-C filter of the filled in region; (c) The output as it would look without using the complementary OFF-C image; and (d) The final image which uses the OFF-C filter to control the effects of the leakage leaving only the desired boundary signal.

points (Figure 18B). Without taking the OFF-C network information into account, the output could cause difficulties for a pattern recognizer since the objects would not be separated (Figure 18C). Combining the two shunting network filters produces the desired separation (Figure 18D). The output of such an operation computes a *double opponent* receptive field (Grossberg, 1987).

A single figure is typically large enough to enclose several injection sites across the set of filling-in networks F_m. Thus, even if the injection site is close to a boundary gap and to a nearby spurious boundary in one network F_m, the injection sites of other filling-in networks will be further from the boundary gap. In some of these networks, double opponent processing can compensate for the gap, and trigger a correct recognition of the figure from the corresponding ART network. If the spurious boundaries are strong enough, the corresponding ART network will remain silent, as in Figure 17J, because the combination of partial figure and background is not similar enough to a previously learned recognition code.

In summary, successful parallel separation of multiple connected figures from other figures and background can be performed with this architecture under noisy conditions by exploiting the possibility of efficient parallel recognition by ART networks of boundary segmentations that are generated by sampling multiple filling-in pespectives.

8. FBF NETWORK EQUATIONS

Input Images

In our computer simulations, the images are 256 * 256 arrays with signal values in the interval [0, 1]. The simulations pictured herein represent maximum signal strength by black and minimum signal strength by white. Noise was generated by randomly choosing a percentage of pixels and setting their values to a random number, or gray level, in the interval [0, 1]. For the simulation pictured herein, 50% of the pixels were randomized. The input pattern $\{I_{ij}\}$ is thus represented as gray levels on a set of square pixels $\{P_{ij}\}$. Pixel P_{ij} attains the value I_{ij} at the set of image points $\{(u, v): i \leq u < i + 1, j \leq v < j + 1\}$.

Step 1 (Discount the Illuminant): ON-C Network

Each node v_{ij} is placed, for notational convenience, at the center of the corresponding pixel P_{ij} where it receives input I_{ij}. The activity x_{ij} of the node v_{ij} at lattice position (i, j) obeys the shunting on-center off-surround equation:

$$\frac{d}{dt}x_{ij} = -Ax_{ij} + (B - x_{ij})C_{ij} - (x_{ij} + D)E_{ij}, \quad (1)$$

where C_{ij} is the on-center interaction and E_{ij} is the off-surround interaction. Each C_{ij} and E_{ij} is a discrete convolution of the input pattern $\{I_{ij}\}$ with a Gaussian

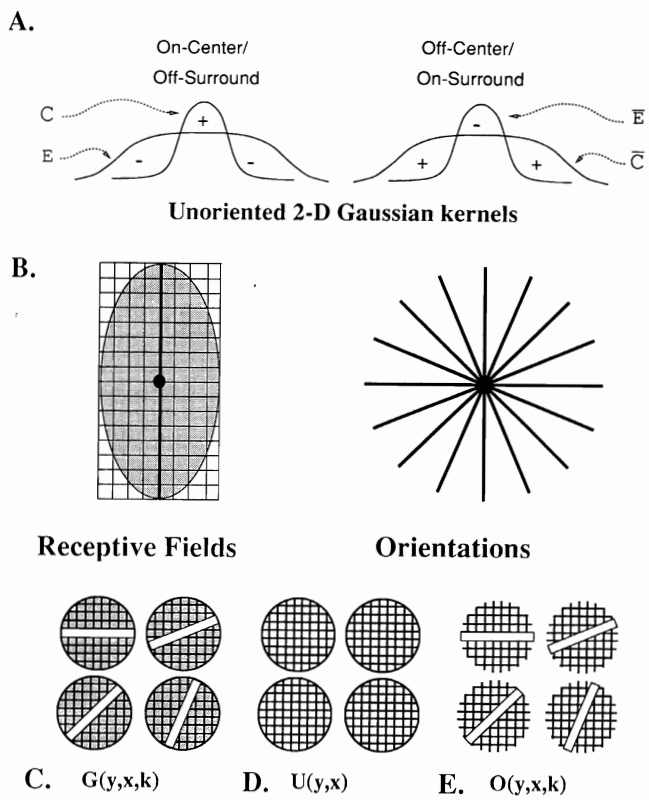

FIGURE 19. The kernels used in the CORT-X Filter. All oriented kernels have orientations at every $\pi/8$ radians.

kernel (Figure 19a). Thus

$$C_{ij} = \sum_{p,q} I_{pq} C_{pqij}, \tag{2}$$

and

$$E_{ij} = \sum_{p,q} I_{pq} E_{pqij}, \tag{3}$$

where

$$C_{pqij} = C \exp\{-\alpha^{-2} \log 2[(p - i)^2 + (q - j)^2]\}, \tag{4}$$

and

$$E_{pqij} = E \exp\{-\beta^{-2} \log 2[(p - i)^2 + (q - j)^2]\}. \tag{5}$$

In our simulations, $A = 134$, $B = 1$, $C = 7$, $D = .5$, $E = 3.333$, $\alpha = 1.3$, and $\beta = 1.875$. For this choice of parameters, the ON-C and OFF-C Gaussians are of equal area. At equilibrium ($dx_{ij}/dt = 0$), (1) yields:

$$x_{ij} = \frac{\Sigma_{(p,q)}(BC_{pqij} - DE_{pqij})I_{pq}}{A + \Sigma_{(p,q)}(C_{pqij} + E_{pqij})I_{pq}}. \tag{6}$$

OFF-C Network

The activity $\bar{x}_{ij}$ of the node v_{ij} at lattice position (i, j) obeys the shunting off-center on-surround equation:

$$\frac{d}{dt}\bar{x}_{ij} = -A(\bar{x}_{ij} - S) + (\bar{B} - \bar{x}_{ij})\bar{C}_{ij} - (\bar{x}_{ij} + \bar{D})\bar{E}_{ij}, \tag{7}$$

where the on-center kernel of $\bar{x}_{ij}$ is the off-surround kernel of x_{ij}, and the off-surround kernel of $\bar{x}_{ij}$ is the on-center kernel of x_{ij}. In particular,

$$\bar{B} = D, \tag{8}$$

$$\bar{C}_{ij} = E_{ij}, \tag{9}$$

$$\bar{D} = B, \tag{10}$$

and

$$\bar{E}_{ij} = C_{ij}. \tag{11}$$

By eqs (8)–(11), (7) may be written as

$$\frac{d}{dt}\bar{x}_{ij} = -A(\bar{x}_{ij} - S) + (D - \bar{x}_{ij})E_{ij} - (\bar{x}_{ij} + B)C_{ij}. \tag{12}$$

At equilibrium,

$$\bar{x}_{ij} = \frac{AS + \Sigma_{(p,q)}(DE_{pqij} - BC_{qpij})I_{pq}}{A + \Sigma_{(p,q)}(C_{pqij} + E_{pqij})I_{pq}}. \tag{13}$$

It follows by summing (6) and (13) that

$$x_{ij} + \bar{x}_{ij} = \frac{AS}{A + \Sigma_{(p,q)}(C_{pqij} + E_{pqij})I_{pq}}, \tag{14}$$

which shows that for images $\{I_{pq}\}$ with constant Gaussianly filtered total activity

$$\sum_{p,q} (C_{pqij} + E_{pqij})I_{pq}, \tag{15}$$

the sum $x_{ij} + \bar{x}_{ij}$ is conserved and maintained at a positive value that increases with the tonic activity level S of $\bar{x}_{ij}$ in the dark. In our simulations, $S = .2$. The value for the parameter A (given the other parameter values) was chosen so that the activation of cell $\bar{x}_{ij}$ takes on values between S and 0 under spatially uniform illumination between 0 and 1 within its receptive field. Under such spatially uniform illumination conditions, $I_{pq} = I$ for all (p, q). Since for our choice of parameters the two Gaussian kernels are of equal area, let $\Phi = \Sigma_{(p,q)} E_{pqij} = \Sigma_{(p,q)} C_{pqij}$. Then Φ factors out so that the equilibrium eqn (13) becomes

$$\bar{x}_{ij} = \frac{AS + I(D - B)\Phi}{A + 2I\Phi}. \tag{16}$$

We derive the value for A by

$$A = \frac{(B - D)\Phi}{S}, \tag{17}$$

so that when $I = 1$, the numerator vanishes. When $I = 0$, $\bar{x}_{ij} = S$ independent of A.

Step 2 (CORT-X Filter)

All simple cell input fields are elliptical. Two sizes of input fields were used, indexed by the subscript s. The smaller scale, $s = 1$, had a major axis of 12 pixels and a minor axis of 6 pixels. The larger scale, $s = 2$, had a major axis of 20 pixels and a minor axis of 10 pixels. Orientations were chosen around the clock spaced by $\pi/8$ degrees. They are indexed below by the subscript k.

Simple Cells

A simple cell with index (i, j) is centered at the lower left hand corner of pixel P_{ij}. By this convention, the nodes v_{pq} of the shunting variables x_{pq} do not lie on the oriented axes that separate the excitatory and inhibitory halves of vertically and horizontally oriented receptive fields. The output of the pair of simple cells of scale s centered at index (i, j) with receptive field orientation k is defined by

$$S_{sL}(i, j, k) = \max[L_s(i, j, k) - \alpha_s R_s(i, j, k) - \beta_s, 0], \tag{18}$$

and

$$S_{sR}(i, j, k) = \max[R_s(i, j, k) - \alpha_s L_s(i, j, k) - \beta_s, 0], \tag{19}$$

where $L_s(i, j, k)$ and $R_s(i, j, k)$ are the contributions of the left-half $l_s(i, j, k)$ and right-half $r_s(i, j, k)$,

respectively, of the oriented input field; that is,

$$L_s(i, j, k) = \frac{\sum_{(p,q)\in l_s(i,j,k)} x_{pq} w_{pq}}{\sum_{(p,q)\in l_b(i,j,k)} w_{pq}} \quad (20)$$

and

$$R_s(i, j, k) = \frac{\sum_{(p,q)\in r_s(i,j,k)} x_{pq} w_{pq}}{\sum_{(p,q)\in r_s(i,j,k)} w_{pq}}, \quad (21)$$

where w_{pq} is a weighting factor equal to the proportion of the area of pixel P_{pq} (taken to be one square unit) covered by the receptive field. An activity x_{pq} was included in L_s or R_s if its pixel had a nonzero intersection with the corresponding half of the receptive field. Parameters α_s are threshold contrast parameters and parameters β_s are threshold noise parameters. We chose $\alpha_1 = 1.4$, $\alpha_2 = 2.0$, and $\beta_1 = \beta_2 = \beta = .012$. Each simple cell in eqns (18) and (19) is sensitive to the opposite direction-of-contrast than its companion, as indicated by the indices L and R in S_{sL} and S_{sR}, respectively.

The ON-C and OFF-C networks each input to a different network of simple cells. The simple cells that receive signals from the ON-C network are denoted by S^+_{sL} and S^+_{sR}. The simple cells that receive signals from the OFF-C network are denoted by S^-_{sL} and S^-_{sR}.

The complex cells pool inputs from all simple cells of like orientation and scale that are centered at the same location, as described below.

Complex Cells

The complex cell output $C_s(i, j, k)$ is defined by

$$C_s(i, j, k) = F[S^+_{sL}(i, j, k) + S^+_{sR}(i, j, k) + S^-_{sL}(i, j, k) + S^-_{sR}(i, j, k)]. \quad (22)$$

Such a cell is sensitive to spatial scale s and amount-of-contrast centered at cell x with orientation k, but it is insensitive to direction-of-contrast. In our simulations, $F = .5$.

Hypercomplex Cells (First Competitive Stage)

The hypercomplex cells $D_s(i, j, k)$ at the first competitive stage receive input from the spatial competition among the complex cells; that is,

$$D_s(i, j, k) = \max\left[\frac{C_s(i, j, k)}{\varepsilon + \mu \sum_m \sum_{p,q} C_s(p, q, m) G_s(p, q, i, j, k)} - \tau, 0\right], \quad (23)$$

where $\varepsilon = .1$, $\mu = 5$, $\tau = .01$. The oriented competition kernel $G_s(y, x, k)$ is normalized so that

$$\sum_{p,q} G_s(p, q, i, j, k) = 1. \quad (24)$$

As in Figure 19c, they are circular. Complex cells at the kernel periphery are weighted by the proportion of their area (taken to be one square unit) that are covered by the kernel. The grey areas in Figure 19c are inhibitory. Any cells whose defining pixel location lies within the one-unit-wide band through the middle of the kernel do not contribute to the inhibition. In our simulations, the small scale kernel is 8 units in diameter, and the large scale is 16 units in diameter.

Hypercomplex Cells (Second Competitive Stage)

The hypercomplex cells $D_s(i, j)$ at the second competitive stage realize a competition among the oriented activities $D_s(i, j, k)$ at each position x. For simplicity, the process is modelled as a winner-take-all competition; namely,

$$D_s(i, j) = D_2(i, j, K) = \max_k D_s(i, j, k), \quad (25)$$

where index K denotes the orientation of the maximally activated cell.

Multiple-Scale Interaction: Boundary Localization and Noise Suppression

The interaction between scales is defined by the equation

$$B_{12}(i, j) = D_1(i, j) \sum_{p,q} D_2(p, q) U(p, q, i, j). \quad (26)$$

The unoriented excitatory kernel $U(p, q, i, j)$ is circular (Figure 19d), and normalized so that

$$\sum_{p,q} U(p, q, i, j) = 1, \quad (27)$$

and had a diameter of 8 units. All cells covered by the kernel contribute to the excitation to the extent that their area (taken to be one square unit) is covered by the kernel. The small-scale hypercomplex cell $D_1(i, j)$ accurately localizes boundary segments and suppresses noise near the boundary. The large-scale hypercomplex cell $D_2(p, q)$ suppresses noise far from the boundary. The product $D_1(i, j)D_2(p, q)$ would simultaneously realize both constraints except that, due to the poor spatial localization of D_2, this term may be zero at boundary points of high curvature, thereby cancelling the good localization properties of $D_1(i, j)$. The effect of $D_2(p, q)$ in the equation is made more spatially diffuse by the kernel $U(i, j, p, q)$. The size of $U(i, j, p, q)$ is chosen to scale with that of $D_2(p, q)$ in order to compensate for the positional uncertanty of $D_2(p, q)$; a larger choice of $D_2(p, q)$ would necessitate a larger choice of $U(i, j, p, q)$. Although term $\Sigma D_2(p, q)U(p, q, i, j)$ in eqn (26) localizes the boundary even less accurately than $D_2(p, q)$ does, the *product* of $D_1(i, j)$ with $\Sigma_{p,q} D_2(p, q)U(p, q, i, j)$ in eqn (26)

restores this loss of boundary localization. Moreover, kernel $U(p, q, i, j)$ causes no harm at locations p, q that are far from the boundary, since $D_2(p, q) = 0$ there (Carpenter, Grossberg & Mehanian, 1989).

Long-Range Cooperation: Boundary Completion

The function $B_{12}(i, j)$ represents the image boundary well except where boundary pixels are missing due to noise. More and larger boundary gaps are generated as the noise level increases.

The large detectors $D_2(i, j)$ can be used to partially overcome this problem. Because of the spatial uncertainty of the large detectors $D_2(i, j)$, they are capable of responding at locations where pixel signal strength has been reduced by noise. Such boundary signals may, however, be poorly localized. To overcome this trade-off between boundary completion and localization, cooperative interactions among the large-scale cells are defined by

$$B_2(i, j) = D_2(i, j) \times \max\left[\sum_{p,q} D_2(p, q, K)O(p, q, i, j, K) - \delta, 0\right]. \quad (28)$$

The oriented kernel $O(p, q, i, j, k)$ is defined by the one-unit-wide white strips in Figure 19E. Any cells with centers that lie within the one unit wide band contributes to the cooperative process. The kernel is normalized so that

$$\sum_{(p,q)\text{ in kernel}} O(p, q, i, j, k) = 1. \quad (29)$$

In our simulations, the length of the kernel was 12 units, and $\delta = .001$.

CORT-X Output

The final output of the CORT-X filter is the rectified sum of the multiple scale interaction and the cooperative interaction:

$$B(i, j) = 1[B_{12}(i, j) + B_2(i, j)], \quad (30)$$

where

$$1[w] = \begin{cases} 1 & \text{if } w > 0 \\ 0 & \text{otherwise} \end{cases} \quad (31)$$

is the Heaviside function.

Step 3: Filling-In

In each filling-in network F_m, an input is injected into a different area, with the shape of either a narrow Gaussian or a single node. For example, let $I = I(m)$ and $J = J(m)$ define the injection indices (I, J) of F_m. Then the injected input pattern to F_m was chosen to be

$$X_{ij}^{(m)} = X \exp\{-\gamma^{-2} \log 2[(I - i)^2 + (J - j)^2]\}. \quad (32)$$

This input pattern triggers filling in within F_m via the nonlinear diffusion equation (Grossberg & Todorović, 1988):

$$\frac{d}{dt} S_{ij}^{(m)} = -MS_{ij}^{(m)} + \sum_{p,q \in N_{ij}} (S_{pq}^{(m)} - S_{ij}^{(m)})P_{pqij} + X_{ij}^{(m)}, \quad (33)$$

where $S_{ij}^{(m)}$ is the activity of the (i, j) node of F_m, and the index set N_{ij} of the summation contains the nearest neighbors of (i, j) The permeability coefficient P_{pqij} is defined by

$$P_{pqij} = \frac{\delta}{1 + \varepsilon(B(p, q) + B(i, j))}, \quad (34)$$

where $B(p, q)$ and $B(i, j)$ are the outputs (30) from the CORT-X filter at the positions (p, q) and (i, j) respectively. Thus, activity spreads poorly, if at all, between cells where boundary signals are large, and easily where boundary signals do not exist. In the simulations, the parameters $M = .0001$, $X = 50$, $\gamma = .5$, $\varepsilon = 100000$, $\delta = 10$.

Step 4: Figure-Ground Separation

Each filling-in network F_m inputs its filled-in image to shunting ON-C and OFF-C networks using the same eqs (1) and (12) as above, with parameters $A = 1$, $B_{ij} = 1$, $C = 18$, $D_{ij} = .5$, $E = 3.333$, $\alpha = 2.96$, $\beta = 7$, and $S = .2$. Here the on-activations $x_{ij}^{(m)}$ and off-activations $\bar{x}_{ij}^{(m)}$ are parameterized by the filling-in network F_m from which they are derived. The boundary representation at position (i, j) of a figure derived from network F_m is defined by

$$R_{ij}^{(m)} = 1[(x_{ij}^{(m)} - \bar{x}_{ij}^{(m)})B(i, j)], \quad (35)$$

where $1[w]$ is the Heaviside function. In other words, the figural boundary that is separated by network F_m is computed from the double opponent filter $(x_{ij}^{(m)} - \bar{x}_{ij}^{(m)})$ of filled-in F_m activation, gated by the CORT-X boundary segmentation $B(i, j)$.

9. CONCLUDING REMARKS: RECOGNITION OF CONJUNCTIVE FEATURES

The FBF network separates figure-from-ground by using regular arrays of feedforward networks to discount the illuminant and to generate boundary segmentations, nearest-neighbor feedback signals for filling-in, and a proliferation of these circuits in parallel copies that input to parallel arrays of ART pattern recognition networks. The FBF networks thus seem to be appropriate designs for implementation in parallel hardware capable of operating at high rates in real time.

It remains to discuss why the biologically motivated FBF network can automatically distinguish between the pair of connected and disconnected Min-

sky–Papert figures in Figure 1 that humans cannot distinguish. The main reason is that FBF networks use internally generated "dye injections." These inputs are topographically distributed across the entire perceptual space such that each dye injection is delivered to a different filling-in domain, or slab. In human perception, by contrast, the same Feature Contour signals that initiate Boundary Contour formation also act as input sources that trigger featural filling-in (Grossberg, 1987; Grossberg & Mingolla, 1985a; Grossberg & Todorović, 1988). The regions used by Minsky and Papert (1969) were all of the same color, of similar overall shape, and occupied essentially the same region of their respective images. They would therefore tend to fill-in the same slab, or set of slabs, when they are being perceived by humans. Hence, they could not be distinguished by filling-in different slabs as a function of their connectivity.

This observation clarifies a recent controversy about human perception; namely whether target figures that differ from distractor figures by more than one type of feature can be separated from them by rapid parallel processing that does not require serial search. Treisman and her colleagues (Treisman & Gelade, 1980; Treisman & Souther, 1985) have suggested that such parallel processing can occur only if the target is distinguished from distractors along a single stimulus dimension, whereas if a target is defined by the conjunction of two or more stimulus dimensions, then it can only be separated from the distractors by a serial search process.

An exception to this rule was discovered by Nakayama and Silverman (1986), who showed that targets which differ from distractors by a combination of disparity and color, or of disparity and motion, can be rapidly separated without serial search. This result is explained by the theory of 3-D vision, called FACADE Theory (see Section 4), which motivated the design of FBF networks. In Grossberg (1987), FACADE representations of different disparity-and-color combinations activate different combinations of filling-in domains. They are *structurally* separated in the representation, and hence can be rapidly detected.

In summary, the difficulty of distinguishing the connected and disconnected Minsky–Papert displays can now be explained by the same mechanisms that explain rapid search of Nakayama–Silverman displays, and that provide the heuristics for designing an FBF network for automatic figure-ground separation.

REFERENCES

Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **8**, 679–698.

Carpenter, G. A., & Grossberg, S. (1987a). Invariant pattern recognition and recall by an attentive self-organizing ART architecture in a nonstationary world. In M. Caudill & C. Butler (Eds.), *Proceedings of the IEEE international conference on neural networks, II* (pp. 737–746). San Diego, CA: SOS Printing.

Carpenter, G. A., & Grossberg, S. (1987b). ART 2: Stable self-organization of pattern recognition codes for analog input patterns. *Applied Optics*, **26**, 4919–4930.

Carpenter, G. A., & Grossberg, S. (1988). The ART of adaptive pattern recognition by a self-organizing neural network. *Computer*, **21**, 77–88.

Carpenter, G. A., Grossberg, S., & Mehanian, C. (1989). Invariant recognition of cluttered scenes by a self-organizing ART architecture: CORT-X boundary segmentation. *Neural Networks*, **2**, 169–181.

Casasent, D., & Psaltis, D. (1976). Position, rotation, and scale invariant optical correlations. *Applied Optics*, **15**, 1793–1799.

Cavanagh, P. (1978). Size and position invariance in the visual system. *Perception*, **7**, 167–177.

Cavanagh, P. (1984). Image transforms in the visual system. In P. C. Dodwell & T. Caelli (Eds.), *Figural synthesis*. Hillsdale, NJ: Erlbaum Associates.

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

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

Grossberg, S. (1988). *Neural networks and natural intelligence*. Cambridge, MA: MIT Press.

Grossberg, S. (1991). 3-D vision and figure-ground separation by visual cortex. Submitted for publication.

Grossberg, S., & Marshall, J. (1989). Stereo boundary fusion by cortical complex cells: A system of maps, filters, and feedback networks for multiplexing distributed data. *Neural Networks*, **2**, 29–51.

Grossberg, S., & Mingolla, E. (1985a). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review*, **92**, 173–211.

Grossberg, S., & Mingolla, E. (1985b). Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations. *Perception and Psychophysics*, **38**, 141–171.

Grossberg, S., Mingolla, E., & Todorović, D. (1989). A neural network architecture for preattentive vision. *IEEE Transactions on Biomedical Engineering*, **36**, 65–84.

Grossberg, S., & Todorović, D. (1988). Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena. *Perception and Psychophysics*, **43**, 241–277.

Grossberg, S., & Wyse, L. (1990). Automatic figure-ground separation by an iterated boundary segmentation and filling-in network. In *Proceedings of the conference on Neural Networks for Automatic Target Recognition* (p. 27). Boston, MA: Wang Institute of Boston University. May 11–13, 1990.

Gschwendtner, A. B., Harney, R. C., & Hull, R. J. (1983). Coherent IR radar technology. In D. K. Killinger & A. Mooradian (Eds.), *Optical and laser remote sensing*. New York: Springer-Verlag.

Harney, R. C. (1980). Infrared airborne radar. *Proceedings of the IEEE 1980 electronic and aerospace systems conference* (pp. 462–471).

Harney, R. C., & Hull, R. J. (1980). Compact infrared radar technology. In CO_2 laser devices and applications. *Proceedings of the Society of Photo-Optical Instrumentation Engineers*, **227**, 162–170.

Harney, R. C. (1981). Military applications of coherent infrared radar. *Physics and Technology of Coherent Infrared Radar* (SPIE Proceedings), **300**, 2–11.

Hull, R. J., & Marcus, S. (1980). A tactical 10.6 micrometer imaging radar. *Proceedings of the IEEE, National Aerospace and Electronics Conference*, **2**, 662–668.

Kolodzy, P. (1987). Multidimensional machine vision using neural networks. In M. Caudill & C. Butler (Eds.), *Proceedings of the IEEE international conference on neural networks*, **II**, 747–758.

Land, E. H. (1977). The retinex theory of color vision. *Scientific American*, **237**, 108–128.

Minsky, M. L., & Papert, S. A. (1969). *Perceptrons*. Cambridge, MA: MIT Press.

Minsky, M. L., & Papert, S. A. (1988). *Perceptrons: Expanded edition*. Cambridge, MA: MIT Press.

Nakayama, K., & Silverman, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature*, **320**, 264–265.

Sullivan, L. J. (1980). Infrared coherent radar. In CO_2 laser devices and applications. *Proceedings of the Society of Photo-Optical Instrumentation Engineers*, **227**, 148–161.

Sullivan, L. J. (1981). Firepond laser radar. *Electro/81 Conference Record*, Session 34.

Sullivan, D. R., Harney, R. C., & Martin, J. B. (1979). Real-time quasi-3-dimensional display of infrared radar images: Real-time signal processing II. *Proceedings of the Society of Photo-Optical Instrumentation Engineers*, **180**, 57–65.

Szu, H. (1986). Holographic coordinate transformations and optical computing. In H. Szu (Ed.), *Optical and hybrid computing*. Bellingham, Washington: International Society for Optical Engineering, Society of Photo-Optical Instrumentation Engineers Vol. 634.

Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, **12**, 97–136.

Treisman, A., & Souther, J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General*, **114**, 285–310.

Walters, D. K. W. (1987). Rho-space: A neural network for the detection and representation of oriented edges. *Proceedings of the ninth annual conference of the Cognitive Science Society* (pp. 455–460). Hillsdale, NJ: Erlbaum Press.

APPENDIX

Image Rendering

Each image has been scaled so that the maximum signals strength is mapped to black, and the minimum signal strength is mapped to white. Intermediate values map linearly onto a grey scale. The maximum signal strengths for the images are: Figure 8: (a) 1.0, (b) .125, (c) .197. Figure 10: (a) .041, (b) .067, (c) .018, (d) .058. Figure 11: (a) .054, (b) .035. Figure 13: (a) .261, (b) .219. Figure 14: (a) .013, (b) .038. Figure 15: (a) 1.0. Figure 17: Row (col. 1) 143.2, (col. 2) 753.3, (col. 3) 195.3; Row 2 (col 1) .324, (col. 2) .330, (col. 3) .325; Row 3 (col. 1) .287, (col. 2) .396, (col. 3) .333; Row 4 (col. 1) 1.0, (col. 2) 1.0, (col. 3) 1.0. Figure 18: (a) .325 (b) .332 (c) 1.0 (d) 1.0.