

Why do parallel cortical systems exist for the perception of static form and moving form?

STEPHEN GROSSBERG

Center for Adaptive Systems, Boston University, Boston, Massachusetts

This article analyzes computational properties that clarify why the parallel cortical systems $V1 \rightarrow V2$, $V1 \rightarrow MT$, and $V1 \rightarrow V2 \rightarrow MT$ exist for the perceptual processing of static visual forms and moving visual forms. The article describes a symmetry principle, called FM symmetry, that is predicted to govern the development of these parallel cortical systems by computing all possible ways of symmetrically gating sustained cells with transient cells and organizing these sustained-transient cells into opponent pairs of on-cells and off-cells whose output signals are insensitive to direction of contrast. This symmetric organization explains how the static form system (static BCS) generates emergent boundary segmentations whose outputs are insensitive to direction of contrast and insensitive to direction of motion, whereas the motion form system (motion BCS) generates emergent boundary segmentations whose outputs are insensitive to direction of contrast but sensitive to direction of motion. FM symmetry clarifies why the geometries of static and motion form perception differ—for example, why the opposite orientation of vertical is horizontal (90°), but the opposite direction of up is down (180°). Opposite orientations and directions are embedded in gated dipole opponent processes that are capable of antagonistic rebound. Negative afterimages, such as the MacKay and waterfall illusions, are hereby explained as are aftereffects of long-range apparent motion. These antagonistic rebounds help to control a dynamic balance between complementary perceptual states of resonance and reset. Resonance cooperatively links features into emergent boundary segmentations via positive feedback in a CC loop, and reset terminates a resonance when the image changes, thereby preventing massive smearing of percepts. These complementary preattentive states of resonance and reset are related to analogous states that govern attentive feature integration, learning, and memory search in adaptive resonance theory. The mechanism used in the $V1 \rightarrow MT$ system to generate a wave of apparent motion between discrete flashes may also be used in other cortical systems to generate spatial shifts of attention. The theory suggests how the $V1 \rightarrow V2 \rightarrow MT$ cortical stream helps to compute moving form in depth and how long-range apparent motion of illusory contours occurs. These results collectively argue against vision theories that espouse independent processing modules. Instead, specialized subsystems interact to overcome computational uncertainties and complementary deficiencies, to cooperatively bind features into context-sensitive resonances, and to realize symmetry principles that are predicted to govern the development of the visual cortex.

1. The Motion Boundary Contour System

This article contributes further evidence for a new theory of biological motion perception that was outlined in Grossberg (1987c) and quantitatively specified and analyzed in Grossberg and Rudd (1989a, 1989c, 1990a, 1990b) and in Grossberg and Mingolla (1990a, 1990b, 1990c). The new theory consists of a neural architecture called a *motion boundary contour system*, or *motion BCS*. The motion BCS consists of several parallel copies, such that each copy is activated by a different range of receptive-field sizes. Each copy is further subdivided into hierar-

chically organized subsystems: a *motion-oriented contrast filter*, or MOC filter, for preprocessing moving images, and a *cooperative-competitive feedback loop*, or CC loop, for generating coherent emergent boundary segmentations of the filtered signals.

These results have provided a computational explanation for the cortical stream $V1 \rightarrow MT$ that joins the areas $V1$ and MT of visual cortex. An earlier model of static form perception, summarized below, modeled aspects of the parallel $V1 \rightarrow V2$ cortical stream. Evidence for the MOC filter includes its ability to explain a variety of classical and recent data about short-range and long-range apparent motion, motion capture, induced motion, and cortical cell properties that have not yet been explained by alternative models. Grossberg and Rudd (1989b) have, moreover, shown how the main properties of other motion perception models can be assimilated into different parts of the motion BCS design.

Work supported in part by the Air Force Office of Scientific Research (AFOSR 90-0175), the Army Research Office (ARO DAAL-03-88-K-0088), DARPA (AFOSR 90-0083), and Hughes Research Labs (S1-903136). The author's mailing address is Center for Adaptive Systems, Boston University, 111 Cummington Street, Boston, MA 02215.

2. Why is a Static Form Perception System Not Sufficient?

The motion BCS model provides a computationally precise answer to the following perplexing question. It is well known that some regions of the visual cortex are specialized for motion processing, notably region MT (Albright, Desimone, & Gross, 1984; Maunsell & van Essen, 1983; Newsome, Gizzi, & Movshon, 1983; Zeki, 1974a, 1974b). On the other hand, even the earliest stages of visual cortex processing, such as the simple cells in area V1, are sensitive to changes in stimulus intensity and to direction of motion (DeValois, Albrecht, & Thorell, 1982; Heggelund, 1981; Hubel & Wiesel, 1962, 1968, 1977; Tanaka, Lee, & Creutzfeldt, 1983). Why has evolution gone to the trouble to generate such specialized regions as MT when even the simple cells of V1 are already change-sensitive and direction-sensitive? What computational properties are achieved by MT that are not already available in V1 and its prestriate projections V2 and V4? In response to this question, many scientists suggest that a motion system needs larger receptive fields. This may be true, but it cannot be the heart of the answer, because V1, V2, and V4 also possess multiple receptive-field sizes.

Our answer evolved, along with a new theory of motion form perception, after some unexpected implications of our previous theory of static form perception were noticed. The latter theory has been called FACADE theory, because its visual representations are predicted to multiplex together properties of Form-And-Color-And-Depth in prestriate cortical area V4. FACADE theory describes the neural architecture of two subsystems, the boundary contour system (BCS) and the feature contour system (FCS), whose properties are computationally complementary (Grossberg, Mingolla, & Todorović, 1989). The BCS generates an emergent three-dimensional (3-D) boundary segmentation of edges, texture, shading, and stereo information at multiple spatial scales (Grossberg, 1987b, 1987c, 1990; Grossberg & Marshall, 1989; Grossberg & Mingolla, 1985a, 1985b, 1987). The FCS compensates for variable illumination conditions and fills in surface properties of brightness, color, and depth among multiple spatial scales (Cohen & Grossberg, 1984; Grossberg, 1987b, 1987c; Grossberg & Mingolla, 1985a; Grossberg & Todorović, 1988).

The BCS provided a new computational rationale as well as a model of the neural circuits governing classical cortical cell types, including simple cells, complex cells, and hypercomplex cells. The theory also predicted a new cell type, the bipole cell (Cohen & Grossberg, 1984; Grossberg & Mingolla, 1985a), whose properties have been confirmed by neurophysiological experiments (Peterhans & von der Heydt, 1989; von der Heydt, Peterhans, & Baumgartner, 1984).

This BCS model, now called the *static* BCS model, consists of several parallel copies, with each copy being activated by a different range of receptive-field sizes, as in the motion BCS. Also as in the motion BCS, each static BCS copy is further subdivided into two hierarchically

BOUNDARY CONTOUR SYSTEM (BCS)

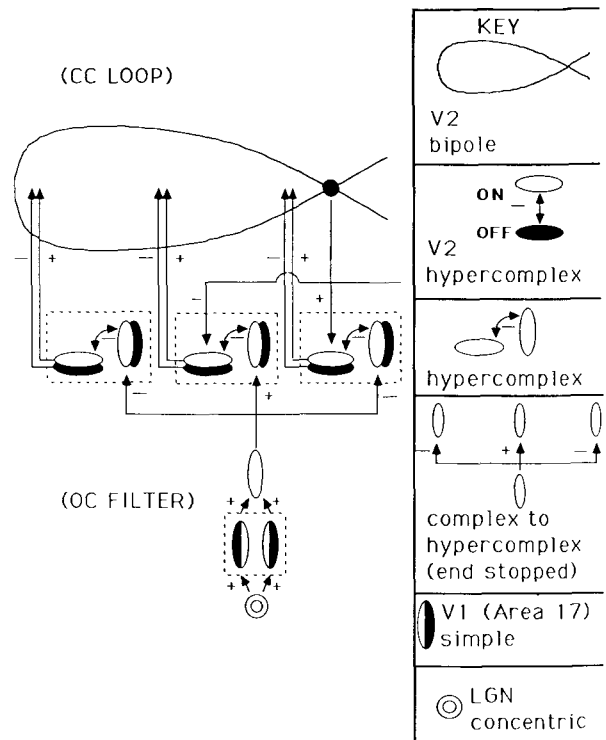


Figure 1. The static boundary contour system circuit described by Grossberg and Mingolla (1985b). The circuit is divided into an oriented contrast-sensitive filter (SOC filter) followed by a cooperative-competitive feedback network (CC loop). Multiple copies of this circuit are used, each corresponding to a different range of receptive-field sizes of the SOC filter. The depicted circuit has been used to analyze data about monocular vision. A binocular generalization of the circuit has also been described (Grossberg, 1987c; Grossberg & Marshall, 1989).

organized systems (Figure 1): a *static oriented contrast* filter, or SOC filter, for preprocessing quasi-static images (the eye never ceases to jiggle in its orbit), and a cooperative-competitive feedback (CC) loop, for generating coherent emergent boundary segmentations of the filtered signals. Thus the motion BCS and static BCS models share many common design features. This important fact, which is not evident in other form and motion theories, enables us to view both models as variations on a common architectural design for visual cortex. A great conceptual simplification is afforded by the fact that variations on a common design can now be used to explain large data bases about form and motion perception that have heretofore been treated separately.

3. Joining Sensitivity to Direction of Motion with Insensitivity to Direction of Contrast

Analysis of the SOC filter design revealed that one of its basic properties made it unsuitable for motion processing: the output of the SOC filter cannot effectively process the direction of motion of a moving figure. This deficiency

arises from the way in which the SOC filter becomes insensitive to direction of contrast at its complex cell level. Insensitivity to the direction of contrast of the SOC filter's complex cells enables the CC loop of the static BCS, which involves feedback interactions between hyper-complex cells and bipole cells (Figure 1), to generate boundary segmentations along scenic contrast reversals (Figures 2 and 3).

The simple cells at the first BCS level are, however, sensitive to direction of contrast (Figure 4). The activities of like-oriented simple cells that are sensitive to opposite directions of contrast are rectified before they generate

outputs to their target complex cells. Because the complex cells pool outputs from both directions of contrast, they are themselves insensitive to direction of contrast.

Figure 4 shows a single pair of simple cells generating inputs to each complex cell. Such an arrangement is not sufficient in general. For example, Grossberg (1987c) and Grossberg and Marshall (1989) have shown that multiple simple cells may input to each complex cell. The number of converging simple cells is predicted to covary in a self-similar manner with the size of the simple-cell receptive fields, and then to trigger nonlinear contrast-enhancing competition at the complex cell level, in order to explain

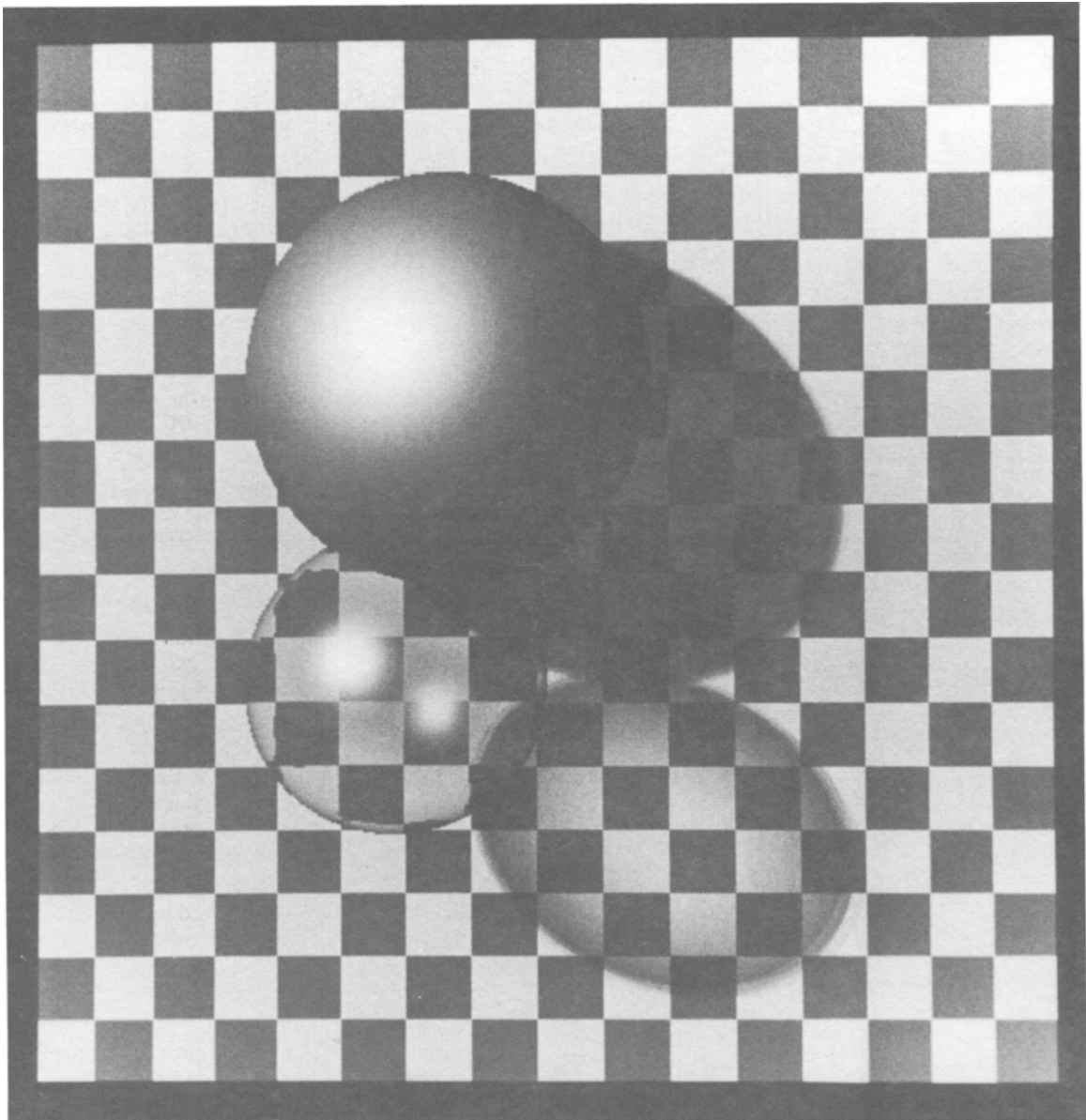


Figure 2. Long vertical and horizontal boundaries are detected despite regular contrast reversals in defining the grid of alternating black and white squares. (From "Perception of Surface Curvature and Direction of Illumination from Patterns of Shading" by E. Mingolla and J. T. Todd, 1983, *Journal of Experimental Psychology: Human Perception and Performance*, 9, p. 586. Copyright 1983 by American Psychological Association Inc. Printed with permission.)

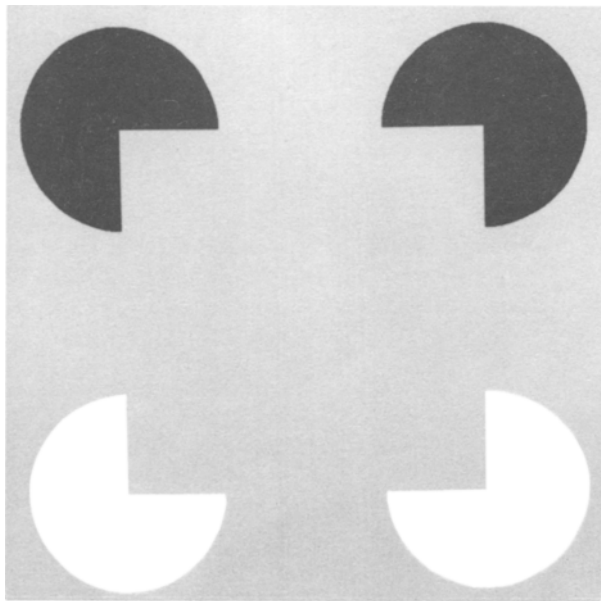


Figure 3. A reverse-contrast Kanizsa square: The BCS is capable of completing illusory boundaries between the vertical dark-light and light-dark contrasts of the Pac Man figures. This boundary completion, or emergent segmentation, process enables the BCS to detect boundaries along contrast reversals, as in Figure 2.

such basic data about binocular vision as the size-disparity correlation and binocular fusion and rivalry.

Inspection of the simple-cell-to-complex cell interaction in Figure 4 shows that a vertically oriented complex cell could respond to, say, a dark-light vertical edge moving to the right *and* to a light-dark vertical edge moving to the left. Thus, the process whereby complex cells become insensitive to direction of contrast has rendered them insensitive to direction of motion in the SOC filter.

The main design problem leading to a MOC filter is to make the minimal changes in the SOC filter that are needed to model an oriented, contrast-sensitive filter whose outputs are insensitive to direction of contrast—a property that is just as important for moving images as it is for static images—yet sensitive to direction of motion—a property that is certainly essential in a motion perception system. The MOC filter, summarized in Figure 5 and Table 1, is rigorously defined in Section 14. It introduces an extra degree of computational freedom which, in one stroke, achieves several important properties: sensitivity to direction of motion, long-range interactions, and binocularity. In particular, the simple cells at the input end of the MOC filter are sensitive to direction of contrast and to stimulus orientation. They are also monocular (except for ocular dominance column overlap)

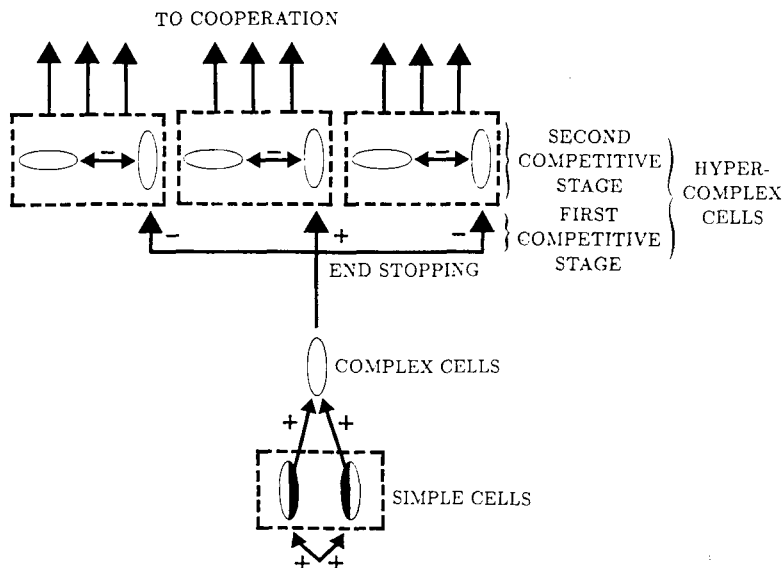


Figure 4. Early stages of SOC filter processing: At each position exist cells with elongated receptive fields (simple cells) of various sizes, which are sensitive to orientation, amount of contrast, and direction of contrast. Pairs of such cells sensitive to like orientation but opposite directions of contrast (lower dashed box) input to cells (complex cells) that are sensitive to orientation and amount of contrast but not to direction of contrast (white ellipses). These cells, in turn, excite like-oriented cells (hypercomplex cells) corresponding to the same position and inhibit like-oriented cells corresponding to nearby positions at the first competitive stage. At the second competitive stage, cells corresponding to the same position but different orientations (higher order hypercomplex cells) inhibit each other via a push-pull competitive interaction.

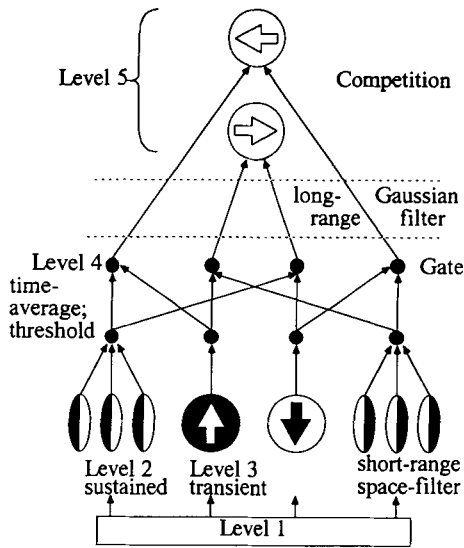


Figure 5. The MOC filter: Level 1 registers the input pattern. Level 2 consists of sustained-response cells with oriented receptive fields that are sensitive to direction of contrast. Level 3 consists of transient response cells with unoriented receptive fields that are sensitive to direction of change in the total cell input. Level 4 cells combine sustained-cell and transient-cell signals to become sensitive to direction of motion and sensitive to direction of contrast. Level 5 cells combine Level 4 cells via a long-range filter to become sensitive to direction of motion and insensitive to direction of contrast.

and interact via spatially short-range interactions. The cells at the output end of the MOC filter play a role analogous to that of SOC filter complex cells. The extra degree of freedom renders these cells insensitive to direction of contrast but sensitive to direction of motion. In particular, these “motion complex cells” can respond to stimulus orientations that are not perpendicular to their preferred direction of motion (Grossberg & Mingolla, 1990a, 1990c). Such a difference between sensitivity to static orientation and sensitivity to motion direction is also found between cortical cells in V1 and MT, respectively (Albright, 1984; Albright et al., 1984; Maunsell & van Essen, 1983; Newsome et al., 1983). In addition, the MOC filter output cells are binocular and have large receptive fields that permit them to engage in the long-range spatial interactions that subservise many apparent motion percepts (Grossberg & Rudd, 1989b, 1990a), as illustrated in Section 15.

4. Why Is a Motion Form Perception System Not Sufficient?

A further analysis of the static BCS and motion BCS poses a new puzzle. This puzzle arises because it seems that the motion BCS has stronger computational properties than does the static BCS. Why, then, has nature not opted to use only a motion BCS? Why has the static BCS not fallen by the wayside of evolution due to disuse? In particular, if nature could design a MOC filter that is sensitive to direction of motion and insensitive to direction of contrast, then why did the SOC filter evolve, in which

insensitivity to direction of contrast comes only at the cost of a loss of sensitivity to direction of motion? This question is perplexing given that animals are usually in relative motion with respect to their visual environment, and that simple cells in V1 are already sensitive to direction of motion.

An answer to this question can be derived by first noting that the SOC filter design described by Grossberg and Mingolla (1985b), Grossberg (1987c), and Grossberg and Marshall (1989) is incomplete. That design omits the processes that would be needed to make the SOC filter sensitive to transient changes in the input pattern. An analysis of how to correct this omission leads herein to an enhanced FACADE theory in which the static BCS and the motion BCS may be viewed as parallel subsystems of a single total system. I predict that this total system unfolds during the development of the visual cortex as an expression of an underlying symmetry principle, called FM symmetry (F = form, M = motion). Many manifestations of symmetry are familiar from our daily experiences with the physical world, and symmetry principles provide an important predictive and explanatory tool in the modern physical sciences. Here I suggest that the static form perception and motion form perception systems are not independent modules that obey different rules. Rather, they express two sides of a unifying organizational principle that is predicted to control the development of visual cortex.

5. A Symmetry Principle for Cortical Development: Sustained-Transient Gating, Opponent Processing, and Insensitivity to Direction of Contrast

This symmetry principle is predicted to control the simultaneous satisfaction of three constraints: (1) multiplicative interaction, or gating, of all combinations of

Table 1
Levels of MOC Filter

Level 1	Input Pattern Preprocess to discount the illuminant
Level 2	Sustained Response Cells Time-averaged and shunted signals from rectified outputs of spatially filtered oriented receptive fields
Level 3	Transient Response Cells Rectified outputs of time-averaged and shunted signals from unoriented change-sensitive cells
Level 4	Local Motion Detectors Pairwise gating of sustained and transient response combinations Sensitive to direction of contrast Sensitive to direction of motion
Level 4 → 5	Long-Range Gaussian Filter
Level 5	Motion-Direction Detectors Contrast-enhancing competition Insensitive to direction of contrast Sensitive to direction of motion

sustained-cell and transient-cell output signals to form four sustained-transient cell types; (2) symmetric organization of these sustained-transient cell types into two opponent on-cell and off-cell pairs, such that (3) output signals from all the opponent cell types are independent of direction of contrast.

Multiplicative gating of sustained cells and transient cells is shown below to generate change-sensitive receptive-field properties of oriented on-cells and off-cells within the static BCS, and direction-sensitive cells within the motion BCS. The constraint that output signals be independent of direction of contrast enables both the static BCS and the motion BCS to generate emergent boundary segmentations along image contrast reversals.

The summary above suggests how the static-form and motion-form systems may both arise. This discussion does not, however, disclose how these systems control different perceptual properties whose behavioral usefulness has preserved their integrity throughout the evolutionary process. The following behavioral implications of the symmetry principle will be explained herein.

6. Different Geometries for Perception of Static Form and Motion Form

We are all so familiar with the different geometries for processing static *orientations* and motion *directions* that we often take them for granted. For example, we all take for granted that the opposite *orientation* of “vertical” is “horizontal,” a difference of 90° ; yet the opposite *direction* of “up” is “down,” a difference of 180° . Why are the perceptual symmetries of static form and motion form different?

7. Negative Afterimages via Antagonistic Rebound

A clue is provided by considering how the 90° and 180° symmetries are reflected in percepts of negative afterimages. These symmetries suggest an opponent organization whereby orientations that differ by 90° are grouped together while directions that differ by 180° are grouped together. Negative aftereffects illustrate a key property of this opponent organization. For example, after sustained viewing of a radial input pattern, looking at a uniform field triggers a percept of a circular afterimage (MacKay, 1957). The orientations within the input and the circular afterimage differ from each other by 90° . After sustained viewing of a downwardly moving image, looking at a uniform field triggers a percept of an upwardly moving afterimage, as in the waterfall illusion (Sekuler, 1975). The directions within the downward input and the upward afterimage differ from each other by 180° .

In summary, the geometries of both static-form perception and motion-form perception include an opponent organization in which offset of the input pattern after sustained viewing triggers onset of a transient *antagonistic rebound*, or activation of the opponent channel.

8. Resonance versus Reset: Cooperative Feature Linking Without Destructive Smearing

Antagonistic rebound within opponent channels is needed to control the complementary perceptual processes of *resonance* and *reset*. Within the CC loop (Figure 1), positive feedback signals between the hypercomplex cells and bipole cells can cooperatively link similarly oriented features at approximately colinear locations into emergent boundary segmentations (Grossberg & Mingolla, 1985a, 1985b, 1987). Several neurophysiological laboratories (Eckhorn et al., 1988; Gray, Konig, Engel, & Singer, 1989) have supported the prediction that such cooperative linking occurs between cortical representations of similarly oriented features.

In the original Grossberg and Mingolla computer simulations of this phenomenon, a lumped version of the CC loop was used in which only “fast” variables were included, for simplicity. In this approximation, the cooperative linking process approaches an equilibrium configuration through time (Figure 6). My student David Somers and I have shown, however, that unlumping the CC loop model by introducing “slow” variables enables emergent segmentations to generate resonant standing waves in which cooperatively linked features oscillate in phase with one another (Grossberg & Somers, 1990). The equilibrium configurations are a limiting, or singular, approximation of these standing waves.

In Grossberg (1976, 1978), it was predicted that perceptual codes in the visual cortex would express themselves either as standing waves—if enough slow variables were operative—or as (approximate) equilibria if they were not. In the Eckhorn et al. (1988) and Gray et al. (1989) experiments, standing waves were reported.

Within the BCS, such a resonant segmentation, whether in standing wave or equilibrium form, derives from the positive-feedback interactions between hypercomplex on-cells and bipole cells of the CC loop. These positive-feedback interactions selectively amplify and sharpen the globally “best” cooperative grouping and provide the activation for inhibiting less favored groupings. The positive-feedback interactions also subserve the coherence, hysteresis, and structural properties of the emergent segmentations.

The positive feedback can, however, maintain itself for a long time after visual inputs terminate. Thus, the very existence of cooperative linking could seriously degrade perception by maintaining long-lasting positive afterimages, or smearing, of every percept.

Although some smearing can occur, it is known to be actively limited by inhibitory processes that are triggered by changing images (Hogben & DiLollo, 1985). Herein I suggest how antagonistic rebounds between opponently organized on-cells and off-cells can actively inhibit CC loop resonances when the input pattern changes. This inhibitory process *resets* the resonance and enables the CC

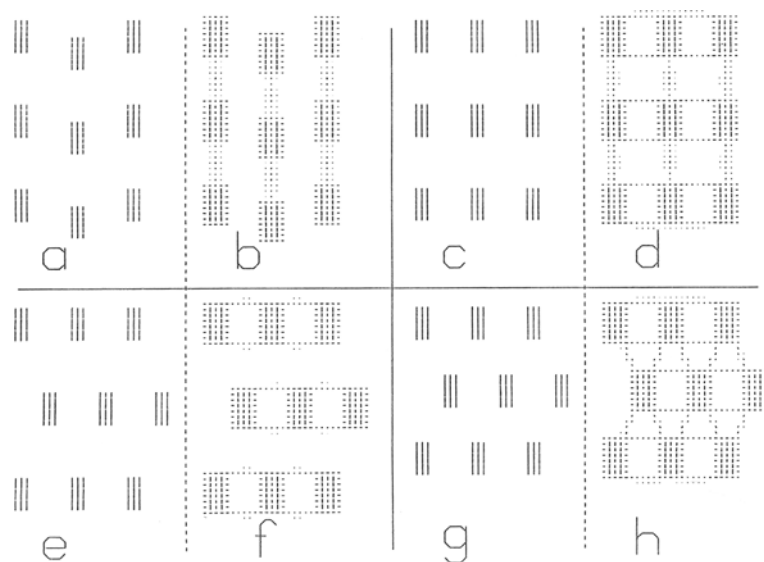


Figure 6. Computer simulations of processes underlying textural grouping: The length of each line segment is proportional to the activation of a network node responsive to 1 of 12 possible orientations. Parts a, c, e, and g display the activities of oriented cells that input to the CC loop. Parts b, d, f, and h display equilibrium activities of oriented cells at the second competitive stage of the CC loop. A pairwise comparison of (a) with (b), (c) and (d), and so on, indicates the major groupings sensed by the network. (From "Neural dynamics of surface perception: Boundary webs, illuminants, and shape-from-shading" by S. Grossberg and E. Mingolla, 1987, *Computer Vision, Graphics, & Image Processing*, 37, p. 133. Copyright 1987 by Academic Press. Reprinted by permission.)

loop to flexibly establish new resonances in response to rapidly changing scenes.

In summary, the symmetry principle that is predicted to control the parallel development of the static-form and motion-form systems enables these systems to rapidly reset their resonant segmentations in response to rapidly changing inputs.

9. Resonance and Reset Control Stable Autonomous Learning

The control of complementary states of resonance and reset during emergent segmentation within the BCS appears to be a special case of a more general principle of brain design. Adaptive resonance theory, or ART, clarifies how the brain can continue to learn new perceptual and cognitive recognition codes throughout life without experiencing unselective forgetting of previously learned, but still effective, recognition codes (Carpenter & Grossberg, 1987a, 1987b, 1988, 1990; Grossberg, 1976, 1980, 1982, 1987a, 1988). The ability to stably preserve previous memories while engaging in rapid new learning is called a solution to the *stability-plasticity dilemma*.

ART solves the stability-plasticity dilemma by suggesting how the brain autonomously controls complementary states of resonance and reset. Within ART, the resonant state focuses attention upon predictive groupings of perceptual features. This attentive resonant state also triggers learning of new recognition codes or selective refine-

ment of previously learned recognition codes. The reset event drives a rapid-search, or hypothesis-testing, cycle for more appropriate recognition codes when top-down learned expectations do not adequately match bottom-up perceptual data. This hypothesis-testing process prevents unselective forgetting of recognition codes from occurring by rapidly initiating reset and search before the bottom-up data can be associated with the incorrect recognition category. Within ART, an opponent organization capable of antagonistic rebound helps to trigger the reset events that drive the search process, and the same model of opponent processing, called a *gated dipole*, is used in both the BCS and ART reset schemes (Section 22).

BCS resonance and reset play a functional role similar to ART resonance and reset. In fact, the prediction that standing waves of cooperatively linked features should exist in visual cortex was made in the context of ART resonance (Grossberg, 1976, 1978). It remains to directly test whether the standing waves reported by Eckhorn et al. (1988) and Gray et al. (1989) utilize BCS bipole cells, as suggested here, or more general ART adaptive filters. This can be accomplished by combining the tests of Eckhorn et al. (1988) and Gray et al. (1989) with those of von der Heydt et al. (1984). In both the BCS and ART, the resonance subserves a perceptual event that may be attentively modulated. In particular, top-down learned ART expectations from, say, inferotemporal cortex are predicted to attentively modulate BCS segmentations that are pre-

dicted to arise in cortical area V2 (Grossberg, 1987b, Figure 2; Grossberg & Mingolla, 1985b, Figure 1). In ART, resonance triggers learning. I hypothesize that BCS resonance also controls a learning process. In the BCS, such learning would enable appropriate synaptic linkages to be selected and stably learned between bipole cells and hypercomplex cells of the CC loop. This type of learning would be driven by statistical regularities, such as edges, curves, and angles, in the visual environment. Marshall (1990) has, in fact, already used ART learning principles to train the synapses of a motion segmentation network. His model does not, however, use a CC loop and cannot yet explain various data which CC loop dynamics can explain. It remains an open problem to demonstrate adaptive tuning of CC loop synapses during a resonant boundary segmentation.

10. Combining Rapid Reset and Spatial Impenetrability Predictions

The network design that controls rapid reset of a CC loop resonance also constrains which combinations of features can resonate together, and thereby helps to structure the geometry of perceptual space. In particular, rapid reset of a resonating segmentation uses on-cells and off-cells of a given orientation that generate excitatory inputs and inhibitory inputs, respectively, to bipole cells of the same orientation (Figure 1). When on-cells lose their input, an antagonistic rebound activates off-cells that inhibit bipole cells and terminate the resonance.

Grossberg and Mingolla (1985b) have shown that the same mechanism can also generate the property of *spatial impenetrability* whereby emergent segmentations, during the resonance phase, are prevented from penetrating figures whose boundaries are built up from noncolinear

orientations (Figure 7). In particular, in a cartoon drawing of a person standing in a grassy field, the horizontal contours where the ground touches the sky do not generate horizontal emergent boundaries that cut the person's vertical body in half. This property follows from the fact that vertical on-cells inhibit vertical off-cells, which disinhibit horizontal off-cells (Figure 1). The horizontal off-cells, in turn, inhibit horizontal bipole cells, and thereby undermine horizontal segmentations that might otherwise have penetrated the vertical figure.

The hypothesis that these reset and impenetrability mechanisms are one and the same may be tested by variants of the prediction that sudden offset of a previously sustained figure that contains many vertically oriented lines may facilitate, rather than block, the propagation of horizontal emergent boundary segmentations between the horizontally oriented lines that surround the location of the figure on both sides. Such a facilitation would be due to antagonistic rebounds that activate horizontal orientations when the vertical orientations terminate. Then these horizontal orientations could cooperate with the horizontal orientations of the background to facilitate formation of a horizontal segmentation that spans the region where the vertical figure had been.

11. Perception of Moving Form in Depth: The V1 → V2 → MT Pathway

An additional consequence of the symmetry principle clarifies why an indirect, V1 → V2 → MT, cortical pathway from V1 to MT exists in addition to the direct V1 → MT pathway (DeYoe & van Essen, 1988). Outputs from the MOC filter sacrifice a measure of orientational specificity in order to effectively process direction of motion. However, precisely oriented binocular matches are

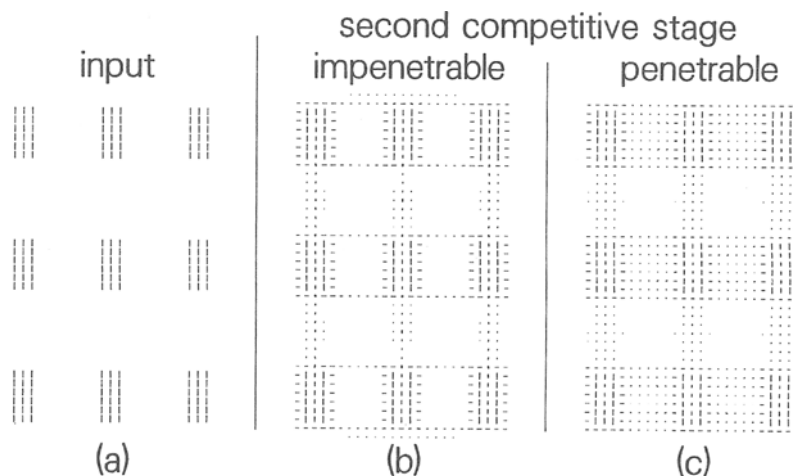


Figure 7. Computer simulations of processing underlying textural grouping. The length of each line segment is proportional to the activation of a network node responsive to 1 of 12 possible orientations. Part a displays the activity of oriented cells that input to the CC loop. Part b displays the groupings sensed by our actual model network. Part c displays the resulting flooding of boundary activity that occurs when the model's mechanism for spatial impenetrability is removed. (From "Computer Simulation of Neural Networks for Perceptual Psychology" by S. Grossberg and E. Mingolla, 1986, *Behavior Research Methods, Instruments, & Computers*, 18, p. 606. Copyright 1986 by Psychonomic Society Inc. Reprinted by permission.)

important in the selection of cortical cells that are tuned to the correct binocular disparities (von der Heydt, Hányi, & Dürsteler, 1981). The static BCS can carry out such precise oriented matches; the motion BCS cannot. This fact suggests that a pathway from the static BCS to the motion BCS exists in order to help the motion BCS to generate its motion segmentations at correctly calibrated depths.

Such a pathway needs to arise after the level of BCS processing at which cells capable of binocular fusion are chosen and binocularly rivalrous cells are suppressed. This occurs within the hypercomplex cells and bipole cells of the static BCS (Grossberg, 1987c; Grossberg & Marshall, 1989), and hence within the model analog of prestriate cortical area V2 (Figure 8). Thus the existence of a pathway from V2 to MT is consistent with the different functional roles of the static BCS and the motion BCS.

According to this reasoning, the V2 → MT pathway should occur at a processing stage prior to the one at which several orientations are pooled into a single direction of motion within each spatial scale. Thus, the pathway ends in the MOC filter at a stage no later than Level 4 in Figure 5. Such a pathway would join like orientations within like spatial scales between the static BCS and the motion BCS. It would thereby enhance the activation within the motion BCS of those spatial scales and orientations that are binocularly fused within the static BCS.

12. Apparent Motion of Illusory Figures

This interpretation of the V2 → MT pathway helps to explain the percept of apparent motion of illusory figures—a type of “doubly illusory” percept. Ramachandran, Rao, and Vidyasagar (1973) and Ramachandran (1985) have studied this phenomenon using the display shown in Figure 9. Frame 1 of this display generates the percept of an illusory square using a Kanizsa figure. Frame 2 generates the percept of an illusory square using a different combination of image elements. When Frame 2 is flashed on after Frame 1 shuts off, the illusory square is seen to move continuously from its location in Frame 1 to its location in Frame 2. Because matching of image elements between the two frames is impossible, the experiment demonstrates that the illusory square, not the image elements that generate it, is undergoing apparent motion.

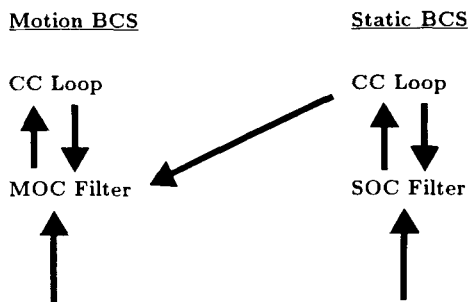


Figure 8. Model analog of V1 → V2 → MT pathway: Stereosensitive emergent segmentations from the static CC loop help to select the depthfully correct combinations of motion signals in the MOC filter.

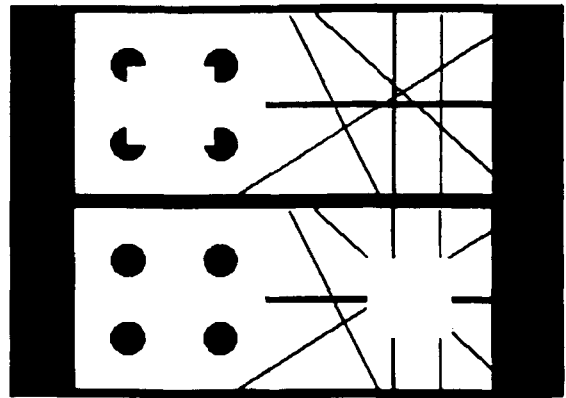


Figure 9. Images used to demonstrate that apparent motion of illusory figures arises through interactions of the static illusory figures, not from the inducing elements. Frame 1 (row 1) is temporally followed by Frame 2 (row 2). (From “Apparent motion of subjective surfaces” by V. S. Ramachandran, 1985, *Perception*, 14, p. 129. Copyright 1985 by Plon Ltd. Reprinted by permission.)

This phenomenon can be explained using the pathway from the CC loop of the static BCS to Level 4 of the MOC filter. The CC loop is capable of generating an illusory square in response to Frame 1 and Frame 2 (Grossberg & Mingolla, 1985b; van Allen & Kolodzy, 1987). Successive inputs to Level 4 of the MOC filter can induce continuous apparent motion if they are properly timed and spatially arranged (Grossberg & Rudd, 1989b, 1990a), as I will indicate below, in Section 15. When this happens, the two static illusory squares can induce a continuous wave of apparent motion at Level 5 of the MOC filter.

This explanation of apparent motion of illusory figures can be used to test whether the V2 → MT pathway plays the role suggested above. One possible approach is to train a monkey to respond differently when the two illusory figures appear to move and when they do not. Then a (reversible) lesion of the V2 → MT pathway should abolish the former behavior but not the latter.

13. Augmenting the Static BCS

The design of the motion BCS, and the symmetry principle that combines the static BCS and motion BCS into a unified theory, both came into view by noting and correcting incomplete features of the static BCS model that was introduced by Grossberg and Mingolla (1985b, 1987). These features can be understood by inspecting Figure 1.

A. *Insensitivity to direction of motion.* As shown in Figure 1, although the simple cells of the static BCS are sensitive to direction of contrast, or contrast polarity, the complex cells are rendered insensitive to direction of contrast by receiving inputs from pairs of simple cells with opposite direction of contrast. Such a property is also true of the simple cells and complex cells in area V1 (DeValois et al., 1982; Poggio, Motter, Squatrito, & Trotter, 1985; Thorell, DeValois, & Albrecht, 1984).

This property is useful for extracting boundary structure that is independent of contrast polarity, as in Figures 2 and 3. As remarked in Section 3, however, the output of the SOC filter is unable to differentiate direction of motion. A key property of the MOC filter described in Section 14 is that it joins the property of insensitivity to direction of contrast, which is also needed to process moving segmentations, with sensitivity to direction of motion. The fact that a modest change of the static BCS leads to a motion BCS that can explain a large body of data concerning motion perception provides additional support for both the static BCS model and the motion BCS model by showing that both models may be considered variations on a single neural architectural theme.

Although this modification of the static BCS is computationally modest, it is based upon a conceptually radical departure of FACADE theory from previous vision models. Indeed, the property of insensitivity to direction of contrast in the static BCS reflects one of the fundamental new insights of FACADE theory. Insensitivity to direction of contrast is possible within the BCS because all boundary segmentations within the BCS are perceptually invisible. Visibility is a property of the FCS, whose computations are sensitive to direction of contrast. A vision theory built up from independent processing modules could not articulate the heuristics or the mechanisms of the motion BCS, because it could not articulate the fact that the BCS and FCS are computationally *complementary* subsystems of a single larger system rather than being independent modules for the processing of form and color (Grossberg et al., 1989).

B. Insensitivity to input transients. In Figure 1, the simple cells of the SOC filter are modeled as oriented sustained-response cells. Sustained-response cells can respond with a sustained output to a constant input. In contrast, simple cells *in vivo* are sensitive to transient changes in input patterns, including changes due to moving images.

When the SOC filter is modified to be sensitive to image transients, it may be compared with the MOC filter, which is obviously also sensitive to image transients. Such a comparison led to the discovery of the symmetry principle, and to the realization that both filters might be viewed as parallel halves of a larger system design.

The answer to how the SOC filter computes image transients was suggested by another incomplete feature of the original SOC filter model.

C. No simple and complex off-cells. The simple cells and complex cells in Figure 1 are all on-cells; they are activated when external inputs turn on. No simple or complex off-cells are represented. In contrast, the hypercomplex cells in Figure 1 include both on-cells and off-cells. This asymmetry in the network raises the question of how to design simple off-cells and complex off-cells to interact with the hypercomplex off-cells.

It turns out that problems B and C, which seem to be two distinct problems, have the same solution. Multiplicative coupling of transient on-cells and transient off-cells with oriented sustained on-cells define simple on-cells and off-cells, as well as complex on-cells and off-cells, that are sensitive to image transients.

In summary, including off-cells and the property of sensitivity to image transients in the SOC filter leads to a more symmetric SOC filter model which, when compared with the MOC filter model, reveals a deeper principle of symmetric design. The remainder of this article shows how problems A-C may be solved, and describes how these solutions lead to the data implications summarized in Sections 1-12. On a first reading, the reader may wish to skip to Section 17 in order to follow the qualitative analysis of how the SOC filter may be augmented and the FM symmetry described.

14. Design of a MOC Filter

This section suggests a solution to the problem raised in Section 13A. The equations for a one-dimensional MOC filter were described in Grossberg and Rudd (1989b). The MOC filter's five processing levels are described qualitatively below for the more general 2-D case. The equations used for the 1-D theory are also described to provide a basis for rigorously defining the augmented SOC filter and FM symmetry.

Level 1: Preprocess input pattern. The image is preprocessed before it activates the MOC filter. For example, it is passed through a shunting on-center off-surround net to compensate for variable illumination, or to "discount the illuminant" (Grossberg & Todorović, 1988).

In the 1-D theory, I_i denotes the input at position i .

Level 2: Sustained-cell short-range filter. Four operations occur here, as illustrated in Figure 10.

(1) *Space average.* Inputs are processed by individual sustained cells with oriented receptive fields.

(2) *Rectify.* The output signal from a sustained cell grows with its activity above a signal threshold.

(3) *Short-range spatial filter.* A spatially aligned array of sustained cells with like orientation and direction of contrast pool their output signals to activate the next cell level. This spatial pooling plays the role of the short-range

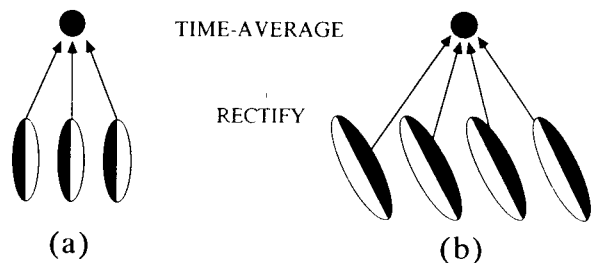


Figure 10. The sustained-cell short-range filter combines several spatially contiguous receptive fields of like orientation via a spatial filter with a fixed directional preference. The orientation perpendicular to the direction is preferred, but nonorthogonal orientations can also be grouped in a prescribed direction.

motion limit D_{max} (Braddick, 1974). The breadth of spatial pooling scales with the size of the simple-cell receptive fields (Figures 10a and 10b). Thus, " D_{max} " is not independent of the spatial frequency content of the image (Anderson & Burr, 1987; Burr, Ross, & Morrone, 1986; Nakayama & Silverman, 1984, 1985), and is not a universal constant.

The direction of spatial pooling may not be perpendicular to the oriented axis of the sustained-cell receptive field (Figure 10b). The target cells are thus sensitive to a movement direction that may not be perpendicular to the sustained cell's preferred orientation.

(4) *Time average*. The target-cell time-averages the directionally sensitive inputs that it receives from the short-range spatial filter. This operation has properties akin to the "visual inertia" during apparent motion that was reported by Anstis and Ramachandran (1987); see Figure 19.

In the 1-D theory, only horizontal motions are considered. It therefore suffices to consider two types of such cells that filter the input pattern I_i , one of which responds to a light-dark luminance contrast (designated by L , for left) and the other of which responds to a dark-light luminance contrast (designated by R , for right). Output pathways from like cells converge (Figure 10) to generate inputs J_{iL} and J_{iR} at each position i . The activity x_{ik} of the i th target cell at Level 2 obeys a membrane equation,

$$\frac{d}{dt} x_{ik} = -Ax_{ik} + (1 - Bx_{ik})J_{ik}, \quad (1)$$

where $k = L, R$, which performs a time average of the input J_{ik} .

Level 3: Transient-cell filter. In parallel with the sustained-cell filter, a transient-cell filter reacts to input increments (on-cells) and decrements (off-cells) with positive outputs (Figure 11). This filter uses four operations too:

(1) *Space average*. This is accomplished by a receptive field that sums inputs over its entire range.

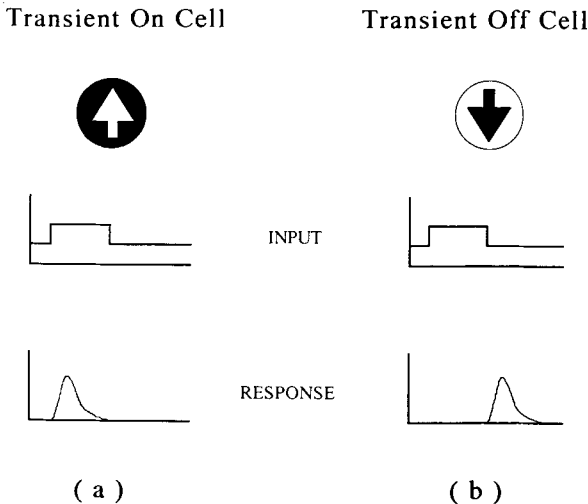


Figure 11. The transient-cell filter consists of on-cells which react to input increments and off-cells which react to input decrements.

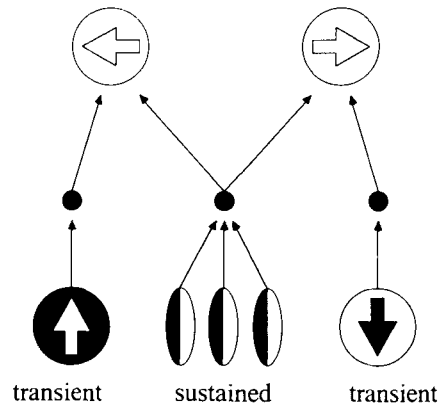


Figure 12. Sustained-transient gating generates cells that are sensitive to direction of motion as well as to direction of contrast.

(2) *Time average*. This sum is time-averaged to generate a gradual growth and decay of total activation.

(3) *Transient detector*. The on-cells are activated when the time average increases (Figure 11a). The off-cells are activated when the time average decreases (Figure 11b).

(4) *Rectify*. The output signal from a transient cell grows with its activity above a signal threshold.

In the 1-D theory, the activities of the transient cells were computed as the rectified time derivatives of an unoriented space-time average x_i of the input pattern I_i . The time derivative is given by the membrane equation

$$\frac{d}{dt} x_i = -Cx_i + (D - Ex_i) \sum_j I_j F_{ji}, \quad (2)$$

where F_{ji} is the unoriented spatial kernel that represents a transient-cell receptive field.

Positive and negative half-wave rectifications of the time derivative were performed independently by defining

$$y_i^+ = \max\left(\frac{d}{dt} x_i - \Gamma, 0\right), \quad (3)$$

and

$$y_i^- = \max\left(\Omega - \frac{d}{dt} x_i, 0\right), \quad (4)$$

where Γ and Ω are constant thresholds. The activity y_i^+ models the response of a transient on-cell; the activity y_i^- models the response of a transient off-cell.

Level 4: Sustained-transient gating yields direction-of-motion sensitivity and direction-of-contrast sensitivity. Maximal activation of a Level 2 sustained-cell filter is caused by image contrasts moving in either of two directions that differ by 180°. Multiplicative gating of each Level 2 sustained-cell output with a Level 3 transient-cell on-cell or off-cell removes this ambiguity (Figure 12). For example, consider a sustained-cell output from vertically oriented dark-light receptive fields that are joined together in the horizontal direction by the short-range spatial filter (Figure 10a). Such a sustained-cell output is maximized by a dark-light image contrast moving to the right or to

the left. Multiplying this Level 2 output with a Level 3 transient on-cell output generates a Level 4 cell that responds maximally to motion to the left.

In the 1-D theory, there are two types of sustained cells (corresponding to the two antisymmetric directions of contrast), and also two types of transient cells (the on-cells and the off-cells). Consequently, there are four types of gated responses that can be computed. Two of these produce cells that are sensitive to local rightward motion: the $(L, +)$ cells that respond to x_{iLY}^{\dagger} and the $(R, -)$ cells that respond to x_{iRY}^{\dagger} . The other two produce cells that are sensitive to local leftward motion: the $(L, -)$ cells that respond to $x_{iLY}^{\bar{\dagger}}$ and the $(R, +)$ cells that respond to $x_{iRY}^{\bar{\dagger}}$. All of these cells inherit a sensitivity to the direction of contrast of their inputs from the Level 2 sustained cells from which they are constructed.

The cell outputs from Level 4 are sensitive to direction of contrast. Level 5 consists of cells that pool outputs from Level 4 cells that are sensitive to the same direction of motion but to opposite directions of contrast.

Level 5: Long-range spatial filter and competition. Outputs from Level 4 cells sensitive to the same direction of motion but opposite directions of contrast activate individual Level 5 cells via a long-range spatial filter that has a Gaussian profile across space (Figure 13). This long-range filter groups together Level 4 cell outputs that are derived from Level 3 short-range filters with the same directional preference but different simple-cell orientations. Thus, the long-range filter provides the extra degree of freedom that enables Level 5 cells to function as "direction" cells rather than "orientation" cells.

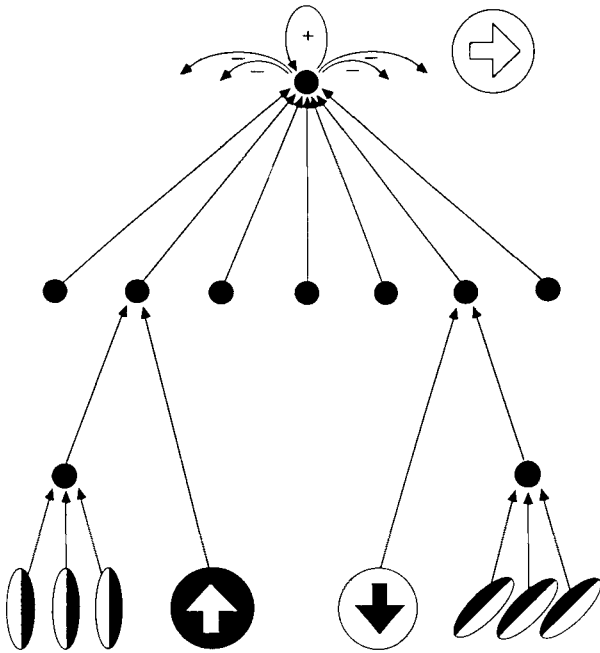


Figure 13. The long-range spatial filter combines sustained-transient cells with the same preference for direction of motion, including cells whose sustained-cell inputs are sensitive to opposite directions of contrast and to different orientations.

The long-range spatial filter broadcasts each Level 4 signal over a wide spatial range in Level 5. Competitive, or lateral inhibitory, interactions within Level 5 contrast-enhance this input pattern to generate spatially sharp Level 5 responses. A winner-take-all competitive network (Grossberg, 1973, 1982) can transform even a very broad input pattern into a focal activation at the position that receives the maximal input. The winner-take-all assumption is a limiting case of how competition can restore positional localization. More generally, we suggest that this competitive process partially contrast-enhances its input pattern to generate a motion signal whose breadth across space increases with the breadth of its inducing pattern. A contrast-enhancing competitive interaction has also been modeled at the complex cell level of the SOC filter (Grossberg, 1987c; Grossberg & Marshall, 1989). The Level 5 cells of the MOC filter are, in other respects too, computationally homologous to the SOC filter complex cells.

In the 1-D theory, we define the transformation from Level 4 to Level 5 by letting

$$r_i = x_{iLY}^{\dagger} + x_{iRY}^{\bar{\dagger}} \quad (5)$$

and

$$l_i = x_{iLY}^{\bar{\dagger}} + x_{iRY}^{\dagger} \quad (6)$$

be the total response of the local right-motion and left-motion detectors, respectively, at position i of Level 4. Signal r_i increases if either a light-dark or a dark-light contrast pattern moves to the right. Signal l_i increases if either a light-dark or a dark-light contrast pattern moves to the left.

These local motion signals are assumed to be filtered independently by a long-range operator with a Gaussian kernel,

$$G_{ji} = H \exp[-(j-i)^2/2K^2], \quad (7)$$

that defines the input fields of the Level 5 cells. Thus, there exist two types of direction-sensitive cells at each position i of Level 5. The activity at i of the right-motion-sensitive cell is given by

$$R_i = \sum_j r_j G_{ji}, \quad (8)$$

and the corresponding activity of the left-motion sensitive cell is given by

$$L_i = \sum_j l_j G_{ji}. \quad (9)$$

We assume that contrast-enhancing competitive, or lateral inhibitory, interactions within Level 5 generate the activities that encode motion information. In the simplest case, the competition is tuned to select that population whose input is maximal, as in

$$x_i^{(R)} = \begin{cases} 1 & \text{if } R_i > R_j, j \neq i \\ 0 & \text{otherwise,} \end{cases} \quad (10)$$

and

$$x_i^{(L)} = \begin{cases} 1 & \text{if } L_i > L_j, j \neq i \\ 0 & \text{otherwise.} \end{cases} \quad (11)$$

In the simulations summarized below, the above assumption was made for simplicity. The functions $x_i^{(R)}$ and $x_i^{(L)}$ change through time in a manner that idealizes the parametric properties of many apparent motion phenomena. See Grossberg and Rudd (1989b, 1990a) for details. More generally, we suggest that the competitive process idealized by Equations 10 and 11 performs a partial contrast enhancement of its input pattern and thereby generates a motion signal whose breadth across space increases with the breadth of its inducing pattern.

The total MOC filter design is summarized in Figure 5.

15. Continuous Motion Paths from Spatially Stationary Flashes

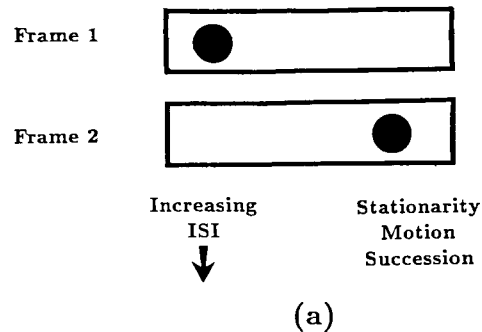
The model equations listed in Section 14 provide an answer to long-standing questions in the vision literature concerning why individual flashes do not produce a percept of long-range motion although long-range interaction between spatially discrete pairs of flashes can produce a spatially sharp percept of continuous motion. Such apparent motion phenomena are a particularly useful probe of motion mechanisms because they describe controllable experimental situations in which nothing moves, yet a compelling percept of motion is generated. For example, two brief flashes of light, separated in both time and space, create an illusion of movement from the location of the first flash to that of the second when the spatio-temporal parameters of the display are within the correct range (Figure 14a).

Outstanding theoretical issues concerning apparent motion include the resolution of a tradeoff that exists between the long-range spatial interaction that is needed to generate the motion percept and the localization of the perceived-motion signal that smoothly interpolates the inducing flashes. If a long-range interaction between the flashes must exist in order to generate the motion percept, then why is it not perceived when only a single light is flashed? Why are outward waves of motion-carrying signals not induced by a single flash? What kind of long-range influence is generated by each flash, such that a perceived-motion signal is triggered only when at least two flashes are activated? What kind of long-range influence from individual flashes can generate a smooth motion signal between flashes placed at variable distances from one another? How does the motion signal speed up to smoothly interpolate flashes that occur at larger distances but at the same time lag (Kolers, 1972)? How does the motion signal speed up to smoothly interpolate flashes when they occur at the same distance but with shorter time lags (Kolers, 1972)?

Variants of apparent motion include *phi motion*, or the *phi phenomenon*, whereby a "figureless" or "objectless" motion signal propagates from one flash to another; *beta motion*, whereby a well-defined form seems to move smoothly and continuously from one flash to the other;

TWO TYPES OF APPARENT MOTION DISPLAYS

TWO FLASH DISPLAY (Wertheimer, 1912)



TERNUS DISPLAY (Ternus, 1926)

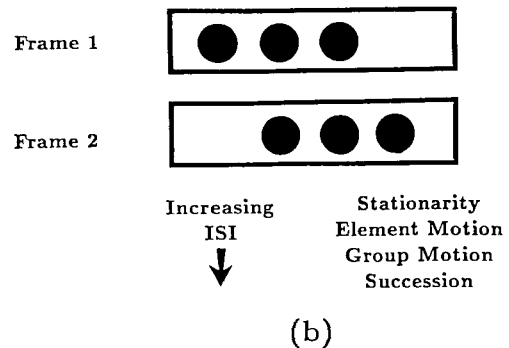


Figure 14. Two types of apparent motion displays in which the two frames outline the same region in space into which the dots are flashed at successive times: In (a), a single dot is flashed, followed by an interstimulus interval (ISI), followed by a second dot. At small ISIs, the two dots appear to flicker in place. At longer ISIs, motion from the position of the first dot to that of the second is perceived. (b) In the Ternus display, three dots are presented in each frame such that two of the dots in each frame occupy the same positions. At short ISIs, all the dots appear to be stationary. At longer ISIs, the dots at the shared positions appear to be stationary, while apparent motion occurs from the left dot in Frame 1 to the right dot in Frame 2. At still longer ISIs, the three dots appear to move from Frame 1 to Frame 2 as a group.

gamma motion, the apparent expansion at onset and contraction at offset of a single flash of light; *delta motion*, or backward motion from a more intense second flash to a less intense first flash; and *split motion*, or simultaneous motion paths from a single first flash to a simultaneous pair of second flashes (Bartley, 1941; Kolers, 1972).

Another well-known apparent motion display, which originated with Ternus (1926/1950), illustrates the fact that not only the existence of a motion percept, but also its figural identity may depend on subtle aspects of the display, such as the interstimulus interval, or ISI, between the offset of the first flash and the onset of the second flash (Figure 14b). In the Ternus display, a cyclic alter-

nation of two stimulus frames gives rise to competing visual movement percepts. In Frame 1, three elements are arranged in a horizontal row on a uniform background. In Frame 2, the elements are shifted to the right in such a way that the positions of the two leftmost elements in Frame 2 are identical to those of the two rightmost elements in Frame 1. Depending on the stimulus conditions, the observer will see either of two bistable motion percepts. Either the elements will appear to move to the right as a group between Frames 1 and 2 and then back again during the second half of a cycle of the display or, alternatively, the leftmost element in Frame 1 will appear to move to the location of the rightmost element in Frame 2, jumping across two intermediate elements which appear to remain stationary. The first percept is called "group" motion, and the second percept, "element" motion. At short ISIs, there is a tendency to observe element motion. At longer ISIs, there is a tendency to observe group motion.

Remarkably, formal analogs of all these and many other motion phenomena occur at Level 5 of the motion MOC filter in response to sequences of flashes presented to Level 1 (Grossberg & Rudd, 1989b, 1990a). Intuitively, a signal for motion will arise when a continuous wave of activation connects the locations corresponding to the flashes—that is, when a connected array of the functions $x_i^{(R)}$, $x_{i+1}^{(R)}$, $x_{i+2}^{(R)}$, ... are activated sequentially through time or, alternatively, the functions $x_i^{(L)}$, $x_{i-1}^{(L)}$, $x_{i-2}^{(L)}$, ... are activated sequentially through time. Each activation, $x_i^{(R)}$ or $x_i^{(L)}$, represents the peak, or maximal activity, of a broad spatial pattern of activation across the network.

The broad activation pattern (Figure 15b) is generated by the long-range Gaussian filter G_{ji} in Equation 7 in response to a spatially localized flash to Level 1 (Figure 15a). The sharply localized response function $x_i^{(R)}$ is due to the contrast-enhancing action of the competitive network within Level 5 (Figure 15c). A stationary localized $x_i^{(R)}$ response is hereby generated in response to a single flashing input.

In contrast, suppose that two input flashes occur with the following spatial and temporal separations. Let the positions of the flashes be $i=1$ and $i=N$. Let the activity $r_1(t)$, in Equation 5, caused by the first flash start to decay as the activity $r_N(t)$, in Equation 5, caused by the second flash starts to grow. Suppose, moreover, that the flashes are close enough for their spatial patterns $r_1 G_{1i}$ and $r_N G_{Ni}$ to overlap. Then the total input

$$R_i = r_1 G_{1i} + r_N G_{Ni} \quad (12)$$

to the i th cell in Level 5 can change in such a way that the maximum value of the spatial pattern $R_i(t)$ through time, namely $x_i^{(R)}(t)$ in Equation 10, first occurs at $i=1$, then $i=2$, then $i=3$, and so on until $i=N$. A percept of continuous motion from the position of the first flash to that of the second will result.

This basic property of the MOC filter is illustrated by the computer simulations from Grossberg and Rudd (1989b) summarized in Figures 16–18. Figure 16 depicts

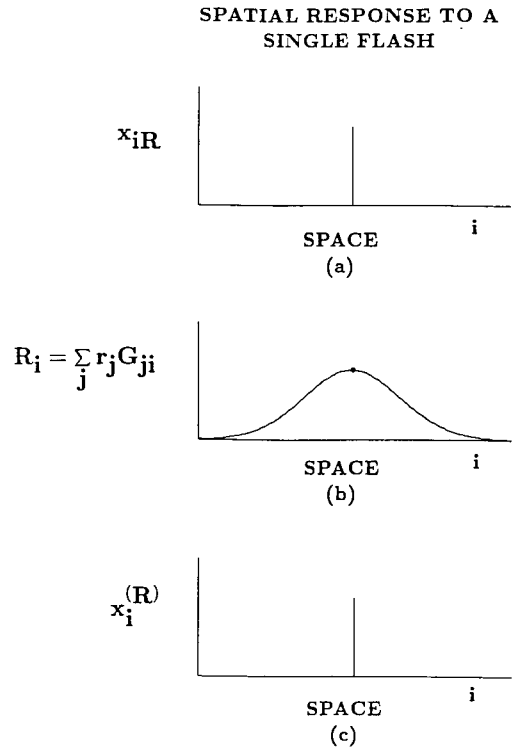


Figure 15. Spatial response of the MOC filter to a point input. (a) Sustained activity of a Level 2 cell. (b) Total input pattern to Level 5. (c) Contrast-enhanced response at Level 5. (From "A neural architecture for visual motion perception: Group and element apparent motion" by S. Grossberg and M. E. Rudd, 1989, *Neural Networks*, 2, p. 425. Copyright 1989 by Pergamon Press. Reprinted by permission.)

the temporal response to a single flash at Position 1 of Level 1. The sustained-cell response at Position 1 of Level 2 undergoes a gradual growth and decay of activation (Figure 16b), although the position of maximal activation in the input to Level 5 does not change through time (Figure 16c). The temporal decay of activation in Figure 16b may be compared with the "visual inertia" by Anstis and Ramachandran (1987, Figure 6) during their experiments on apparent motion (Figure 19).

Figure 17 illustrates an important implication of the fact that the Level 2 cell activations persist due to temporal averaging after their Level 1 inputs shut off. If a flash at Position 1 is followed, after an appropriate delay, by a flash at Position N , then the sustained response to the first flash [e.g., $x_{1R}(t)$] can decay while the response to the second flash [e.g., $x_{NR}(t)$] grows.

Assume, for simplicity, that the transient signals defined by Equations 3 and 4 are held constant, and consider how the waxing and waning of sustained-cell responses control the motion percept. Then the total input pattern R_i to Level 5 can change through time in the manner depicted in Figure 18. Each row of Figure 18a illustrates the total input to Level 5 caused, at a prescribed time t , by $x_{1R}(t)$ alone, by $x_{NR}(t)$ alone, and by both

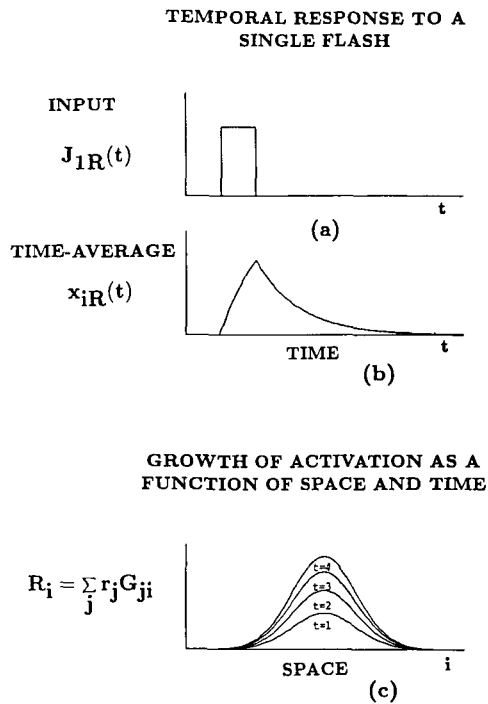


Figure 16. Temporal response of sustained-response cells to a point input: (a) The input is presented for a brief duration at location 1. (b) The activity of the sustained-response cell gradually builds up after input onset, then decays after input offset. (c) Growth of the input pattern to Level 5 through time with transient-cell activity held constant. The activity pattern retains a Gaussian shape centered at the location of the input. (From "A neural architecture for visual motion perception: Group and element apparent motion" by S. Grossberg and M. E. Rudd, 1989, *Neural Networks*, 2, p. 429. Copyright 1989 by Pergamon Press. Reprinted by permission.)

flashes together. Successive rows plot these functions at equally spaced later times. As $x_{1R}(t)$ decays and $x_{NR}(t)$ grows, the maximum value of $R_i(t)$ moves continuously to the right. Figure 18b depicts the position $x_i^{(R)}(t)$ of the maximum value at the corresponding times.

16. Feature Integration and Spatial Attention Shifts

The conditions under which such a traveling wave of activation can occur are proved in Grossberg and Rudd (1989b) to be quite general. The phenomenon can arise whenever a decaying trace of one activation adds to an increasing trace of a second activation via spatially long-range Gaussian receptive fields before the sum is contrast-enhanced. Such a traveling wave may, for example, subserve certain shifts in spatial attention (Eriksen & Murphy, 1987; LaBerge & Brown, 1989; Remington & Pierce, 1984). It remains for future analyses to determine whether discrete jumps of spatial attention and continuous shifts of attention may receive a unified analysis in terms of the same formal constraints that explain how discrete flashes or continuous apparent motion are perceived.

Within the general theoretical framework of FACADE theory and ART, mechanisms of "preattentive feature integration" by BCS segmentation and FCS filling-in, "at-

tentive feature integration" through resonance with top-down learned ART expectations after reset and search terminate, and shifts in spatial attention due to mechanisms similar in formal properties to preattentive mechanisms of motion perception come together in a unified computational theory. Such a theory provides an alternative framework to Treisman's seminal account of feature integration (Treisman & Gelade, 1980; Treisman & Souther, 1985), whose conceptual difficulties and the demands of new data have gradually led to qualitative theoretical approaches more in harmony with the quantitative mechanisms of FACADE theory and ART (Duncan & Humphreys, 1989; Nakayama & Silverman, 1986; Pashler, 1987; Treisman & Gormican, 1988).

17. Design of Simple On-Cells and Off-Cells

This section suggests a solution to the problems raised in Sections 13B and 13C. As noted there, hypercomplex cells in Figure 1 are organized into opponent on-cells and off-cells, yet the SOC filter explicitly depicts only pathways to the hypercomplex on-cells from the simple on-cells via complex on-cells. Moreover, all of these cells are of the sustained-cell type. Interactions with simple off-cells, complex off-cells, and transient cells are not described. It is now shown how multiplicative gating of sustained cells with transient on-cells and transient off-

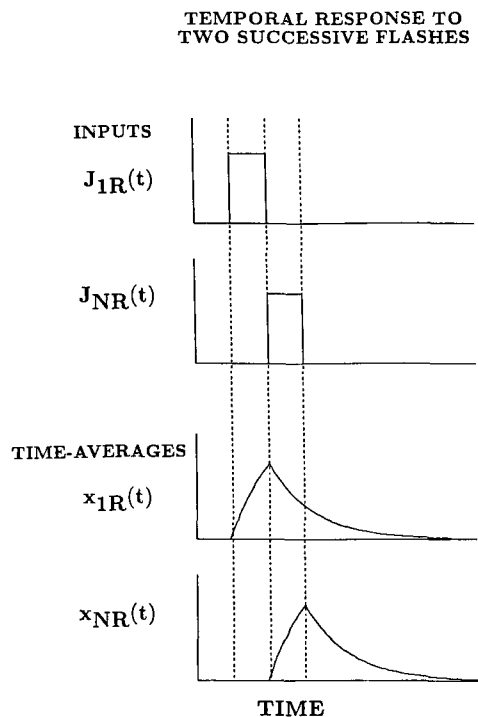


Figure 17. Temporal response of the sustained-response cells at Level 2 to two successive point inputs. One input is presented briefly at location 1, followed by a second input at location N. For an appropriately timed display, the decaying response at position 1 overlaps the rising response at position N. (From "A neural architecture for visual motion perception: Group and element apparent motion" by S. Grossberg and M. E. Rudd, 1989, *Neural Networks*, 2, p. 429. Copyright 1989 by Pergamon Press. Reprinted by permission.)

LONG-RANGE INTERACTION

SHARP MOTION SIGNAL

$$R_i = \sum_j r_j G_{ji}$$

$$x_i^{(R)} = \begin{cases} 1 & \text{if } R_i > R_j, j \neq i \\ 0 & \text{otherwise} \end{cases}$$

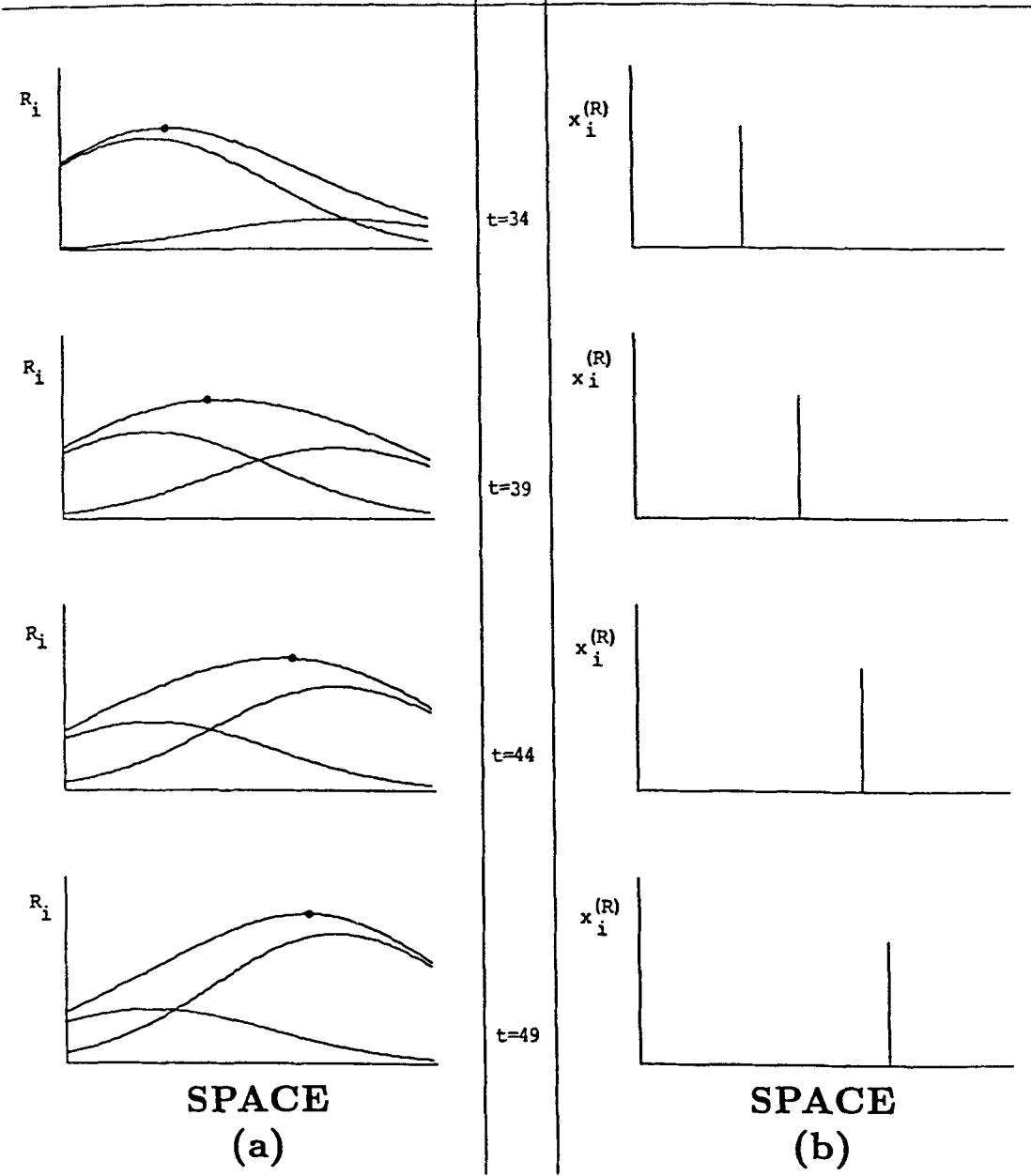


Figure 18. MOC filter simulation in response to a two-flash display. Successive rows correspond to increasing times: (a) The two lower curves in each row depict the total input to Level 5 caused by each of the two flashes. The input due to the left flash decreases, while the input due to the right flash increases. The total input due to both flashes is a traveling wave whose maximum value moves from the location of the first flash to that of the second flash. (b) Position of the contrast-enhanced response at Level 5. (From "A neural architecture for visual motion perception: Group and element apparent motion" by S. Grossberg and M. E. Rudd, 1989, *Neural Networks*, 2, p. 430. Copyright 1989 by Pergamon Press. Reprinted by permission.)

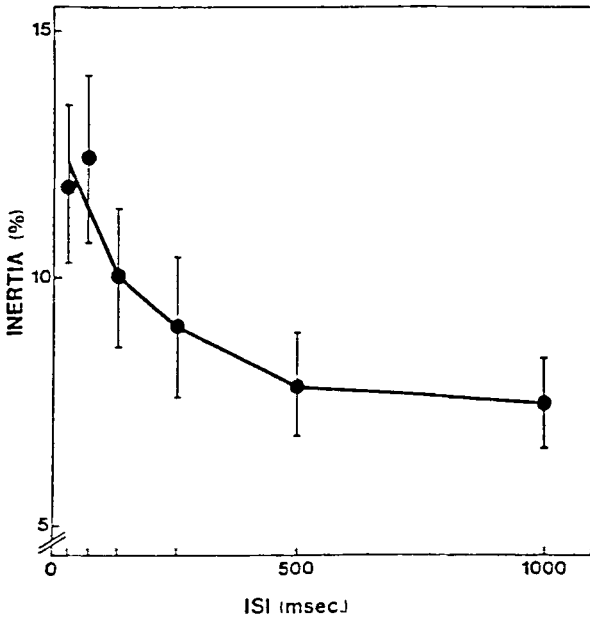


Figure 19. Strength of visual inertia as a function of the timing of dots that prime the direction of a subsequent apparent motion. (From "Visual inertia in apparent motion" by S. Anstis and V. S. Ramachandran, 1987, *Vision Research*, 27, p. 759. Copyright 1987 by Pergamon Press. Reprinted by permission.)

cells solves both of these problems, and reveals that the modified SOC filter and the MOC filter are parallel sub-systems of a symmetric total system.

In Figures 1 and 4, pairs of like-oriented simple cells that are sensitive to opposite directions of contrast input to a single complex cell that is insensitive to direction of contrast. Our task is to preserve this fundamental property while rendering the simple cells sensitive to image transients and defining both on-cells and off-cells of simple and complex types. To define simple on-cells that are sensitive to image transients, let a transient on-cell, as defined in Equation 3, multiply each sustained cell in the pair of like-oriented cells depicted in Figures 1 and 4. This gives rise to a pair of like-oriented simple on-cells (Figure 20a) that are sensitive to opposite directions of contrast and are activated when properly oriented and positional inputs are turned on. In the 1-D model notation of Equations 1 and 3, these cell responses are defined by x_{iLY}^\dagger and x_{iRY}^\dagger rather than by x_{iL} and x_{iR} alone. As in Figures 1 and 4, a pair of simple on-cells with like-orientation but opposite direction of contrast inputs to a complex on-cell, as in Figure 20a. The complex on-cell is defined by summing the rectified outputs of the simple on-cells, as in the equation

$$c_i^\dagger = x_{iLY}^\dagger + x_{iRY}^\dagger. \quad (13)$$

Likewise, a pair of simple off-cells can be defined by gating the pair of like-oriented sustained cells in Figure 1 with a transient off-cell, as defined in Equation 4. The

pair of simple off-cell responses is thus defined by x_{iLY}^\dagger and x_{iRY}^\dagger , rather than by x_{iL} and x_{iR} alone. The off-cells are activated when properly oriented and positioned inputs shut off. Such a pair of simple off-cells is depicted in Figure 20b, where it gives rise to a complex off-cell through the interaction

$$c_i^\dagger = x_{iLY}^\dagger + x_{iRY}^\dagger. \quad (14)$$

By construction, both complex on-cells and off-cells are insensitive to direction of contrast.

Let the complex on-cell in Figure 20a input to hyper-complex on-cells, as in Figure 1. In a similar fashion, let the complex off-cell in Figure 20b input to the hyper-complex off-cells in Figure 1. The process of gating sustained cells by transient cells to generate on-cells and off-cells in the static BCS thus makes the overall design of this architecture more symmetric by showing how simple and complex on-cells and off-cells fit into the scheme.

18. FM Symmetry

The above refinement of the SOC filter merely adds sensitivity to image transients in a manner consistent with Figure 1. Having done so, a comparison of the modified SOC filter with the MOC filter reveals the FM symmetry

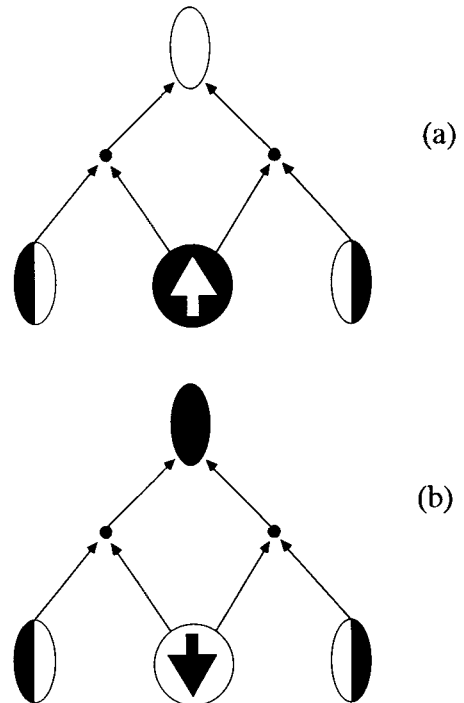


Figure 20. (a) A complex/orientation/on-cell: Pairs of rectified sustained cells with opposite direction of contrast are gated by rectified transient on-cells to generate simple sustained-transient on-cells before the gated responses are added. (b) A complex/orientation/off-cell: Pairs of rectified sustained cells with opposite direction of contrast are gated by rectified transient off-cells to generate simple sustained-transient off-cells before the gated responses are added.

principle that was introduced in Section 5. FM symmetry is embodied in the following set of four equations, the first two from the MOC filter and the last two from the enhanced SOC filter.

Left-direction motion complex on-cell:

$$r_l = x_{iLY}^{\dagger} + x_{iRY}^{\bar{\dagger}} \tag{5}$$

Right-direction motion complex on-cell:

$$l_r = x_{iLY}^{\bar{\dagger}} + x_{iRY}^{\dagger} \tag{6}$$

Vertical-orientation static complex on-cell:

$$c_l^{\dagger} = x_{iLY}^{\dagger} + x_{iRY}^{\dagger} \tag{13}$$

Vertical-orientation static complex off-cell:

$$c_r^{\bar{\dagger}} = x_{iLY}^{\bar{\dagger}} + x_{iRY}^{\bar{\dagger}} \tag{14}$$

These equations describe all possible ways of symmetrically gating an opponent pair (x_{iL}, x_{iR}) of sustained cells with transient cells to generate two opponent pairs, $(c_l^{\dagger}, c_r^{\bar{\dagger}})$ and (r_l, l_r) , of output signals that are insensitive to direction of contrast. One opponent pair of outcomes $(c_l^{\dagger}, c_r^{\bar{\dagger}})$ contains cell pairs that are insensitive to direction of mo-

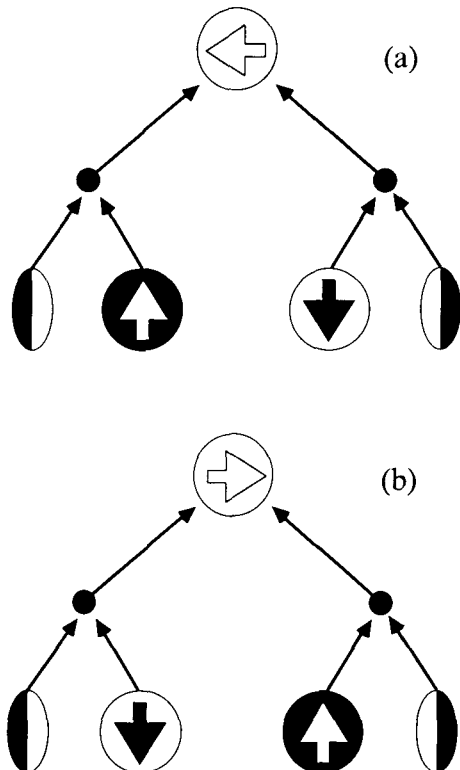


Figure 21. (a) A complex/direction/left cell: Pairs of rectified sustained cells with opposite direction of contrast are gated by pairs of rectified transient on-cells and off-cells, before the gated responses are added. (b) A complex/direction/right cell: Same as in (a), except that sustained cells are gated by the opposite transient cell.

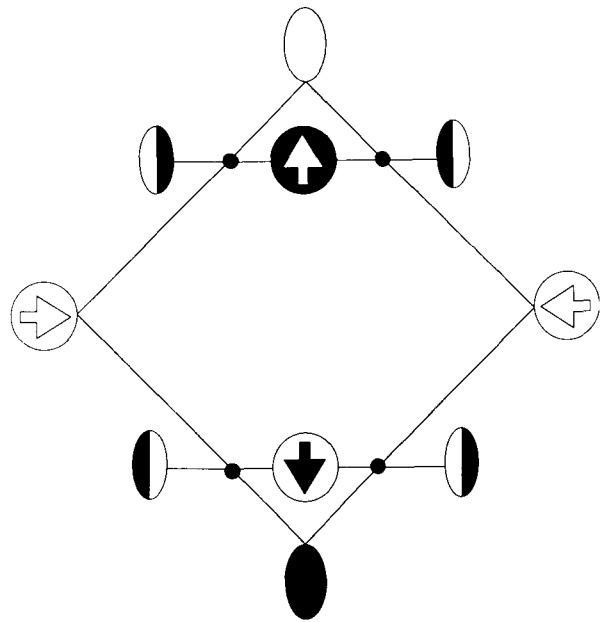


Figure 22. FM symmetry: Symmetric unfolding of pairs of opponent-orientation cells and opponent-direction cells whose outputs are insensitive to direction of contrast. The gating combinations from Figures 20 and 21 are combined to emphasize their underlying symmetry.

tion but sensitive to either the onset or the offset of an oriented contrast difference. These cells may be called complex/orientation/on cells and complex/orientation/off cells, respectively, as in Equations 13 and 14. They belong to the SOC filter.

The other opponent pair of outcomes (r_l, l_r) contains the MOC filter cell pairs, schematized in Figure 21, which are sensitive to opposite directions of motion. These cells may be called complex/direction/left cells and complex/direction/right cells, as in Equations 5 and 6. When both sets of pairs are combined into a single symmetric diagram, the result is as shown in Figure 22. Figure 22 suggests that parallel, but interdependent, streams of static form and motion form processing arise in visual cortex because the cortex develops by computing all possible sustained-transient output signals that are independent of direction of contrast and organized into opponent on-cells and off-cells. Experimental tests of this prediction will require a coordinated analysis of cell types and processing levels.

19. 90° Orientations: From V1 to V2

An important consequence of the abstract symmetry described in Figure 22 is the familiar fact from daily life that opposite static orientations are 90° apart, whereas opposite motion directions are 180° apart, as summarized in Section 6.

The 90° symmetry of opposite orientations may be explained by the way in which perpendicular *end cuts* are generated at the hypercomplex cells of the static BCS, as

analyzed in Grossberg and Mingolla (1985b). This perpendicularity property arises from the fact that the opponent feature of a complex/orientation/on cell is a complex/orientation/off cell. This mechanism is reviewed below for completeness. For those familiar with the end-cut concept, the 90° symmetry may tersely be summarized as follows: Suppose that a vertical line end excites a complex/vertical/on cell in Figure 1. Suppose that the end-stopped competition inhibits hypercomplex/vertical/on cells at positions beyond the line end. Hypercomplex/horizontal/on cells at these positions are thereby activated, and a horizontally oriented end cut is generated. In addition, hypercomplex/horizontal/off cells at these positions are inhibited by the opponent interaction. As a result, a net excitatory input is generated from the horizontally oriented hypercomplex cells to the horizontally oriented bipole cells of the CC loop at that position. These excitatory end-cut inputs cooperate across positions to generate, along the entire line end, a horizontal emergent segmentation that is perpendicular to the vertical line.

The more complete summary below of how end cuts are generated also provides an occasion for summarizing various data and predictions based on the Grossberg and Mingolla (1985b) prediction that end cuts exist. Readers familiar with this analysis may wish to skip to Section 21.

20. End Cuts: Cortical Simple Cells, Complex Cells, and Hypercomplex Cells as a Module for Hierarchical Resolution of Uncertainty

To effectively build up boundaries, the SOC filter must be able to determine the orientation of a boundary at every position. To accomplish this, the cells at the first stage of the SOC filter possess orientationally tuned simple-cell receptive fields. Such simple cells, or cell populations, are selectively responsive to orientations that activate a prescribed small region of the retina, and their orientations lie within a prescribed band of orientations with respect to the retina. A collection of such orientationally tuned cells is assumed to exist at every network position, such that each cell type is sensitive to a different band of oriented contrasts within its prescribed small region of the scene, as in the hypercolumn model, which was developed to explain the responses of simple cells in area V1 of the striate cortex (Hubel & Wiesel, 1977).

These oriented receptive fields are oriented *local-contrast* detectors, rather than edge detectors. This property enables them to fire in response to a wide variety of spatially nonuniform image contrasts, including edges, spatially nonuniform densities of unoriented textural elements, and spatially nonuniform densities of surface gradients. Thus, by sacrificing a certain amount of spatial resolution in order to detect oriented local contrasts, these masks achieve a general detection characteristic that can respond to edges, textures, and surfaces.

The fact that the receptive fields of the SOC filter are *oriented* greatly reduces the number of possible groupings into which their target cells can enter. On the other hand, in order to detect oriented local contrasts, the recep-

tive fields must be elongated along their preferred axis of symmetry. Then the cells can preferentially detect differences of average contrast across this axis of symmetry, yet remain silent in response to differences of average contrast that are perpendicular to the axis of symmetry. Such receptive-field elongation creates greater positional uncertainty about the exact locations within the receptive field of the image contrasts that fire the cell. This positional uncertainty becomes acute during the processing of image line ends and corners.

Oriented receptive fields cannot easily detect the ends of thin scenic lines or corners (Grossberg & Mingolla, 1985b). This property illustrates a basic uncertainty principle that says: Orientational "certainty" implies positional "uncertainty" at the ends of scenic lines whose widths are neither too small nor too large with respect to the dimensions of the oriented receptive field. If no BC signals are elicited at the ends of lines, however, then in the absence of further processing within the BCS, boundary contours will not be activated to prevent color and brightness signals from flowing out of line ends within the FCS during filling-in. Many percepts would hereby become badly degraded by featural flow. Thus, orientational certainty implies a type of positional uncertainty, which is unacceptable from the perspective of featural filling-in requirements.

Later processing stages within the BCS are needed to recover both the positional and orientational information that are lost in this way, so that the boundaries at line ends and corners can be completed before they are mapped into the FCS to control filling-in of surface brightness, color, and depth. Grossberg and Mingolla (1985b) have called the process that completes the boundary at a line end an *end cut*. End cuts actively reconstruct the line end at a processing stage higher than that of the oriented receptive field much as they do to form a circular Ehrenstein figure (Figure 23). To emphasize the paradoxical nature of this process, we say that *all line ends are illusory*. Interactions between simple cells, complex cells, and hypercomplex cells were predicted to generate these perpendicular end cuts.

The processing stages that are hypothesized to generate end cuts are diagramed in Figure 4. First, oriented simple-cell receptive fields of like position and orientation but opposite direction of contrast generate rectified output signals that summate at the next processing stage to activate complex cells whose receptive fields are sensitive to the same position and orientation as their own, but are insensitive to direction of contrast. In order to generate boundary detectors that can detect the broadest possible range of luminance or chromatic contrasts, in particular boundaries capable of bridging contrast reversals, these complex cells maintain their sensitivity to *amount* of oriented contrast, but not to the *direction* of this oriented contrast.

The rectified output from the complex cells activates a second filter, which is composed of two successive stages of spatially short-range competitive interaction whose net effect is to generate end cuts (Figure 4). First, a cell of prescribed orientation excites like-oriented cells

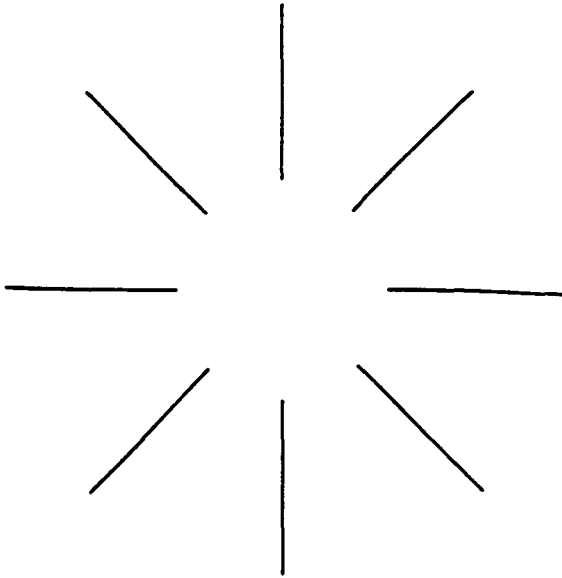


Figure 23. An Ehrenstein figure: A bright circular disk is perceived even though all white areas of the image are equally luminant. It is suggested that end cuts formed perpendicular to the line ends by the SOC filter cooperate within the CC loop to form a circular illusory boundary. This boundary separates two regions within the FCS whose filled-in activity levels differ. This difference is perceived as a difference in brightness.

corresponding to its location and inhibits like-oriented cells corresponding to nearby locations at the next processing stage. In other words, an on-center off-surround organization of like-oriented cell interactions exists around each perceptual location. This mechanism is analogous to the neurophysiological process of *end stopping*, whereby hypercomplex cell receptive fields are fashioned from interactions of complex cell output signals (Hubel & Wiesel, 1965; Orban, Kato, & Bishop, 1979). The outputs from this competitive mechanism interact with the second competitive mechanism. Here, hypercomplex cells that represent different orientations, notably perpendicular orientations, compete at the same perceptual location. This competition defines a push-pull opponent process. If a given orientation is excited, then its perpendicular orientation is inhibited. If a given orientation is inhibited, then its perpendicular orientation is excited via disinhibition.

The combined effect of these two competitive interactions generates end cuts as follows. The strong vertical activations along the edges of a scenic line inhibit the weak vertical activations near the line end. In turn, these inhibited vertical activations disinhibit horizontal activations near the line end, thereby generating a horizontal end cut that is perpendicular to its inducing vertical line end. Thus, the positional uncertainty generated by orientational certainty is eliminated at a subsequent processing level by the interaction of two spatially short-range competitive mechanisms which convert complex cells into two distinct populations of hypercomplex cells. This analysis of end cuts suggests that the simple cells, complex cells, and

hypercomplex cells function as a unitary network for achieving hierarchical resolution of uncertainty.

These model mechanisms have successfully predicted and helped to explain a variety of neural and perceptual data. For example, the Grossberg and Mingolla (1985b) complex cell model accords with the complex cell model that was independently derived by Spitzer and Hochstein (1985) from their neurophysiological experiments on cats. The full BCS model also clarifies, in a manner that goes beyond the Spitzer-Hochstein model, why many complex cells receive inputs from several different classes of color opponent cells in the lateral geniculate nucleus (DeValois et al., 1982). This convergence of opponent cells generates a chromatically broad-band boundary detector. The BCS model hereby clarifies how inputs to complex cells from simple cells with chromatically opponent receptive fields may be attenuated in response to isoluminant stimuli, without denying that BCS boundaries are sensitive to color inputs. This model of complex cells also needs to be extended to explain basic facts of binocular fusion and rivalry (Grossberg, 1987c; Grossberg & Marshall, 1989).

An important prediction of the theory anticipated the report by von der Heydt, Peterhans, and Baumgartner (1984) that cells in prestriate visual cortex respond to perpendicular line ends, whereas cells in striate visual cortex do not. These cell properties also helped to explain why color is sometimes perceived to spread across a scene, as in the phenomenon of neon color spreading (Grossberg, 1987b; Grossberg & Mingolla, 1985a; Redies & Spillmann, 1981), by showing how some of the boundaries that would otherwise have been generated by image contrasts may be inhibited by the competitive mechanisms underlying end cuts in response to certain scenes, thereby allowing colors to spread beyond these image contrasts. The end-cut process also exhibits properties of hyperacuity that have been used (Grossberg, 1987b) to explain subsequent psychophysical data about spatial localization and hyperacuity (Badcock & Westheimer, 1985a, 1985b; Watt & Campbell, 1985). A similar double-filter model has been used to analyze data about texture segregation (Sutter, Beck, & Graham, 1989) in a way that supports the texture analyses of Grossberg and Mingolla (1985b). The latter texture segregation analyses also utilized the cooperative-competitive feedback interactions of the CC loop (Figure 1) to generate emergent boundary segmentations, such as the Kanizsa square generated in response to the four Pac Man figures in Figure 9.

21. 180° Opponent Directions from V1 to MT

The fact that opponent directions differ by 180°, rather than 90°, follows from the fact, diagrammed in Figure 21, that the opposite feature of a complex/direction cell is another complex/direction cell whose direction preference differs from it by 180°. When the latter property is organized into a network topography, one finds the type of direction hypercolumns that were described in MT by Albright et al. (1984). A schematic explanation of how direction hypercolumns in MT may be generated from the

orientation hypercolumns of V1 is shown in Figure 24. This explanation suggests that the pathways from V1 to MT combine signals from sustained cells and transient cells, as in Figure 21, in a different way than the pathways from V1 to V2, as in Figure 20.

22. Opponent Rebounds: Rapid Reset of Resonating Segmentations

A final refinement of the SOC filter and MOC filter designs assumes that the opponent cell pairs shown in Figures 20 and 21 are capable of *antagonistic rebound*; that is, offset of one cell in the pair after its sustained activation can trigger an antagonistic rebound that transiently activates the opponent cell in the pair. A minimal neural model of such an opponent rebound, illustrated in Figure 25, is called a *gated dipole* (Grossberg, 1972, 1982, 1988). Such an antagonistic rebound, when appropriately embedded in an SOC filter or an MOC filter, can reset a resonating segmentation in response to rapid

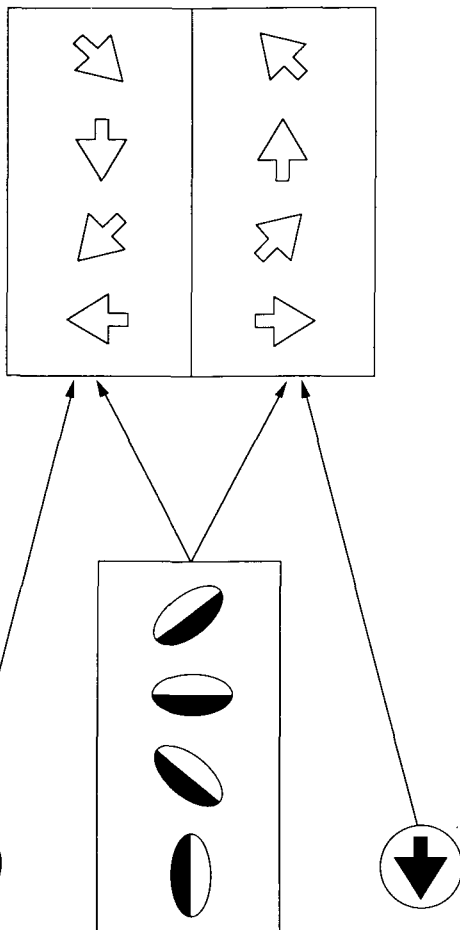


Figure 24. Orientation and direction hypercolumns: A single hypercolumn of orientation cells (say in V1) can give rise to a double hypercolumn of opponent-direction cells (say in MT) through gating with opponent pairs of transient cells.

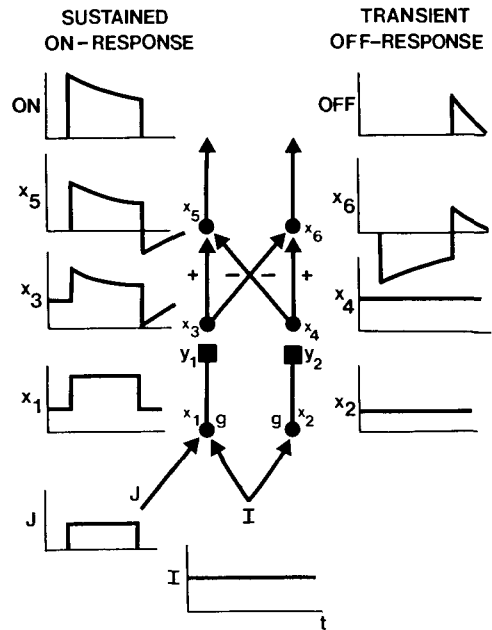


Figure 25. Example of a feedforward gated dipole: A sustained habituating on-response (top left) and a transient off-rebound (top right) are elicited in response to onset and offset, respectively, of a phasic input J (bottom left), when tonic arousal, I (bottom center), and opponent processing (diagonal pathways) supplement the slow gating actions (square synapses).

changes in the stimulus, as discussed in Section 8. For example, suppose that horizontally oriented hypercomplex cells in the SOC filter are cooperating with horizontally oriented bipole cells to generate a horizontal boundary segmentation in the CC loop (Figure 1) when the input pattern is suddenly shut off. In the absence of opponent processing, the positive feedback signals between the active hypercomplex on-cells and bipole cells could maintain the boundary segmentation for a long time after input offset, thereby causing massive smearing of the visual percept in response to rapidly changing scenes.

Suppose, however, that due to opponent processing by gated dipoles, offset of the horizontal complex on-cells can trigger an antagonistic rebound that activates the horizontal hypercomplex off-cells. The horizontal hypercomplex off-cells would then generate inhibitory signals to the horizontal bipole cells, as in Figure 1. These inhibitory signals would actively shut off the resonating segmentation, thereby preventing too much smearing from occurring. Such inhibitory signals from hypercomplex cells to bipole cells are predicted to be one of the inhibitory processes that control the amount of smearing caused by a moving image in the experiments of Hogben and DiLollo (1985).

This analysis of how antagonistic rebounds can reset a resonating segmentation leads to the prediction that gated dipoles occur at the complex cell level or the hyper-

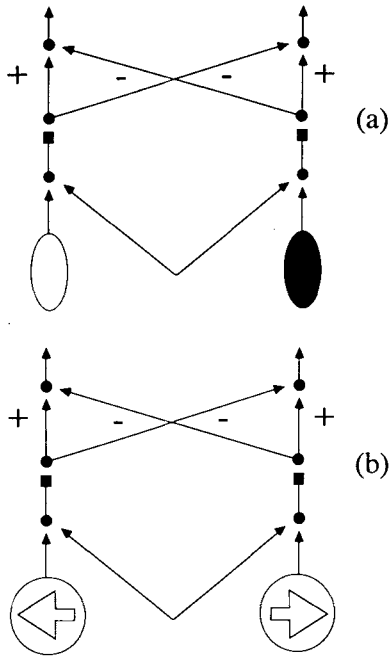


Figure 26. Opponent rebounds: When orientationally tuned complex cells in the SOC filter are organized into gated dipole opponent circuits, as in (a), offset of a complex on-cell can transiently activate like-oriented complex off-cells, as well as perpendicular hyper-complex on-cells at the second competitive stage (see text). Offset of directionally tuned complex cells within the MOC filter, as in (b), can transiently activate complex cells tuned to the opposite direction.

complex cell level (Figure 26) in the static BCS and motion BCS.

23. MacKay Afterimages, the Waterfall Effect, and Long-Range MAE

The previous sections argued that some positive after-effects might be partly due to a lingering resonance in a CC loop, and that some negative aftereffects might be partly due to an antagonistic rebound that resets such a resonance. Within the static BCS, negative aftereffects tend to activate perpendicular segmentations via the same 90° symmetry of the SOC filter that generates perpendicular end cuts (Section 20). Due to this symmetry, sustained inspection of a radial image can induce a circular aftereffect if a blank field is subsequently inspected (MacKay, 1957). In a similar fashion, it follows from the 180° symmetry of the MOC filter, diagrammed in Figure 24, that sustained inspection of a waterfall can induce an upward-moving motion aftereffect (MAE) if a blank field is subsequently attended (Sekuler, 1975).

The assumption that a level of gated dipoles occurs at, or subsequent to, Level 5 of the MOC filter also provides an explanation of how a long-range MAE can occur between the locations of two flashes that previously generated apparent motion between themselves (von Grünau, 1986). As discussed in Section 15, a wave of apparent motion is synthesized at Level 5 as a result of interactions of the flashes through the long-range Gaussian filter

described in Equations 5–11. The gated dipoles at, or subsequent to, Level 5 will habituate to the wave of apparent motion much as they would in response to a “real” motion signal expressed at Level 5.

24. Concluding Remarks: The Inadequacy of Independent Visual Modules

FACADE theory clarifies that although specialization of function surely exists during visual perception, it is not the type of specialization that may adequately be described by independent neural modules for the processing of edges, textures, shading, depth, motion, and color information. In particular, FACADE theory provides an explanation of many data that do not support the modular approach advocated by Marr (1982).

A basic conceptual problem faced by a modular approach may be described as follows. Suppose that specialized modules capable of processing edges, or textures, or shading, and so forth, are available. Typically, each of these modules is described using different mathematical rules that are not easily combined into a unified theory. Correspondingly, the modules do not respond well to visual data other than the type of data they were designed to process. In order to function well, either the visual world that such a module is allowed to process must be restricted, whence the module could not be used to process realistic scenes, or a smart preprocessor is needed to sort scenes into parts according to the type of data that each module can process well, and to expose a module only to that part of a scene for which it was designed. Such a smart preprocessor would, however, embody a vision model that was more powerful than the modules themselves, which would render the modules obsolete. In either case, modular algorithms do not provide a viable approach to the study of real-world vision.

The task of such a smart preprocessor is, in any case, more difficult than that of sorting scenes into parts which contain only one type of visual information. This is because each part of a visual scene often contains locally ambiguous information about edges, textures, shading, depth, motion, and color, all overlaid together. Humans are capable of using these multiple types of visual information cooperatively to generate an unambiguous 3-D representation of Form-And-Color-And-Depth; hence the term FACADE representation. The hyphens in “Form-And-Color-And-Depth” emphasize the well-known fact that changes in perceived color can cause changes in perceived depth and form, changes in perceived depth can cause changes in perceived brightness and form, and so on. Every stage of visual processing *multiplexes* together several key properties of the scenic representation. It is a central task of biological vision theories to understand how the organization of visual information processing regulates which properties are multiplexed together at each processing stage, and how the stages interact to generate these properties.

FACADE theory became possible through the discovery of several new uncertainty principles—that is, principles which show what combinations of visual properties

cannot, in principle, be computed at a single processing stage (Grossberg, 1987b; Grossberg & Mingolla, 1985b). The theory has by now described how to design parallel and hierarchical interactions that can resolve these uncertainties using several processing stages. The hierarchical computations of end cuts to overcome orientational uncertainty is illustrative (Section 20). These interactions occur within and between two subsystems, the BCS and the FCS, whose computations are computationally complementary. In addition, principles of symmetry seem to govern the organization of these subsystems, as the designs of SOC and MOC filters illustrate. Resonance principles are also operative, as in the design of the CC loop.

Issues concerning uncertainty principles, complementarity, symmetry, and resonance lie at the foundations of quantum mechanics and other physical theories. Mammalian vision systems are also quantum systems in the sense that they can generate visual percepts in response to just a few light quanta. How the types of uncertainty, complementarity, symmetry, and resonance that are resolved by biological vision systems for purposes of macroscopic perception may be related to concepts of uncertainty, complementarity, symmetry, and resonance in quantum mechanics or other physical theories is a theme of considerable importance for future research.

The themes of uncertainty, complementarity, symmetry, and resonance show the inadequacy of the modular and rule-based approaches from a deeper information theoretic perspective. Although the BCS, FCS, and their individual processing stages are computationally specialized, their interactions overcome computational uncertainties and complementary deficiencies to generate useful visual representations, rather than properties that may be computed by independent processing modules. Context-sensitive interactions also determine which combinations of positions, orientations, disparities, spatial scales, and the like will be cooperatively linked through resonance. Likewise, the symmetry principle that integrates static form and motion form properties cannot be stated as a property of independent modules for form or motion perception, because the static BCS and the motion BCS each process aspects of both form and motion, and the design of each of these networks can best be understood as parts of a single larger system, as in Figure 22.

Such an interactive theory precludes the sharp separation between formal algorithm and mechanistic realization that Marr (1982) proposed. How computational uncertainties can be overcome, how complementary processing properties can be interactively synthesized, and how particular combinations of multiplexed properties may resonate or be symmetrically organized are all properties of particular classes of mechanistic realizations. Many workers in the field of neural networks summarize this state of affairs by saying that "the architecture is the algorithm." Future tasks in theoretically understanding biological vision promise to require that we replace algorithmic rules

and independent modules by architectural designs whose emergent properties constitute intelligence as we know it.

REFERENCES

- ALBRIGHT, T. D. (1984). Direction and orientation selectivity of neurons in visual area MT of the macaque. *Journal of Neurophysiology*, **52**, 1106-1130.
- ALBRIGHT, T. D., DESIMONE, R., & GROSS, C. G. (1984). Columnar organization of directionally sensitive cells in visual area MT of the macaque. *Journal of Neurophysiology*, **51**, 16-31.
- ANDERSON, S. J., & BURR, D. C. (1987). Receptive field size of human motion detection units. *Vision Research*, **27**, 621-635.
- ANSTIS, S., & RAMACHANDRAN, V. S. (1987). Visual inertia in apparent motion. *Vision Research*, **27**, 755-764.
- BADCOCK, D. R., & WESTHEIMER, G. (1985a). Spatial location and hyperacuity: The centre/surround localization contribution function has two substrates. *Vision Research*, **25**, 1259-1267.
- BADCOCK, D. R., & WESTHEIMER, G. (1985b). Spatial location and hyperacuity: Flank position within the centre and surround zones. *Spatial Vision*, **1**, 3-11.
- BARTLEY, S. H. (1941). *Vision, a study of its basis*. New York: Van Nostrand.
- BRADDICK, O. J. (1974). A short range process in apparent motion. *Vision Research*, **14**, 519-527.
- BURR, D. C., ROSS, J., & MORRONE, M. C. (1986). Smooth and sampled motion. *Vision Research*, **26**, 643-652.
- CARPENTER, G. A., & GROSSBERG, S. (1987a). ART 2: Stable self-organization of pattern recognition for analog input patterns. *Applied Optics*, **26**, 4919-4930.
- CARPENTER, G. A., & GROSSBERG, S. (1987b). A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics, & Image Processing*, **37**, 54-115.
- 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. (1990). ART 3: Hierarchical search using chemical transmitters in self-organizing pattern recognition architectures. *Neural Networks*, **3**, 129-152.
- COHEN, M. A., & GROSSBERG, S. (1984). Neural dynamics of brightness perception: Features, boundaries, diffusion, and resonance. *Perception & Psychophysics*, **36**, 428-456.
- DEVALOIS, R. L., ALBRECHT, D. G., & THORELL, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, **22**, 545-559.
- DEYOE, E. A., & VAN ESSEN, D. C. (1988). Concurrent processing streams in monkey visual cortex. *Trends in Neuroscience*, **11**, 219-226.
- DUNCAN, J., & HUMPHREYS, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, **96**, 433-458.
- ECKHORN, R., BAUER, R., JORDAN, W., BROSCHE, M., KRUSE, W., MUNK, M., & REITBOECK, H. J. (1988). Coherent oscillations: A mechanism of feature linking in the visual cortex? *Biological Cybernetics*, **60**, 121-130.
- ERIKSEN, C. W., & MURPHY, T. D. (1987). Movement of attentional focus across the visual field: A critical look at the evidence. *Perception & Psychophysics*, **42**, 29-305.
- GRAY, C. M., KONIG, P., ENGEL, A. K., & SINGER, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, **338**, 334-337.
- GROSSBERG, S. (1972). A neural theory of punishment and avoidance: II. Quantitative theory. *Mathematical Biosciences*, **15**, 253-285.
- GROSSBERG, S. (1973). Contour enhancement, short-term memory, and constancies in reverberating neural networks. *Studies in Applied Mathematics*, **52**, 217-257.
- GROSSBERG, S. (1976). Adaptive pattern classification and universal recoding: II. Feedback, expectation, olfaction, and illusions. *Biological Cybernetics*, **23**, 187-202.

- GROSSBERG, S. (1978). A theory of visual coding, memory, and development. In E. Leeuwenberg & H. Buffart (Eds.), *Formal theories of visual perception* (pp. 7-26). New York: Wiley.
- GROSSBERG, S. (1980). How does a brain build a cognitive code? *Psychological Review*, **87**, 1-51.
- GROSSBERG, S. (1982). *Studies of mind and brain: Neural principles of learning, perception, development, cognition, and motor control*. Boston: Reidel Press.
- GROSSBERG, S. (Ed.) (1987a). *The adaptive brain: II. Vision, speech, language, and motor control*. Amsterdam: Elsevier/North-Holland.
- GROSSBERG, S. (1987b). Cortical dynamics of three-dimensional form, color, and brightness perception: I. Monocular theory. *Perception & Psychophysics*, **41**, 87-116.
- GROSSBERG, S. (1987c). Cortical dynamics of three-dimensional form, color, and brightness perception: II. Binocular theory. *Perception & Psychophysics*, **41**, 117-158.
- GROSSBERG, S. (Ed.) (1988). *Neural networks and natural intelligence*. Cambridge, MA: MIT Press.
- GROSSBERG, S. (1990). *3-D vision and figure-ground separation by visual cortex*. Manuscript 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 & Psychophysics*, **38**, 141-171.
- GROSSBERG, S., & MINGOLLA, E. (1986). Computer simulation of neural networks for perceptual psychology. *Behavior Research Methods, Instruments, & Computers*, **18**, 601-607.
- GROSSBERG, S., & MINGOLLA, E. (1987). Neural dynamics of surface perception: Boundary webs, illuminants, and shape-from-shading. *Computer Vision, Graphics, & Image Processing*, **37**, 116-165.
- GROSSBERG, S., & MINGOLLA, E. (1990a). Neural dynamics of motion segmentation. In *Proceedings of Graphics Interface/Vision Interface '90*, Halifax, Nova Scotia, May 14-18 (pp. 112-119). Toronto: Canadian Information Processing Society.
- GROSSBERG, S., & MINGOLLA, E. (1990b). Neural dynamics of motion segmentation: Direction fields, apertures, and resonant grouping. In M. Caudill (Ed.), *Proceedings of the International Joint Conference on Neural Networks*, January, **1**, 11-14. Hillsdale, NJ: Erlbaum.
- GROSSBERG, S., & MINGOLLA, E. (1990c). *Neural dynamics of motion segmentation: Direction fields, apertures, and resonant grouping*. Manuscript submitted for publication.
- GROSSBERG, S., MINGOLLA, E., & TODOROVIĆ, D. (1989). A neural network architecture for preattentive vision. *IEEE Transactions on Biomedical Engineering*, **36**, 65-84.
- GROSSBERG, S., & RUDD, M. E. (1989a). A neural architecture for visual motion perception: Group and element apparent motion. In M. Caudill (Ed.), *Proceedings of the International Joint Conference on Neural Networks*, June, **1**, 195-199. Piscataway, NJ: IEEE.
- GROSSBERG, S., & RUDD, M. E. (1989b). A neural architecture for visual motion perception: Group and element apparent motion. *Neural Networks*, **2**, 421-450.
- GROSSBERG, S., & RUDD, M. E. (1989c). Neural dynamics of visual motion perception: Group and element apparent motion. *Investigative Ophthalmology Supplement*, **30**, 73.
- GROSSBERG, S., & RUDD, M. E. (1990a). *Cortical dynamics of visual motion perception: Short-range and long-range motion*. Manuscript submitted for publication.
- GROSSBERG, S., & RUDD, M. E. (1990b). Cortical dynamics of visual motion perception: Short- and long-range motion. *Investigative Ophthalmology Supplement*, **31**, 529.
- GROSSBERG, S., & SOMERS, D. (1990). *Synchronized oscillations during cooperative feature linking in a cortical model of visual perception*. Manuscript submitted for publication.
- GROSSBERG, S., & TODOROVIĆ, D. (1988). Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena. *Perception & Psychophysics*, **43**, 241-277.
- HEGGLUND, P. (1981). Receptive field organization of simple cells in cat striate cortex. *Experimental Brain Research*, **42**, 89-98.
- HOGGEN, J. H., & DiLollo, V. (1985). Suppression of visual persistence in apparent motion. *Perception & Psychophysics*, **38**, 450-460.
- HUBEL, D. H., & WIESEL, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, **160**, 106-154.
- HUBEL, D. H., & WIESEL, T. N. (1965). Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *Journal of Neurophysiology*, **28**, 229-289.
- HUBEL, D. H., & WIESEL, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, **195**, 215-243.
- HUBEL, D. H., & WIESEL, T. N. (1977). Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London (B)*, **198**, 1-59.
- KOLERS, P. A. (1972). *Aspects of motion perception*. Oxford: Pergamon Press.
- LABERGE, D., & BROWN, V. (1989). Theory of attentional operations in shape identification. *Psychological Review*, **96**, 101-124.
- MACKEY, D. M. (1957). Moving visual images produced by regular stationary patterns. *Nature*, **180**, 849-850.
- MARR, D. (1982). *Vision*. San Francisco: Freeman.
- MARSHALL, J. (1990). Self-organizing neural networks for perception of visual motion. *Neural Networks*, **3**, 45-74.
- MAUNSELL, J. H. R., & VAN ESSEN, D. C. (1983). Response properties of single units in middle temporal visual area of the macaque. *Journal of Neurophysiology*, **49**, 1127-1147.
- NAKAYAMA, K., & SILVERMAN, G. H. (1984). Temporal and spatial characteristics of the upper displacement limit for motion in random dots. *Vision Research*, **24**, 293-299.
- NAKAYAMA, K., & SILVERMAN, G. H. (1985). Detection and discrimination of sinusoidal grating displacements. *Journal of the Optical Society of America*, **2**, 267-273.
- NAKAYAMA, K., & SILVERMAN, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature*, **320**, 264-265.
- NEWSOME, W. T., GIZZI, M. S., & MOVSHON, J. A. (1983). Spatial and temporal properties of neurons in macaque MT. *Investigative Ophthalmology & Visual Science*, **24**, 106.
- ORBAN, G. A., KATO, H., & BISHOP, P. O. (1979). Dimensions and properties of end-zone inhibitory areas in receptive fields of hypercomplex cells in cat striate cortex. *Journal of Neurophysiology*, **42**, 833-849.
- PASHLER, H. (1987). Detecting conjunctions of color and form: Re-assessing the serial search hypothesis. *Perception & Psychophysics*, **41**, 191-201.
- PETERHANS, E., & VON DER HEYDT, R. (1989). Mechanisms of contour perception in monkey visual cortex, II. Contours bridging gaps. *Journal of Neuroscience*, **9**, 1749-1763.
- POGGIO, G. F., MOTTER, B. C., SQUATRITO, S., & TROTTER, Y. (1985). Responses of neurons in visual cortex (V1 and V2) of the alert macaque to dynamic random-dot stereograms. *Vision Research*, **25**, 397-406.
- RAMACHANDRAN, V. S. (1985). Apparent motion of subjective surfaces. *Perception*, **14**, 127-134.
- RAMACHANDRAN, V. S., RAO, V. M., & VIDYASAGAR, T. R. (1973). Apparent motion with subjective contours. *Vision Research*, **13**, 1399-1401.
- REDIES, C., & SPILLMANN, L. (1981). The neon color effect in the Ehrenstein illusion. *Perception*, **10**, 667-681.
- REMINGTON, R., & PIERCE, L. (1984). Moving attention: Evidence for time-invariant shifts of visual selective attention. *Perception & Psychophysics*, **35**, 393-399.
- SEKULER, R. (1975). Visual motion perception. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception: Vol. V. Seeing* (pp. 387-430). New York: Academic Press.
- SPITZER, H., & HOCHSTEIN, S. (1985). A complex-cell receptive field model. *Journal of Neurophysiology*, **53**, 1266-1286.
- SUTTER, A., BECK, J., & GRAHAM, N. (1989). Contrast and spatial variables in texture segregation: Testing a simple spatial-frequency channels model. *Perception & Psychophysics*, **46**, 312-332.

- TANAKA, M., LEE, B. B., & CREUTZFELDT, O. D. (1983). Spectral tuning and contour representation in area 17 of the awake monkey. In J. D. Mollon & L. T. Sharpe (Eds.), *Colour vision* (pp. 269-276). New York: Academic Press.
- TERNUS, J. (1926). Experimentelle Untersuchungen über phänomenale Identität. *Psychologische Forschung*, **7**, 81-136. [Abstracted and translated in W. D. Ellis (Ed.), *A sourcebook of Gestalt psychology*. New York: Humanities Press, 1950.]
- THORELL, L. G., DEVALOIS, R. L., & ALBRECHT, D. G. (1984). Spatial mapping of monkey V1 cells with pure color and luminance stimuli. *Vision Research*, **24**, 751-769.
- TREISMAN, A., & GELADE, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, **12**, 97-136.
- TREISMAN, A., & GORMICAN, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, **95**, 15-48.
- TREISMAN, A., & SOUTHER, J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General*, **114**, 285-310.
- VAN ALLEN, E. J., & KOLODZY, P. J. (1987). Application of a boundary contour neural network to illusions and infrared sensor imagery. In M. Caudill & C. Butler (Eds.), *Proceedings of the IEEE First International Conference on Neural Networks, IV*, 193-197. Piscataway, NJ: IEEE Press.
- VON DER HEYDT, R., PETERHANS, E., & BAUMGARTNER, G. (1984). Illusory contours and cortical neuron responses. *Science*, **224**, 1260-1262.
- VON DER HEYDT, R., HÄNNY, P., & DÜRSTELER, M. R. (1981). The role of orientation disparity in stereoscopic perception and the development of binocular correspondence. In E. Grastyán & P. Molnár (Eds.), *Advances in physiological sciences: Vol. 16. Sensory functions*. Elmsford, NY: Pergamon Press.
- VON GRÜNAU, M. W. (1986). A motion aftereffect for long-range stroboscopic apparent motion. *Perception & Psychophysics*, **40**, 31-38.
- WATT, R. J., & CAMPBELL, F. W. (1985). Vernier acuity: Interactions between length effects and gaps when orientation cues are eliminated. *Spatial Vision*, **1**, 31-38.
- ZEKI, S. M. (1974a). Cells responding to changing image size and disparity in the cortex of the rhesus monkey. *Journal of Physiology*, **242**, 827-841.
- ZEKI, S. M. (1974b). Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *Journal of Physiology*, **236**, 549-573.

(Manuscript received January 12, 1990;
revision accepted for publication August 3, 1990.)