Perception & Psychophysics
1993, 53 (3), 243-278

Neural dynamics of motion perception: Direction fields, apertures, and resonant grouping

STEPHEN GROSSBERG and ENNIO MINGOLLA
Boston University, Boston, Massachusetts

A neural network model of global motion segmentation by visual cortex is described. Called the motion boundary contour system (BCS), the model clarifies how ambiguous local movements on a complex moving shape are actively reorganized into a coherent global motion signal. Unlike many previous researchers, we analyze how a coherent motion signal is imparted to all regions of a moving figure, not only to regions at which unambiguous motion signals exist. The model hereby suggests a solution to the global aperture problem. The motion BCS describes how preprocessing of motion signals by a motion oriented contrast (MOC) filter is joined to long-range cooperative grouping mechanisms in a motion cooperative-competitive (MOCC) loop to control phenomena such as motion capture. The motion BCS is computed in parallel with the static BCS of Grossberg and Mingolla (1985a, 1985b, 1987). Homologous properties of the motion BCS and the static BCS, specialized to process motion directions and static orientations, respectively, support a unified explanation of many data about static form perception and motion form perception that have heretofore been unexplained or treated separately. Predictions about microscopic computational differences of the parallel cortical streams V1→MT and V1→V2→MT are made—notably, the magnocellular thick stripe and parvocellular interstripe streams. It is shown how the motion BCS can compute motion directions that may be synthesized from multiple orientations with opposite directions of contrast. Interactions of model simple cells, complex cells, hypercomplex cells, and bipole cells are described, with special emphasis given to new functional roles in direction disambiguation for endstopping at multiple processing stages and to the dynamic interplay of spatially short-range and long-range interactions.

1. Introduction: Neural Architectures for Segmenting Static Forms and Moving Forms

In our everyday perceptions, the unity and persistent identity of objects undergoing motion is so immediate that we may all too readily take for granted the subtle problems posed in discriminating such unity within the scintillating mosaic of visual stimulation. In the familiar example of trying to detect a leopard moving through a forest canopy, the change in the optic array over time is a jumble of contrast changes whose local components point in a variety of directions due to the vagaries of the motions of rustling leaves and of the occlusions and disocclusions of markings on the leopard's body. For a leopard to be spotted (by a visual system) in such contexts, a key problem must be solved: How can the locally ambiguous local motion signals corresponding to the many parts of the leopard's body be *reorganized* into a coherent object whole with a unitary motion?

Our investigation of the global reorganization of local motion signals is conducted within a theoretical framework that has already analyzed analogous problems within the domain of static form perception. This latter theory has been called FACADE theory, because its visual representations are predicted to multiplex together properties of Form-And-Color-And-DEpth in a model analog of the 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, 1987a, 1987b, 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, 1987a, 1987b, 1987c; Grossberg & Mingolla, 1985a; Grossberg & Todorović, 1988).

This BCS model is now called the *static* BCS in order to distinguish it from the *motion* BCS that is analyzed in

S. Grossberg was supported in part by the Air Force Office of Scientific Research (AFOSR 90-0175), DARPA (AFOSR 90-0083), and the Office of Naval Research (ONR N00014-91-J-4100). E. Mingolla was supported in part by the Air Force Office of Scientific Research (AFOSR 90-0175). The authors wish to thank Cynthia E. Bradford for her expert assistance in the preparation of the manuscript. Correspondence may be addressed to the authors at Center for Adaptive Systems, Department of Cognitive and Neural Systems, Boston University, 111 Cummington St., Boston, MA 02215.

the present article. The static BCS models properties of the visual processing stream that originates at parvocellular cells of the lateral geniculate nuclei (LGN) and passes through interblob (area V1) and interstripe (area V2) cortical networks on its way to cortical area V4 (Figure 1). The FCS models the processing stream that originates at LGN parvocellular cells and passes through blob (area V1) and thin stripe (area V2) cortical networks on its way to cortical area V4.

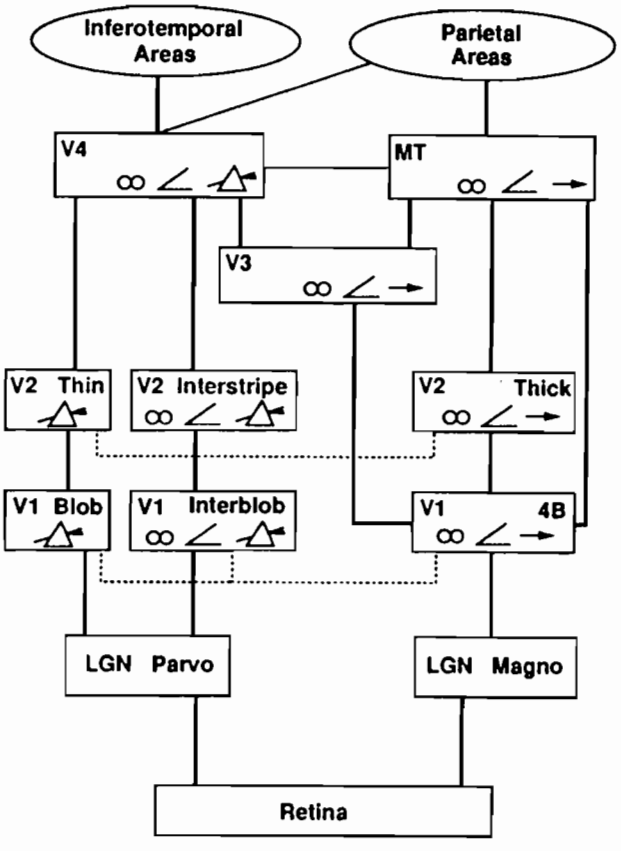


Figure 1. Schematic diagram of anatomical connections and neuronal selectivities of early visual areas in the macaque monkey. LGN = lateral geniculate nucleus (parvocellular and magnocellular divisions). Divisions of V1 and V2: blob = cytochrome oxidase blob regions; interblob = cytochrome oxidase-poor regions surrounding the blobs; 4B = lamina 4B; thin = thin (narrow) cytochrome oxidase strips; interstripe = cytochrome oxidase-poor regions between the thin and thick stripes; thick = thick (wide) cytochrome oxidase strips; V3 = visual area 3; V4 = visual area(s) 4; MT = middle temporal area. Areas V2, V3, V4, and MT have connections to other areas not explicitly represented here. Area V3 may also receive projections from V2 interstripes or thin stripes. Heavy lines indicate robust primary connections, and thin lines indicate weaker, more variable connections. Dotted lines represent observed connections that require additional verification. Icons: rainbow = tuned and/or opponent wavelength selectivity (incidence at least 40%); angle symbol = orientation selectivity (incidence at least 20%); spectacles = binocular disparity selectivity and/or strong binocular interactions (V2) (incidence at least 20%); pointing arrow = direction of motion selectivity (incidence at least 20%). (From "Concurrent processing streams in monkey visual cortex" by E. A. DeYoe and D. C. van Essen, 1988, *Trends in Neuroscience*, 11, 219-226. Copyright 1988 by Elsevier Trends Journals. Adapted by permission.)

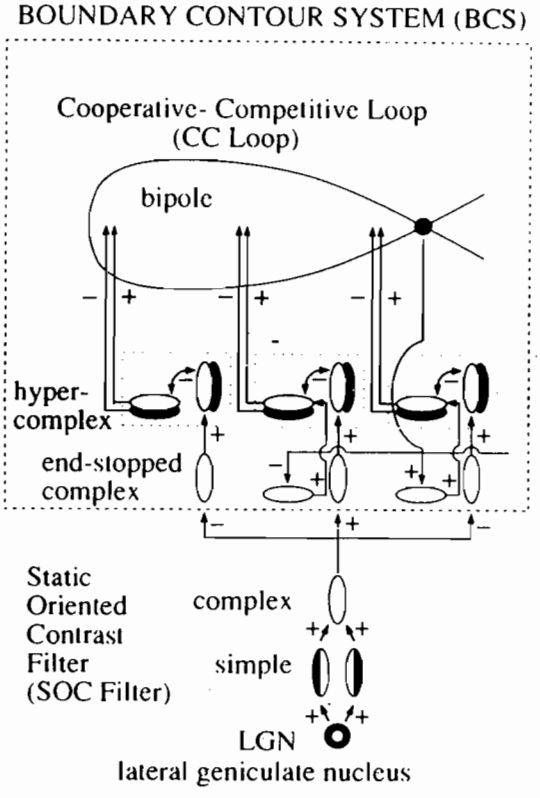


Figure 2. The static boundary contour system consists of two main parts, the static oriented contrast (SOC) filter and the cooperative-competitive (CC) loop. The SOC filter determines the locally preferred orientations of contrast differences in the input, while the CC loop groups, selects, and sharpens contours, both "real" and "illusory." (Reprinted, with permission, from Grossberg, 1991.)

The static BCS provides a new computational rationale as well as a model of the neural circuits governing classical cortical cell types, such as simple cells, complex cells, and hypercomplex cells in the interblob and interstripe regions of cortical areas V1 and V2. The theory also predicted a new cell type, the bipole cell (Cohen & Grossberg, 1984; Grossberg, 1984; Grossberg & Mingolla, 1985a), whose properties have been supported by neurophysiological experiments (Peterhans & von der Heydt, 1989; von der Heydt, Peterhans, & Baumgartner, 1984). The static BCS model consists of several parallel copies, such that each copy is activated by a different range of receptive-field sizes. Each static BCS copy is further subdivided into two hierarchically organized systems (Figure 2): 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 loop*, or *CC loop*, for generating coherent emergent boundary segmentations of the filtered signals. The SOC filter contains simple, complex, and hypercomplex cells. The CC loop contains hypercomplex and bipole cells. Interactions between the static BCS and FCS model how a FACADE representation emerges in cortical area V4. Many heretofore unexplained phenomena about preattentive vision and its attentive modulation have been clari-

fied by FACADE theory, and its concepts and mechanisms have been tested by a number of laboratories (Beck, Graham, & Sutter, 1991; Beck, Rosenfeld, & Ivry, 1990; Brown & Weisstein, 1988; Buckley, Frisby, & Mayhew, 1989; Dresp, Lorenceau, & Bonnet, 1990; Eskew, 1989; Eskew, Stromeyer, Picotte, & Kronauer, 1991; Graham, Beck, & Sutter, 1992; Humphreys, Quinlan, & Riddoch, 1989; Kellman & Shipley, 1991; Meyer & Dougherty, 1987; Mikaelian, Linton, & Phillips, 1990; Nakayama, Shimojo, & Ramachandran, 1990; Paradiso & Nakayama, 1991; Peterhans & von der Heydt, 1989; Prinzmetal, 1990; Prinzmetal & Boaz, 1989; Ramachandran, 1992; Shipley & Kellman, 1992; Sutter, Beck, & Graham, 1989; Takeichi, Shimojo, & Watanabe, 1992; Takeichi, Watanabe, & Shimojo, in press; Todd & Akerstrom, 1987; Watanabe & Cavanagh, 1991; Watanabe & Sato, 1989; Watanabe & Takeichi, 1990).

After the development of the static BCS reached a certain level of clarity, it focused attention on the question: How do the cortical systems that process static forms and moving forms differ computationally? Some regions of 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). However, even the earliest stages of visual cortical processing, such as simple cells in V1, require stimuli that change through time for their maximal activation and are direction-sensitive (DeValois, Albrecht, & Thorell, 1982; Heggelund, 1981; Hubel & Wiesel, 1962, 1968, 1977; Tanaka, Lee, & Creutzfeldt, 1983). Why could not V1 be used as the gateway to a single cortical stream that processes all aspects of form and motion? What computational properties are achieved by the magnocellular V1→MT cortical stream that are not achieved by the parallel parvocellular V1→V2→V4 cortical stream?

Analysis of the SOC filter design revealed that one of its basic properties made it unsuitable for motion processing. In particular, SOC filter output signals cannot adequately discriminate the direction of motion of a moving figure. In particular, SOC filter complex cells can respond equally well to motion in opposite directions, as is also true of many V1 complex cells in vivo (Foster, Gaska, Nagler, & Pollen, 1985). Further analysis of how this happens and how it can be overcome led to a new theory of biological motion perception that was outlined in Grossberg (1987b) and quantitatively specified and analyzed in Grossberg (1990, 1991) and Grossberg and Rudd (1989a, 1989b, 1992). The results of this analysis suggested that the motion form perception system shares many design features with the static form perception system, but that it has incorporated the minimal differences needed to achieve sensitivity to both local and global properties of direction of motion. In fact, Grossberg (1990, 1991) has suggested that the two systems are parallel subsystems of a larger, symmetric system design, called *form-motion* (FM) *symmetry*, that is predicted to govern the development of visual cortex. This symmetric organization provides an explanation, for example, of why the opposite orientation of vertical is horizontal—a difference of 90°—whereas the opposite direction of up is down—a difference of 180°.

Correspondingly, this new theory of biological motion perception 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, as in the static BCS. Also as in the static BCS, each motion BCS copy is further subdivided into hierarchically organized subsystems: a *motion oriented contrast filter*, or *MOC filter*, for preprocessing moving images, and a *motion cooperative-competitive feedback loop*, or *MOCC loop*, for generating coherent emergent boundary segmentations of the filtered signals. A great conceptual simplification is afforded by the fact that variations on a common BCS design can now be used to explain large data bases about form and motion perception that have heretofore been treated as wholly separate. The properties of the static BCS and the motion BCS suggest a computational explanation for why parallel parvocellular and magnocellular visual processing streams exist (Figure 1). As noted above, the static BCS models the parvo → interblob → interstripe → V4 stream. The motion BCS models the magno → 4B → thick stripe → MT stream. Tests of FM symmetry can be made by comparing static BCS and motion BCS model properties with neurobiological properties of these parvocellular and magnocellular streams—notably, how they support overlapping but distinguishable combinations of visual properties (Logothetis, Schiller, Charles, & Hurlbert, 1990; Schiller, Logothetis, & Charles, 1990).

The computer simulations of motion perception data that were reported by Grossberg and Rudd (1989a, 1989b, 1992) used a 1-D version of the MOC filter. Evidence for this MOC filter includes its ability to explain many classical and recent data about short- and long-range motion properties, and about cortical cell properties, that have not yet been explained by alternative models. Grossberg and Rudd (1989a) have also shown how the main properties of other motion perception models can be assimilated into different parts of the motion BCS design. The present article simulates data that require a 2-D version of the MOC filter. In addition, we outline how outputs from the 2-D MOC filter input to a 2-D version of the MOCC loop. To motivate these results, we first discuss key motion segmentation and grouping phenomena, including the aperture problem, barberpole illusion, and motion capture. Then we discuss how these phenomena suggest the existence of a MOCC loop that is analogous to the static CC loop of Grossberg and Mingolla (1985a, 1985b, 1987) but is specialized to process moving images. We then provide a self-contained description of the MOC filter of Grossberg and Rudd (1989a, 1992) and show how the MOC filter model can be extended to effectively process 2-D moving images. These extensions include hypotheses concerning the role of endstopped simple cells, the spatial layout of simple-cell receptive fields, and competition among signals sensitive to different directions of

motion. We illustrate these concepts through computer simulations that study how the motion BCS responds to changes in the bounding orientations, shapes, and motion directions of an object. These results are used to explain data about the aperture problem, barberpole illusion, and motion capture.

A central theme of both the static BCS and the motion BCS is that a specialized multistage filter inputs to a specialized grouping network. In the SOC filter (Figure 2), oriented simple cells filter outputs from the model LGN. The simple cell outputs are, in turn, rectified and combined at complex cells. The complex cell outputs are filtered again to generate inputs to the hypercomplex cells. Such a double-filter model has subsequently been used by a number of investigators to fit psychophysical data about texture segmentation (Beck et al., 1991; Beck et al., 1990; Graham et al., in press; Sutter et al., 1989).

In the case of the MOC filter, and more generally the motion BCS, this multistage filtering process sheds new light on a long-standing problem in the motion perception literature. As shown in Section 6 below, the MOC filter possesses analogs of both a short-range motion filter and a long-range motion filter. It should not, however, be concluded from this distinction that the MOC filter proposes separate short-range and long-range motion *systems*, as many authors have previously proposed (Anstis, 1980, 1986; Braddick, 1980). Quite the contrary is true. The distinction between short- and long-range spatial filters clarifies computationally how distinct short-range and long-range motion properties can occur, but these properties emerge in the MOC filter as part of a unitary motion processing system. Thus, the MOC filter of Grossberg and Rudd (1989a), and our refinement thereof, sheds new light on the recent conclusion that the weight of experimental evidence indeed supports the idea that a unitary motion system exists (Bischof & DiLollo, 1990; Cavanagh & Mather, 1989; Cleary & Braddick, 1990; Sperling, 1989).

In particular, properties of the MOC filter and the MOCC loop mechanistically clarify the conclusion of Cavanagh and Mather (1989) that "the observed differences forming the basis of the claims for two motion processes may be more easily attributed to differences in the stimuli used in traditional short-range and long-range experiments than to differences between two motion processes" (p. 105). Cavanagh and Mather instead distinguish between first- and second-order stimuli, the former defined by displacement of luminance and color differences, and the latter defined by displacement of statistical properties of texture, motion, or binocular disparity with mean luminance and color held fixed. The motion BCS can process both types of stimuli in a way consistent with data of Cavanagh and Mather showing that "both first- and second-order motion processes . . . require access to approximately the same amount of a grating cycle to determine the direction of stimulus motion" (p. 119). Many of the other perceptual properties observed in response to first- and second-order stimuli are also consistent with the analyses of Grossberg and Rudd (1989a, 1992) and those offered below, particularly the differences in perceptual properties caused by the short-range filter, the long-range filter, and the MOCC loop across multiple spatial scales.

The hypercomplex cells of the static BCS interact to activate higher order hypercomplex cells. In this interaction, similarly oriented hypercomplex cells cooperate, but dissimilarly oriented hypercomplex cells compete, with competition maximal between perpendicular orientations (Grossberg & Mingolla, 1987). The cooperating hypercomplex cells tend to perform a *vector average* of their input orientations. The higher order hypercomplex cells, in turn, input to bipole cells whose feedback helps to choose the emergent boundary segmentation that forms the best compromise among the image constraints that survive the filtering process.

The motion BCS has a similar organization, although it is specialized to deal with motion directions rather than static orientations. It is shown below how several differently oriented responses may be pooled to generate a stable estimate of motion direction. Orientational sensitivity of the motion-direction cells is not entirely lost in this process, much as motion-sensitive cells in MT preserve some orientational sensitivity (Albright et al., 1984). Thus, when we speak of directional sensitivity below, we do not imply an absence of orientational tuning.

Given this caveat, we can say that the motion BCS has homologs of all the processing stages of the static BCS. In particular, there is a double filter, followed by direction averaging and competition stages, followed by a feedback network for generating emergent segmentations of motion form. The term *motion form* is used, rather than *motion*, because the motion BCS is not a "motion" processing system. It is a system for generating emergent segmentations that are sensitive to coherent properties of the directions of motion of moving forms. These similarities between the static BCS and motion BCS segmentation systems clarify neurophysiological data showing a considerable degree of functional overlap between the response capabilities of the parvocellular and magnocellular processing streams, including the ability of both systems to process coarse form, coarse stereo, slow motion, and flicker at high contrast (Logothetis et al., 1990; Schiller et al., 1990). The need for a separate motion segmentation system can best be understood by analyzing the types of processing ambiguities and errors that can occur without it.

2. Global Segmentation and Grouping: From Locally Ambiguous Motion Signals to Coherent Object Motion Signals

The *aperture problem* describes the percept whereby a straight edge or grating viewed through a circular aperture appears to be moving perpendicular to its orientation of contrast regardless of its true motion direction (Wallach, 1976). While this description of the aperture problem is at the molar perceptual level, any early visual cell with a localized receptive field, whether concentric or elongated, experiences its own "aperture problem," to the extent that its estimate of motion direction is a

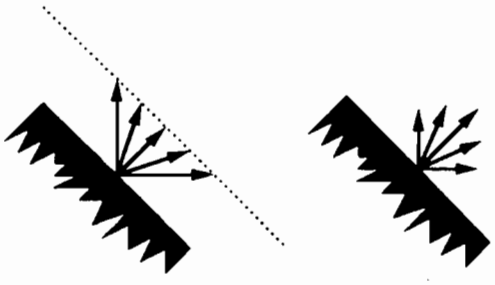


Figure 3. (a) This "velocity space" construction, adapted from Adelson and Movshon (1982), illustrates that the motion of a given point on a moving edge could be along any of the trajectories whose arrows end on the dashed line. (b) A complementary way to view this direction ambiguity is to consider that local motion signals from a straight edge are *strongest* in the perpendicular direction, while covering a range of possible directions. (From "Phenomenal coherence of moving visual patterns" by E. H. Adelson and J. A. Movshen, 1982, *Nature*, 300, 523-525. Copyright 1982 by Macmillan Magazines Limited. Reprinted by permission.)

spatiotemporally weighted function of changes in local stimulation only. In their discussion of "velocity space," Adelson and Movshon (1982) introduced diagrams similar to Figure 3a to illustrate the ambiguity of local motion direction and speed from information confined to an aperture. In this figure, the length of arrows codes possible trajectories of the point *A* that would be consistent with the measured change of contrast over time of the cell in question. For this reason, it is sometimes said that such a cell is sensitive to only the normal component of velocity. A complementary way of thinking about this situation is illustrated in Figure 3b, in which the length of arrows is roughly proportional to the cell's "prior probability distribution" for interpreting changing stimulation as occurring in one of several directions. The direction perpendicular to the cell's receptive field's orientation is locally preferred. In this conception, a cell with an oriented receptive field, such as a simple cell, may be stimulated by a moving edge that is not perfectly aligned with its receptive field's dark-to-light contrast axis. However, it is assumed that within a hypercolumn of cells that are tuned to similar position, spatial frequency, contrast, and temporal parameters but vary in preferred orientation (Hubel & Wiesel, 1977), some other cell whose preferred orientation was more nearly aligned with the moving edge would generate a stronger signal than would the cell in question. Thus, the distribution of local motion signals across cells tuned to all orientations at a given position would favor the direction perpendicular to the orientation of the edge.

The barberpole illusion (Wallach, 1976) shows that these local preferences can be readily overridden by global factors related to the forms within which stationary and moving segments of a display are located (Figure 4). This illusion indicates that, to the extent that locally *unambiguous* motion signals exist at line ends or corners, those unambiguous signals can somehow enforce an interpretation of motion direction consistent with their own throughout the length of a contour bounded by those ends or corners (Figure 5). Clearly, some long-range grouping process subsequent to the early generation of local motion signals is at work. This process appears to cooperatively "choose and sharpen" within relatively large domains of ambiguous signals on the basis of relatively localized, but unambiguous, signals. We will illustrate below how this is achieved by the MOCC loop.

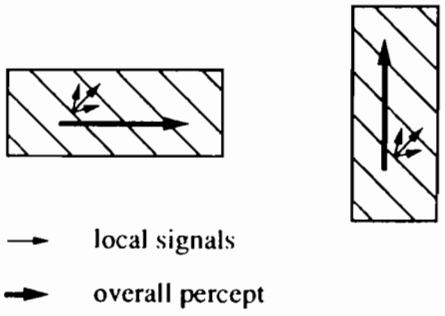


Figure 4. In the barberpole illusion, a striped pattern is perceived as moving in the direction of elongation of a rectangular frame (or aperture). Local motion signals, indicated by thin arrows at only one location within each rectangle, are generated throughout the interior lengths of each diagonal line. As indicated in Figure 2b, these local motion signals express both ambiguity and preference regarding direction of motion. Despite the large area covered by the ambiguous and diagonal-motion-preferring local signals, the resulting percept is horizontal or vertical, depending on the configuration of a visible frame (or aperture).

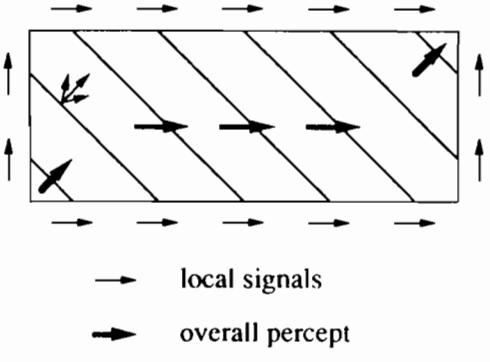


Figure 5. A closer look at the barberpole illusion reveals that for sufficiently large displays, whose length-to-width ratio is not too far from unity, the overall motion percept may vary across different areas of the display. The perceived motion near the lower left and upper right corners of the rectangle may be diagonal, while horizontal motion is seen through the bulk of the display. Unambiguous motion signals are generated in the region where each diagonal line meets the horizontal and vertical contours of the rectangle. (These unambiguous signals are diagrammed *outside* the rectangle for clarity.) Evidently, such unambiguous signals exert an influence on the percept that is disproportionate to their areal extent, since the overall percept throughout a diagonal line tends to be a resultant (horizontal plus vertical to diagonal or horizontal plus horizontal to horizontal) of those signals.

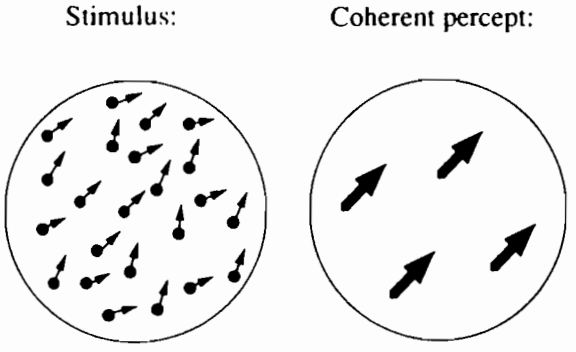


Figure 6. Williams and Sekuler (1984) found evidence for cooperativity in motion perception using an ingenious variant of the random dot cinematogram paradigm. The local motions of individual dots support a percept of coherent global motion, as well as the perception of individual dot trajectories. See text for details. (Reprinted with permission from *Vision Research*, 24, D. Williams and R. Sekular, "Coherent global motion percepts from stochastic local motions," Copyright 1984 by Pergamon Press Ltd.)

Cooperativity among motion signals was also studied by Lappin and Bell (1976) and Williams and Sekuler (1984), using displays such as that described in Figure 6. Random dot cinematograms were displayed in which successive displacements of dots over frames were not uniform across all dots, but rather were sampled from a rectangular distribution of possible directions. On any frame, each dot's displacement was independent of both its own history of displacements and of the displacement of other dots in that frame. For appropriate parameter choices, observers reported perceiving a coherent global motion in the direction of the mean of sampled local dot motions, as well as perceiving the individual motions of each dot. Williams and Sekuler also reported hysteresis effects, whereby the percept of coherent motion persisted while motion distribution parameters were altered from a relatively narrow range, which easily supported the coherent percept, to a wider range, for which coherent motion was not ordinarily seen if presented initially. Thus, cooperativity among motion signals involves averaging of directional information and hysteresis, as well as sharpening and choice.

Ramachandran (1981) and Ramachandran and Inada (1985) provided additional evidence for cooperation among motion signals in a phenomena termed *motion capture*. As described in Figure 7, strong and unambiguous motion signals can actively reorganize motion signals in ambiguous regions where there are no locally preferred motion directions. From this perspective, the barberpole illusion can be analyzed as displaying its own form of "motion capture," as indicated in Figure 8. That is, signals from the ambiguous interior region of the diagonal line segment are being "captured" by signals from the unambiguous ends. This enhancement of horizontal signals in the interior region is accompanied by the suppression of diagonal and vertical signals. Long-range cooperation is thus accompanied by short-range competition in order to enforce a clear choice of percept all along the line. As noted below, this cooperative-competitive interaction is modeled by a MOCC loop whose processes are analogous to that of the static CC Loop.

The phenomenal coherence of moving plaid patterns can also be analyzed in terms of the cooperative-competitive

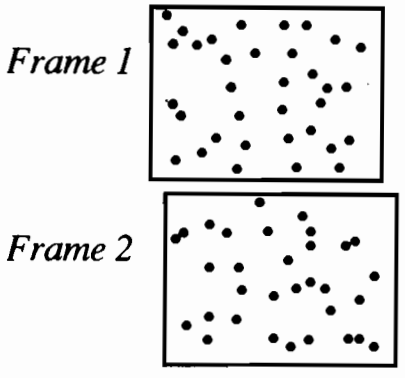
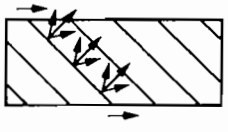
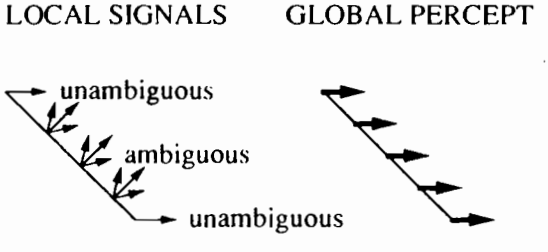


Figure 7. Ramachandran and Inada (1985) describe a study by Ramachandran (1981) in which a two-frame apparent-motion display contains unambiguous and ambiguous regions. For appropriate displacement, flash duration, and interstimulus intervals, the outer rectangular contour provides strong and unambiguous signals for rightward motion. Since the interior region contains random dots whose distributions are uncorrelated between frames, local signals are generate in all directions within the interior. Notwithstanding, the percept is described as being one of rightward motion for *both* the rectangular contour *and* the interior dots. (Reprinted by permission from *Spatial Vision*, 1(1), 57-67. Copyright 1985 by VSP, The Netherlands.)



(a)



(b)

Figure 8. (a) By focusing on the distribution of motion signals on a single diagonal line, as indicated in the barberpole illusion, the phenomenon can be analyzed as a form of motion capture. (b) The unambiguous signals from the line ends help to enhance signals of like direction (cooperation) and suppress signals of different directions (competition) within the interior of the line segment.

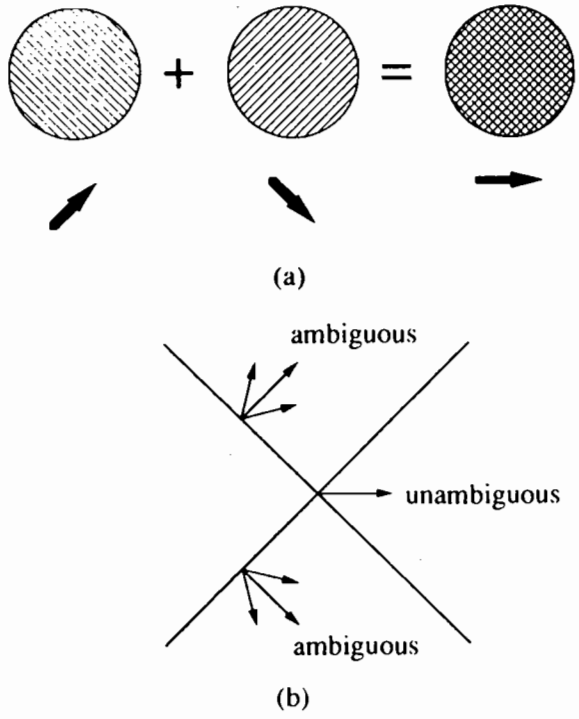


Figure 9. (a) As described by Adelson and Movshon (1982), under appropriate conditions of velocity, contrast, and spatial frequency, a moving plaid pattern consisting of a superposition of two oriented gratings will exhibit a coherent, rigid motion in a direction not perceived when either grating is displayed alone. (b) The phenomenal coherence of such plaids can be analyzed as another example of motion capture, insofar as unambiguous signals from regions of intersection can dominate and reorganize the signals coming from segments of only one component of the plaid. See text for details. (From "Phenomenal coherence of moving visual patterns" by E. H. Adelson and J. A. Movshon, 1982, *Nature*, 300, 523-525. Copyright 1982 by Macmillan Magazines Limited.)

interactions of the motion BCS. As indicated in Figure 9, under appropriate conditions of velocity, contrast, and spatial frequency content, observers often report experiencing coherent and rigid motion of the entire pattern in a direction consistent with the "velocity space" interpretation (Figure 3a) of each component's possible motions (Adelson & Movshon, 1982). Movshon, Adelson, Gizzi, and Newsome (1985) reviewed several lines of evidence for rejecting the hypothesis that the visual system merely tracks "features" formed by local maxima in luminance at the component intersections. In one experiment, "one-dimensional" noise, consisting of parallel stripes of varying widths, was superimposed on plaid patterns, at times in the orientation parallel to one of the component gratings and at times in the orientation perpendicular to the coherent plaid motion. They measured threshold elevation for detection of coherent motion and found that noise within about 20° of one of the component orientations was much more efficacious in masking the resulting percept than was noise that was perpendicular to the direction of plaid motion. They concluded that "the mechanisms responsible for the phenomenal coherence of moving plaids belong to a pathway which, at some point, passes through a state of orientation selective spatial analysis" (p. 130). In another experiment, they also measured the effects of adaptation on coherence, using a factorial design with which single gratings or plaids were used as both adaptation and test stimuli. The strongest adaptation effects were found in the conditions where identical component orientations appeared in both adaptation and test stimuli, again implicating an orientation-specific site of early processing.

The velocity space interpretation of these data does not always predict the same result as the vector averaging of motion directions that occurs at the direction pooling stages of the motion BCS (Sections 1 and 11). In addition, motion BCS computations also suggest that the processing stages for generating coherent motion directions and their interactions over time occur prior to the processing stages that compute velocity signals as such. The distinction between velocity space and vector averaging was experimentally investigated by Mingolla, Todd, and Norman (1992), who examined how ambiguous velocity measures along smooth contours are spatially integrated to obtain a globally coherent perception of motion. Observers viewed displays containing a large number of apertures, with each aperture containing one or more contours whose orientations and velocities could be independently specified. The total pattern of the contour trajectories across individual apertures was manipulated to produce globally coherent motions, such as rotations, expansions, or translations. When the displays contained only straight contours extending to the circumferences of the apertures, observers' reports of global motion direction were biased whenever the sampling of contour orientations was asymmetric relative to the direction of motion. Performance was improved by the presence of identifiable features, such as line ends or crossings, whose trajectories could be tracked over time. The reports of these observers were consistent with the motion BCS prediction. In particular, a pooling process computes a vector average of the velocity components normal to contour orientation, rather than the intersection of constraints that define velocity space. Wilson (1991) has reported related data and a model for pooling motion signals of different directions that is similar to the motion BCS, including its double-filter and vector averaging stages.

It will be shown in subsequent sections how the motion BCS model is capable of deriving unambiguous motion signals at plaid intersections in a way that survives the Movshon et al. (1985) arguments against feature tracking, while providing an alternative explanation of plaid motion using the same model that has been used to explain many other motion perception data. All of these motion data point to the need for interactions among long-range cooperation and short-range competition within networks capable of enhancing, suppressing, linking, and completing motion signals using feedback signals capable of hysteresis.

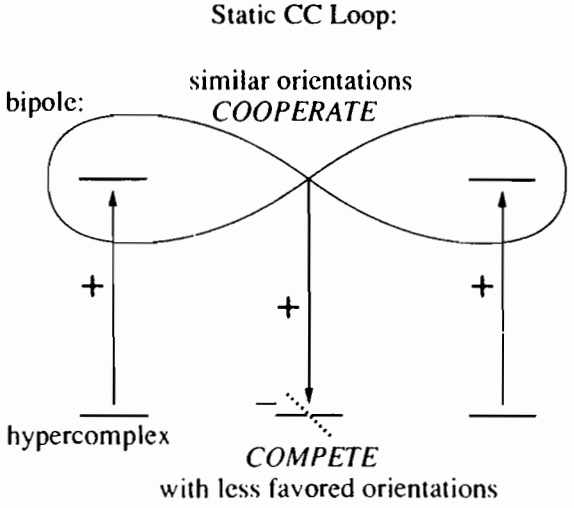


Figure 10. In the static cooperative-competitive feedback loop (CC loop) signals of like orientation (here horizontal) initiated from "bottom-up" input data arrive at two locations separated by some distance (indicated by the horizontal bar above the word *hypercomplex* and the corresponding bar at the lower right). Cooperative bipole cells, identified with projections of V2 cells, can, if they receive sufficient stimulation on *both* sides of their receptive field centers, send positive feedback to cells of like orientation in the hypercomplex layer. There, cells of differing orientation and like position compete with each other to determine the most favorable orientation at each location. Because of the long-range feedback, local orientational preferences are modulated to be consistent with contextual data. While, in this example, the completion is collinear, the "figure eight" shape of the bipole cell indicates that completion can be curvilinear, so long as a consistent co-occurrence of oriented signals favors interpolation at the bipole center. See Figure 31 and Equations A13-A21 of the Appendix of Grossberg and Mingolla (1985b) for a detailed explanation of static bipole cells.

3. Coherent and Resonant Completion: Static and Motion CC Loops

Static BCS mechanisms combine long-range cooperation with short-range competition among orientationally tuned cells within the static CC loop feedback network (Figures 2 and 10). Since publication of the CC loop equations in Grossberg and Mingolla (1985a, 1985b, 1987), increasing physiological evidence for the model has been reported. In particular, Peterhans and von der Heydt (1989) have recently reported evidence for cooperative linking interactions in cortical area V2. Gray, Konig, Engel, and Singer (1989) and Eckhorn et al. (1988) have reported resonant interactions involving rapid phase locking of spike trains of V2 cells with nonoverlapping receptive fields, which is also a property of the long-range cooperative interactions that occur within the CC loop (Grossberg & Somers, 1991, 1992).

Given that the static CC loop can enhance consistent signals and suppress inconsistent or ambiguous signals from orientationally tuned cells, and given that in vivo early orientationally tuned cells, such as simple cells, are sensitive to motion, the question naturally arose whether motion segmentation could be accomplished by the static CC loop. Our daily experiences with the dynamic geometry of form perception for static and moving contours suggest that this is not true. As illustrated in Figure 11, a *static* form system is more concerned with *orientation* of static contours, while a *motion* form system is more concerned with *direction* of moving contours. This distinction must be made with care, however, since both in the neurobiological data concerning these systems (Schiller et al., 1990) and in our modeling of their properties both here and elsewhere (Grossberg, 1987a, 1991; Grossberg & Mingolla, 1985b; Grossberg & Rudd, 1989a, 1992), cells in the static form system exhibit preferred movement directions and cells in the motion form system exhibit preferred image orientations, albeit not with the selectivity of cells in the static form system. Such multiplexing of properties does not support the popular hypothesis (Livingstone & Hubel, 1987) that the visual system is decomposed into modules that are devoted to processing a single visual property, such as orientation, motion, stereo, or color. A rigorously defined computational theory is needed to articulate how such multiplexing occurs and what functional properties are accomplished by its use. It will be shown below, for example, that keeping track of direction as well as orientation requires an additional degree of freedom, since a segment of a given orientation may be moving in any of several directions, and a given direction of motion can be observed for image contrasts that are moving with any of several orientations. This is just another way of stating the aperture problem.

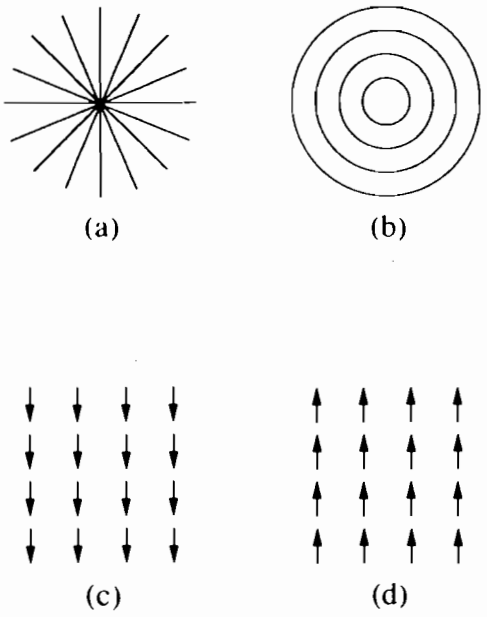


Figure 11. As indicated by such static rebound phenomena as the Mackay effect, in which prolonged fixation on a pattern of radial line segments (a) results in an aftereffect of perceived circular contours (b), the opposite of a static *orientation* signal is a signal of perpendicular orientation. The waterfall illusion, in which prolonged adaptation to motion signals as from a waterfall (c) is followed by the sensation of upward motion when one looks at a neutral scene (d), indicates that the opposite of a motion *direction* signal is not 90° but 180° from that signal.

The striking similarity of cooperative and competitive grouping requirements for both static and moving images, together with the existence of different geometries for static and motion perception, suggested that parallel versions of the CC loop exist within a static BCS and a motion BCS. This taxonomy of "static" and "motion" CC loops does not, however, imply logical exclusivity of function, any more than the existence of distinct systems implies that they compute independent visual properties. When a contour moves, the *static* CC loop may operate to determine the best coherent *orientations* of the moving contour, at the same time that the *motion* CC loop determines the best coherent *directions* of that contour, without denying that each CC loop is capable of processing aspects of both form and motion and that both CC loops interact to form a final unitary percept. However, the better calibration of orientational tuning and orientationally based stereo matching by the static BCS is suggested to help select directionally based boundary segmentations of the motion BCS that are processing image data at correctly calibrated depths (Grossberg, 1991). This interaction has been predicted to occur via the V1→V2→MT pathway (see Section 15).

The proposed function of the V1→V2→MT pathway does not imply that computations within the motion form system are themselves incapable of estimating properties of depth. It is known from neurophysiological experiments (Logothetis et al., 1990; Schiller et al., 1990) and psychophysical experiments (Julesz, 1971; Prazdny, 1985, 1986) that depth may be derived from motion cues in the absence of normal cues about static form. Correspondingly, it has been suggested that the final stages of the MOC filter become binocular in a manner homologous to that used by the binocular complex cells in the parvocellular processing stream of area V1 (Grossberg & Rudd, 1989a, 1992), but can only carry out binocular matches using its larger directionally tuned receptive fields. The V1→V2→MT link is suggested to join the two types of binocular computation in a way that exploits the best properties of each. To the present time, modeling of how the static BCS generates 3-D segmentations capable of supporting figure-ground pop-out (Grossberg, 1987b, 1992, in press; Grossberg & Marshall, 1989) has proceeded considerably beyond a corresponding analysis of 3-D segmentation by the motion BCS. The computational similarities between the two systems promise to accelerate progress on understanding binocular processing within the motion BCS as well.

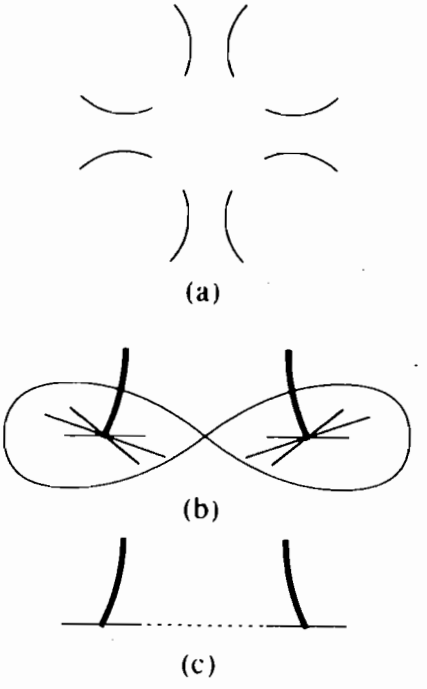


Figure 12. (a) This illusory contour display, adapted from Kennedy (1979), indicates that boundary completion in the static CC loop can choose orientations that are not locally preferred if global organizational factors are sufficiently powerful. (b) The locally preferred directions at the bottom ends of the two top curves of (a) are perpendicular to the ends of the curves and thus tilted off the horizontal. (c) Cooperation among signals that are horizontal enhances and completes horizontal signals (cooperation) while suppressing non-horizontal signals (competition) along the illusory contour. (Kennedy, J. M., Subjective contours, contrast, and assimilation. In C. F. Nodine and D. F. Fisher (Eds.), *Perception and pictorial representation*. Copyright 1979 by Praeger Publishers, an imprint of Greenwood Publishing Group, Inc., Westport, CT. Adapted by permission.)

4. Many Orientations Can Move in the Same Direction

We begin our development of the monocular motion BCS by suggesting how the cooperative bipole cells of the motion CC loop differ from their homologs in the static CC loop of Figure 10. In analyzing the properties of static bipole cells, certain images by John Kennedy proved invaluable (Kennedy, 1979). As shown in Figure 12, static bipole cells are capable of choosing orientations that are not locally preferred for generating positive feedback to cooperatively link and complete static boundaries. Each bipole cell can generate an output signal only if *both* of its oriented receptive fields receive large enough inputs from cells with similarly oriented receptive fields. These input cells at the previous processing level consist of formal analogs of hypercomplex cells (Figure 2). The output signals from the bipole cells feed back to the hypercomplex cells (Figure 10), where they bias the competition among the hypercomplex cells toward the orientations and positions that are most favored by the bipole cells.

The MOCC loop must, however, cope with an additional degree of freedom, since it considers direction as well as orientation. Thus, motion bipole cells are organized in a fashion that differs somewhat from that of their static analogs. As indicated in Figure 13, motion bipole cells are postulated to exist in families sensitive to motion signals of different directions but whose sources are arrayed in patterns of similar orientation. Expressed differently, while some motion bipole cells are assumed to favor the direction of motion perpendicular to the orien-

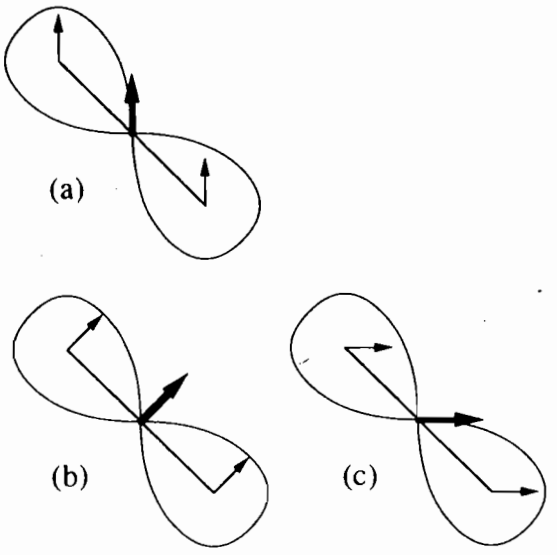


Figure 13. As with static bipole cells, motion bipoles are sensitive to bottom-up direction signals from motion hypercomplex cells (indicated by thin arrows), which send excitatory signals to each lobe of the bipole cell. If sufficient excitatory activity is sensed in *both* lobes, the bipole sends feedback signals of like direction (indicated by thick arrows) to the hypercomplex layer. While analogous to their static counterparts, motion bipole cells are fundamentally different, insofar as they must cope with the additional degree of freedom imposed by the simultaneous determination of a globally consistent motion direction over many possible contour orientations. Thus, families of bipole cells are presumed to exist, such that cells whose major axes are the same (diagonal in the present case) can be maximally sensitive to motion signals of different directions, ranging from (a) vertical to (b) diagonal (perpendicular to the major axis) to (c) horizontal for the three bipoles shown here.

tation induced by the elongated axis of their two lobes (Figure 13b), others have a similar spatial layout but different preferred directions of motion (Figures 13a and 13c).

As in the static CC loop, each bipole cell of the MOCC loop can generate an output signal only if *both* of its receptive fields receive large enough inputs from cells that are sensitive to a similar direction of motion. These input cells at the previous processing level consist of formal analogs of hypercomplex cells. The output signals from the bipole cells feed back to the hypercomplex cells, where they bias the competition among the hypercomplex cells toward the directions and positions that are most favored by the bipole cells.

It is intuitively clear that this is just the sort of cooperative feedback, propagated inward between pairs, or larger numbers, of flanking inducers, that is needed to explain phenomena such as the barberpole illusion (Figures 4 and 5) and motion capture (Figures 6–8). Our analysis elevates this intuition into a computationally precise theory. As a result, both orientationally sensitive grouping during static form perception and directionally sensitive grouping during motion form perception are predicted to utilize bipole cells. The explanatory power of this homology strengthens the case that both the static BCS and the motion BCS architectures model variations of a common cortical design.

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

In order to design a MOCC loop, its hypercomplex cells need to be sensitive to a prescribed direction of motion. These cells may be excited by image contours with different orientations, all moving in the same direction within a prescribed region of perceptual space. The MOC filter thus contains cells that derive their sensitivity to direction of motion by pooling outputs from cells that are maximally sensitive to several different orientations. Typically, orientations that are nearly perpendicular to the preferred motion direction are pooled, as is also observed in MT cells (Albright et al., 1984). Thus, a cell that prefers motion to the right pools signals from oriented cells that prefer vertical, or almost vertical, image contrasts. In order to synthesize directional sensitivity from several different orientations, the MOC filter needs to have a circuit design somewhat different from that of the SOC filter. The reason for this modification is that complex cells of the SOC filter (Figure 2) are insensitive to direction of motion as well as to direction of contrast, in the sense that (say) a vertically oriented SOC filter complex cell can respond to vertically oriented dark-light and light-dark contours moving either to the left or to the right. Complex cells in the parvocellular stream of V1 exhibit similar response properties (Foster et al., 1985; Pollen, Gaska, & Jacobsen, 1989). In this sense, the SOC filter cannot be used to selectively 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 direction of contrast of the SOC filter's complex cells enables the static CC loop (Figure 10) to generate boundary segmentations along scenic contrast reversals.

The simple cells at the first BCS level are, however, sensitive to direction of contrast (Figure 2). 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. Inspection of the simple-cell to complex-cell interaction in Figure 2 shows how a vertically oriented complex cell can respond to a dark-light or light-dark vertical edge moving to the right or to the left. Thus, the process whereby complex cells become insensitive to direction of contrast renders 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 segmenting moving images as for segmenting static images—yet is also *sensitive* to direction of motion—a property that is certainly

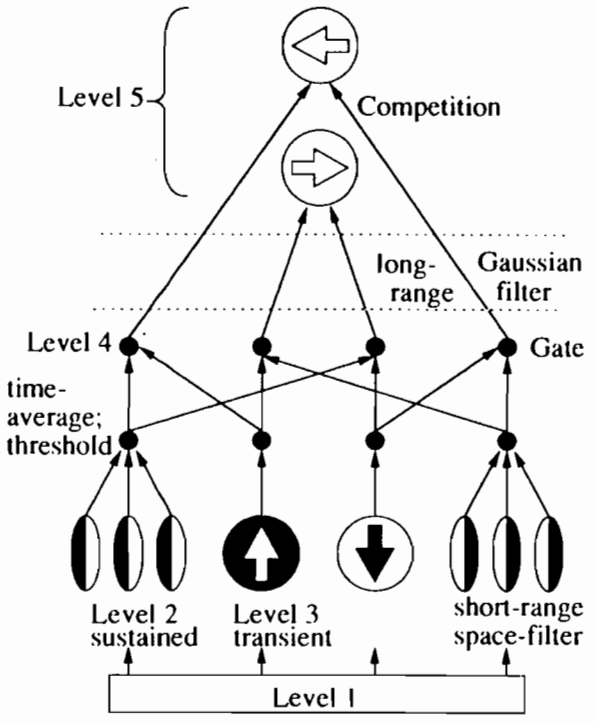


Figure 14. The motion oriented contrast (MOC) filter: Level 1 registers the input pattern. Level 2 consists of sustained response cells with oriented receptive fields that are sensitive to direction of contrast. Level 3 consists of transient response cells with unoriented receptive fields that are sensitive 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 with Level 4 cells to become sensitive to direction of motion and insensitive to direction of contrast. (Reprinted, by permission, from Grossberg, 1991.)

essential in a motion perception system. Along the way, the MOC filter introduces an extra degree of computational freedom that achieves several important properties in one stroke: sensitivity to direction of motion, long-range motion interactions, and binocularity (Grossberg & Rudd, 1989a, 1992).

6. The Grossberg-Rudd MOC Filter

A MOC filter is mathematically defined in Grossberg and Rudd (1989a, 1992). Its five processing stages are qualitatively summarized in Figure 14 and described below. This MOC filter was used to carry out 1-D simulations of motion data. After summarizing the Grossberg-Rudd model, we describe the refinements that are needed to disambiguate orientation and direction using 2-D moving figures. It should be emphasized that variations on these processing stages can be envisaged that give rise to similar computational properties. The *types* and *ordering* of the processing stages and their interactions are needed to explain the full body of data to which they have been applied. Model variations with similar properties will be noted at several places in the following exposition.

Level 1: Preprocess input pattern. The image is preprocessed before activating the 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).

Level 2: Sustained cell short-range filter. Four operations occur here, as illustrated in Figure 15.

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 direction of contrast pool their output signals to activate the next cell level. The breadth of spatial pooling plays the role of the short-range motion limit D_{max} (Braddick, 1974). Several copies of this short-range filter exist, one for each copy of the motion BCS. Each copy corresponds to a different range of receptive-field sizes. The breadth of spatial pooling within each copy scales with the size of the simple-cell receptive fields. Thus, D_{max} within the MOC filter is not independent of the spatial frequency content of the image, as is also true in vivo (Anderson & Burr, 1987; Burr, Ross, & Morrone, 1986; Nakayama & Silverman, 1984, 1985) and is not a universal constant. In addition, the direction in which individual oriented cells are spatially pooled may not be perpendicular to the oriented axis of the sustained cell receptive field. The target cells are thus sensitive to a movement *direction* that may not be perpendicular to the simple 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).

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 16). This filter also uses four operations.

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

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

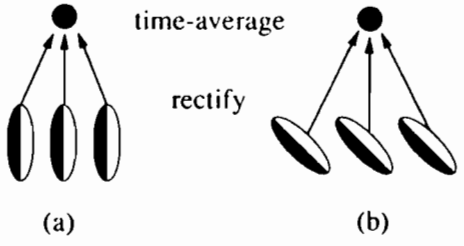


Figure 15. The sustained cell short-range filter combines several spatially contiguous receptive fields of like orientation via a spatial filter with a fixed directional preference. The orientation perpendicular to the direction is preferred, but nonorthogonal orientations can also be grouped in a prescribed direction. (Reprinted, with permission, from Grossberg, 1991.)

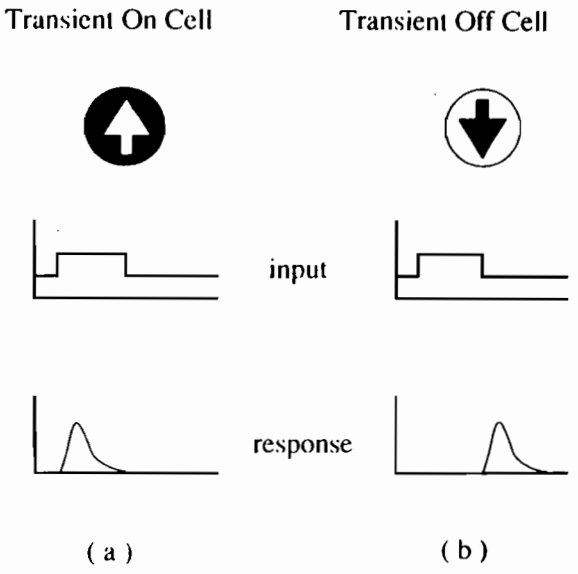


Figure 16. The transient cell filter consists of on-cells that react to input increments and off-cells that react to input decrements. (Reprinted, by permission, from Grossberg, 1991.)

3. *Transient detector*: The on-cells are activated when the time-average increases. The off-cells are activated when the time-average decreases.

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

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 17). For example, consider a sustained cell output from vertically oriented light-dark receptive fields that are joined together in the horizontal direction by the short-range spatial filter. Such a sustained cell output is maximized by a light-dark 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 right. Multiplying it with a Level 3 transient off-cell output generates a Level 4 cell that responds maximally to motion to the left.

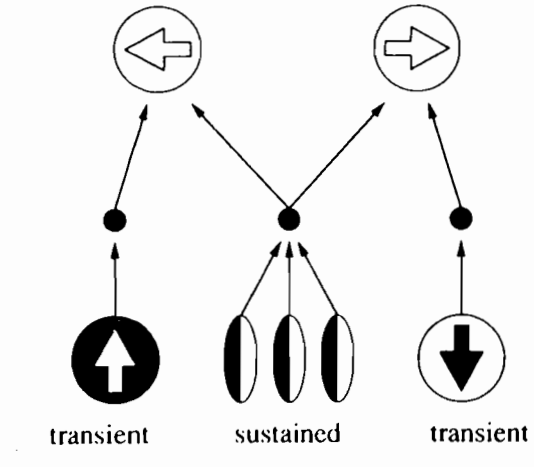


Figure 17. Sustained-transient gating generates cells that are sensitive to direction of motion as well as to direction of contrast. (Reprinted, by permission, from Grossberg, 1991.)

Multiplying a sustained cell with a transient cell is the main operation of the Marr and Ullman (1981) motion detector. Despite this point of similarity, Grossberg and Rudd (1989a) described six basic differences between the MOC filter and the Marr-Ullman model. For example, none of the operations, such as short-range spatial filtering, time-averaging, and rectification, occurs in the Marr-Ullman model. In addition, the rationale of the MOC filter—to design a filter that is sensitive to direction of motion and insensitive to direction of contrast—is not part of the Marr-Ullman model. This difference is fundamental. The Marr-Ullman model is a product of the "independent modules" perspective. The MOC filter's insensitivity to direction of contrast can only be formulated within the framework of a design principle that we have called *BCS/FCS complementarity* (Grossberg et al., 1989): One cannot understand why a boundary filter's output needs to be insensitive to direction of contrast unless there exists a complementary "seeing" system, the Feature Contour System, that is sensitive to direction of contrast.

The remaining processing levels of the MOC filter are also not part of the Marr-Ullman model. In particular, 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, as described below.

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 is Gaussianly distributed across space (Figure 18). This long-range filter groups together Level 4 cell outputs with the same directional preference but different preferred cell orientations and directions of contrast at Level 3. Thus, the long-range filter provides the extra degree of freedom that enables Level 5 cells to function as "direction" cells, rather than as "orientation" cells.

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, this competitive process partially contrast-enhances ("leader-take-most") its input pattern to generate a motion signal whose breadth

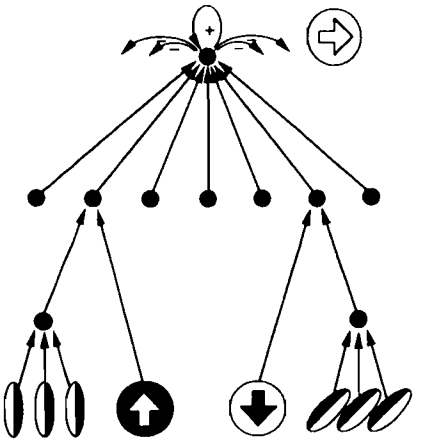


Figure 18. The long-range spatial filter combines sustained-transient cells with the same preference for direction of motion, including cells whose sustained cell inputs are sensitive to opposite directions of contrast and to different orientations. (Reprinted, by permission, from Grossberg, 1991.)

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, 1987b; Grossberg & Marshall, 1989).

The Level 5 cells of the MOC filter are, in other respects too, computationally homologous to SOC filter complex cells. In particular, Grossberg and Rudd (1989a, 1992) have hypothesized that MOC filter cells become binocular no later than Level 5 in order to explain data about interocular transfer of apparent motion. This interaction could, in principle, occur as a result of either short-range spatial filtering or long-range spatial filtering. Binocular short-range filtering would enable sharper spatial and orientational matches to occur in response to dynamic stereo displays than would long-range filtering. In either case, no later than Level 5 of the MOC filter, its model cells possess the properties of directional, binocular, and orientational tuning that are found in lamina 4B cortical cells of the magnocellular processing stream (Figure 1).

As noted above, similarities also occur between the computational processes at higher levels of the motion BCS and static BCS. Do analogous parallels occur in vivo? In particular, the model hypercomplex cells and bipole cells of the static BCS may be compared with known cortical cell types in the interstripe region of area V2 and (possibly) beyond. If the motion BCS corresponds to a similar cortical organization, then the model hypercomplex cells and bipole cells of the motion BCS should have cortical analogs in the thick stripe region of area V2 and (possibly) beyond. An open question concerns whether Level 5 motion complex cells have cortical homologs in lamina 4B of area V1 or in the thick stripes of area V2. Since the model predicts the types and ordering of cortical processing, once the long-range filter is neurobiologically characterized, the final stage of the MOC filter can be precisely identified. Then the predicted hypercomplex cells and bipole cells should occur at subsequent cortical levels, in a prescribed order.

7. Pooling Orientations and Directions of Contrast to Compute 2-D Directions of Motion

The Grossberg-Rudd model was used to explain and simulate motion data that exhibit a natural 1-D symmetry—for example, apparent motion between collinear groups of flashes. The model needs further development to explain data concerning the motion of 2-D shapes, since such shapes may move in directions that may or may not be perpendicular to the orientations of their boundaries. The type of new issues that arise in the 2-D case are illustrated by the following example. Consider the lower right corner of a homogeneous rectangular form of relatively high luminance that is moving diagonally upward and to the right on a homogeneous background of relatively low luminance (Figure 19). Both the regions of horizontal and vertical contrast near the corner provide signals to the MOC filter, provided that the sustained cells of Level 2 (Figures 14 and 15) are spatially laid out as indicated in Figure 20. Here, the direction of motion is diagonal to the orientational preference of the individual sustained cells. These vertically and horizontally oriented cells contribute to the total signal that codes movement in the diagonally upward direction. So too do cells whose orientation is perpendicular to the diagonal direction. On the other hand, cells whose orientation is collinear with the direction of motion should not be included.

Accordingly, each motion detector is assumed to receive weighted inputs from all sustained-transient cells whose orientations differ from the preferred direction of motion by no more than 90°, and whose preferred positions lie within a prescribed distance from the preferred position of the motion detector. As indicated in Figure 21, the long-range filter (Level 5, Figures 14 and 18) can simultaneously accept motion signals from both the horizontal and the vertical edges of the moving corner, despite the gating of one set of signals by transient "luminance-

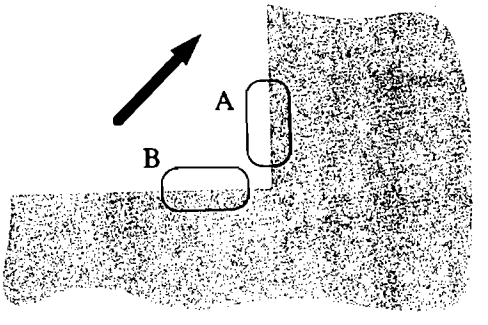


Figure 19. The lower right corner of a horizontally oriented rectangular region of homogeneous high luminance moves diagonally upward and to the right over a background of homogeneous low luminance. In region A, a dark-to-light (luminance increasing over time) transition occurs at a vertical edge, while, in region B, a light-to-dark (luminance decreasing over time) transition occurs at a horizontal edge.

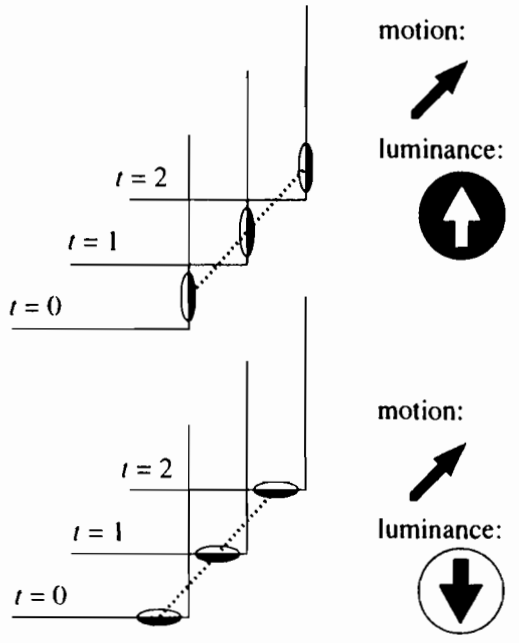


Figure 20. Over three successive time steps, the contours of the rectangle occupy the positions indicated, while luminance increases along the vertical edge and decreases along the horizontal edge. If certain that the sustained cell receptive fields sending inputs to Level 4 of the MOC filter were arranged as indicated, a diagonal motion signal could be generated from both vertically and horizontally oriented cells, in conjunction with luminance-gating signals of opposite signs.

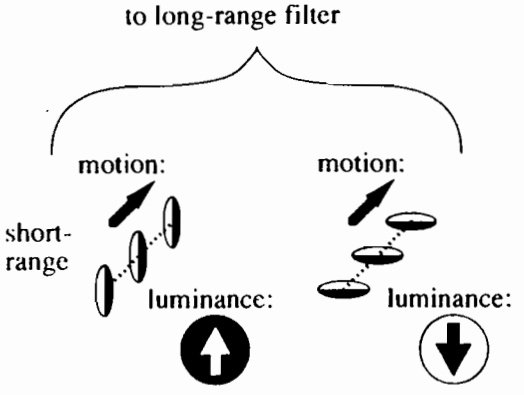


Figure 21. Signals arrive at the long-range filter (to Level 5 of the MOC filter) after several rectification and gating operations. Accordingly, signals that were gated by both increases and decreases in luminance at (necessarily) different places can be combined into a coherent motion signal.

increasing" on-cells and gating of another set by "luminance-decreasing" off-cells (Level 3, Figures 14 and 16). Thus, while simultaneous increase and decrease of luminance is logically impossible in an infinitesimal area, and while a too rapid change from increase to decrease may be unresolvable by sustained cells at Level 2, the simultaneous increase and decrease of luminance at different orientations and different locations, but in the same direction, are pooled by the long-range filter. This sort of long-range filtering by MOC filter complex cells is not the same as the still longer range cooperative grouping by MOCC loop bipole cells.

8. Sustained-Transient Gating before Short-Range Spatial Filtering

In attempting to simulate the 2-D MOC filter on the computer, a computational problem was noticed whose solution was not required in the 1-D simulations of Grossberg and Rudd (1989a, 1992). In particular, Figure 17 shows short-range spatial filtering of sustained cell outputs before gating the result with a transient on-cell or off-cell. The different spatial layouts of sustained cells in Figure 21 made it difficult to select a regular spatial arrangement of transient cells that could be used for all cases. This analysis suggested that each oriented sustained cell is first gated by a transient on-cell and off-cell at its own location before each sustained-transient cell inputs to the several directionally sensitive spatial filters to which it contributes. While these 2-D simulation problems were being confronted, Grossberg (1990, 1991) observed that the FM symmetry principle requires a similar spatial ar-

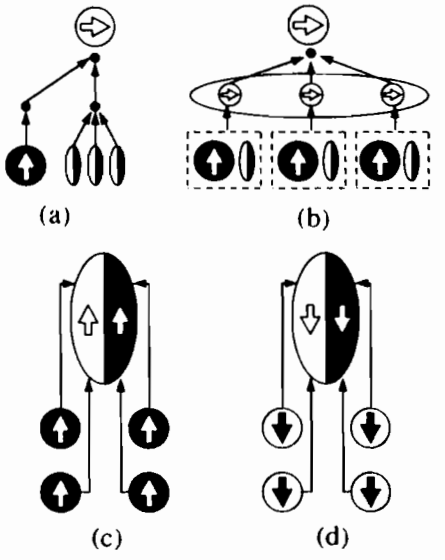


Figure 22. (a) The gating of sustained and transient cells within the MOC filter of Grossberg and Rudd (1989a) was between several sustained cells (three shown) with aligned receptive-field centers and a single transient cell. (b) The revised form of gating is one to one, between transient and sustained cells. Subsequently, signals from spatially aligned gating cells (indicated by three small arrows) are pooled to form a single Level 4 signal. (c) A motion oriented-contrast on-cell formed by oriented pooling of transient on-cell activations, such that the transient cells that input to the white region of the oriented receptive field are excitatory, whereas those inputting to the black region are inhibitory, thereby creating a tendency to respond to rightward motion. (d) A motion oriented-contrast off-cell that also is sensitive to rightward motion but uses transient off-cells to generate the appropriate oriented motion contrast signal.

rangement. We therefore modify the MOC filter, as in Figure 22b, by computing all sustained-transient combinations at each position before combining their outputs via directionally sensitive short-range spatial filters.

It should be noted that parallel computation of oriented sustained cells and transient cells that are subsequently gated is not the only way to compute the desired sustained-transient properties. An alternative is to group the outputs from transient on-cells into oriented receptive fields, thereby generating a motion oriented-contrast on-cell (Figure 22c). In such a cell, transient on-cells excite one half of the oriented receptive field and inhibit the other half of the oriented receptive field. A single serial stream of transient-then-oriented processing here replaces parallel processing by transient cells and sustained cells followed by sustained-transient gating. The output of such a serially generated cell would be positive only if the *oriented temporal contrast* that the cell detected exceeds its rectified output threshold. The duration of transient on-cell activation defines a time window wherein an oriented temporal contrast may be detected. Sufficiently many on-cells from one half of the receptive field would need to be simultaneously more active than those on the other half of the receptive field for the cell to fire. A separate population of oriented off-detectors may likewise be synthesized from combining outputs of transient cells into an oriented receptive field (Figure 22d). Such cells have many of the properties of Barlow and Levick (1965) or Reichardt detectors (Reichardt, 1957; van Santen & Sperling, 1985), even though they do not compute explicit delays between separate input channels.

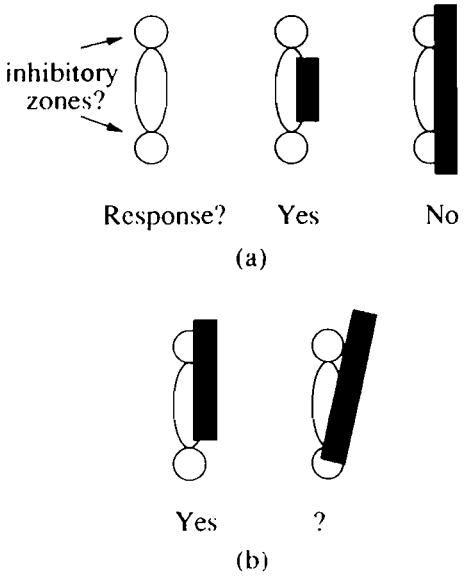


Figure 23. If the inhibitory end zones of endstopped simple cells were laid out as caricatured in part (a), a cell would be unable to distinguish the termination of a line of its preferred orientation from the continuation of a line of an orientation slightly different from the cell's preferred orientation. In the resulting situation, diagrammed in (b), disinhibition of the (top) inhibitory zone occurs while the strength of response from the central region decreases.

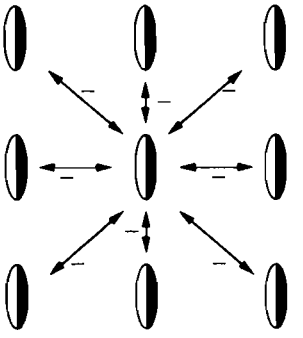


Figure 24. Spatially isotropic inhibition among like-oriented simple cells can generate functional enhancement of activity of oriented receptive fields near line ends or corners, often referred to as *endstopping*.

9. Endstopping: Generation of a Terminator or Corner Advantage in Motion Signals

Another 2-D MOC filter refinement involves an *endstopping* operation. The need for this was illustrated in our discussion of the barberpole illusion in Section 2. There, we noted that motion signals near terminators or corners tend to be better indicators of object motion than do signals generated from a relatively straight interior of a contour. In order for a motion signal at a terminator or corner to be effective, however, it must somehow be translated into a relatively large signal strength, since the region of ambiguous interior motion signals is often larger than the region of unambiguous terminator or corner motion signals.

We suggest that one source of this enhancement involves endstopping the sustained cells and/or the transient cells of the MOC filter. Many simple cells, identified with model cells at Levels 2 and/or 4 of the MOC filter, exhibit endstopping (Dreher, 1972). Endstopping is often informally described as consisting of inhibitory zones that occur along the major axis of orientationally tuned cells, beyond the contrast summing region, as indicated in Figure 23a. If this were all that was involved, then the end of a contrast edge that is oblique to a cell's preferred direction and that partially overlaps the cell could escape the inhibitory zone and stimulate the cell more than could a corresponding edge of the preferred direction (Figure 23b). This problem arises because the inhibition is anisotropic—that is, it occurs only from the directions aligned with the cell's orientational axis.

The anatomical substrate for such anisotropy would be more difficult to implement than would a scheme of isotropic inhibition among orientationally tuned cells. For such an isotropic inhibitory scheme (Figure 24), the *observed* inhibitory zones for stimulation with edge stimuli would still appear only at the ends of the cell's receptive field, owing to the interaction between isotropic inhibition and oriented receptive fields.

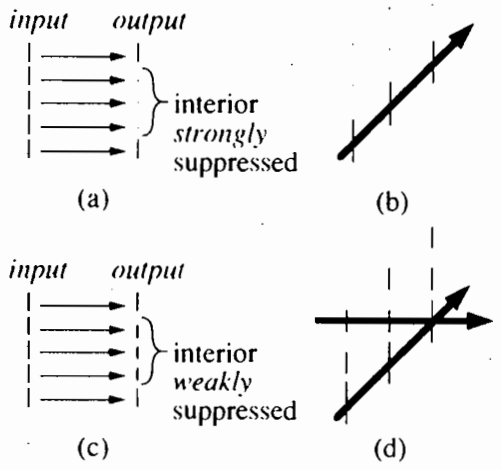


Figure 25. The enhancement of motion signals at line ends and corners can be strong or weak. (a) Strong inhibition can kill interior signals, making the problem of motion segmentation easy. (b) The surviving pools of activity at line ends can be directly tracked. (c) If endstopping is mild enough, however, locally preferred motion directions (perpendicular to edge contrast) survive, as shown in (d).

As illustrated in Figure 25a, strong endstopping can attenuate signals from all but the ends of a contour. Strong endstopping reduces the problem of determining motion direction to one of tracking an isolated region of activity in an upward diagonal direction, as in Figure 25b. Weak endstopping, as in Figure 25c, can partially attenuate interior signals, relative to signals at the ends. If endstopping is too weak, however, surviving signals indicating the "locally preferred" rightward horizontal direction can at least partially confound the computation of an upward diagonal object motion, as illustrated in Figure 25d. We note that Marshall (1990) has also invoked endstopping in his explanation of the barberpole illusion, though in a manner different from that outlined here.

10. Directional Competition and Boundary Completion

The inward cooperative propagation by bipole cells of motion signals to locations between the strong end reactions can help to overcome this problem as it completes the motion signal along the entire contour (Figure 13). The cooperative feedback from the strong end reactions also leads to inhibition of other directional signals via short-range competition between directions at positions along the contour, as occurs during motion capture (see Section 2). Thus, the type of CC loop interactions among hypercomplex cells and bipole cells that help to select the globally preferred orientations in the static BCS, as in Figure 10, also help to select the globally preferred directions in the motion BCS, with the caveat that "orientation" computations are replaced by "direction" computations.

11. Short-Range Versus Long-Range Reactions to Contrast Changes in the Static BCS and Motion BCS

The several BCS processing stages from sustained and transient cells to simple cells, complex cells, hypercomplex cells, and bipole cells form a type of highly specialized pyramid in which an increase in the scale of spatial interaction is combined with a reduction in sensitivity to direction of contrast. For both the static BCS and the motion BCS, these processing stages achieve an *output* signal of orientation or motion direction, respectively, that is insensitive to direction of contrast. These operations do not imply that the systems are insensitive to changes of direction of contrast in the *input*. As pointed out by Prazdny (1984), reversing the contrast relative to a neutral background of one set of dots in a Glass pattern can weaken, or even annihilate, the grouping percept (see Figures 26a and 26b). The paradoxical nature of this result can best be appreciated by juxtaposing it with the striking illusory groupings that can be sustained between the like directions of contrast (Figure 26c) of the Kanizsa square, as well as the opposite directions of contrast (Figure 26d) of the "reverse contrast Kanizsa square" (Cohen & Grossberg, 1984; Grossberg & Mingolla, 1985b; Prazdny, 1984; Shapley & Gordon, 1985). Within the static BCS, this apparent paradox has been resolved as follows (Grossberg & Mingolla, 1985b): The weak grouping percept in response to a reverse-contrast Glass pattern can be traced to spatially short-range interactions within the simple cell receptive fields of the SOC filter, which are sensitive to direction of contrast (Figure 2). The strong grouping percept in response to a reverse-contrast Kanizsa figure is due to spatially long-range interactions within the CC loop, whose hypercomplex cells and bipole cells are capable of combining opposite directions of contrast.

A similar pattern of sensitivity to direction of contrast at short range and insensitivity to direction of contrast at long range also occurs in motion perception. For short-range motion, reversal of dot contrast between frames abolishes a coherent motion percept, whereas long-range apparent motion can occur using displays whose contrast with respect to the background reverses between frames (Anstis & Mather, 1985).

12. 2-D MOC Filter Equations

We have found that several closely related computational realizations of the above heuristics can generate motion fields capable of distinguishing between contour orientation and contour direction of motion. We will here define a model that separates computations involving receptive fields with opposite directions of contrast until these are merged by the long-range Gaussian filter. This is the strategy followed by the Grossberg-Rudd model (Section 6). Analogous properties have been simulated using a variant of this model in which receptive fields with

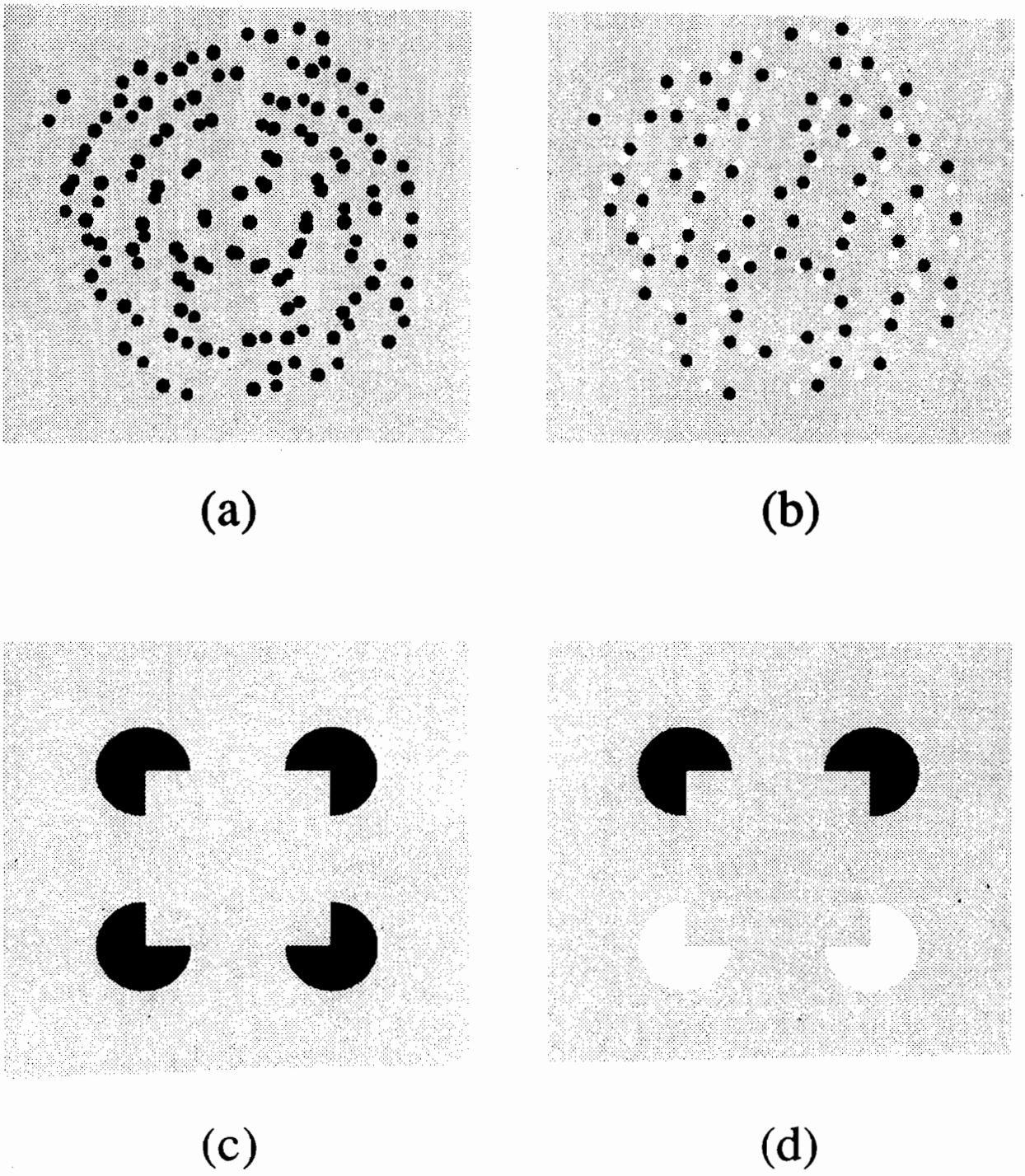


Figure 26. The subtleties of the interaction of spatial scale and direction of contrast are revealed by the juxtaposition of two classical visual phenomena. Parts (a) and (b) illustrate how the formation of Glass patterns is destroyed by reversing the contrast of one field of dots (from "On the perception of glass patterns" by K. Prazdny, 1984, *Perception*, 13, 469-478. Copyright 1984 by Pion, London. Reprinted by permission). The reversal of contrast of inducers in the Kanizsa square configuration from (c) to (d) does not significantly weaken illusory contour formation.

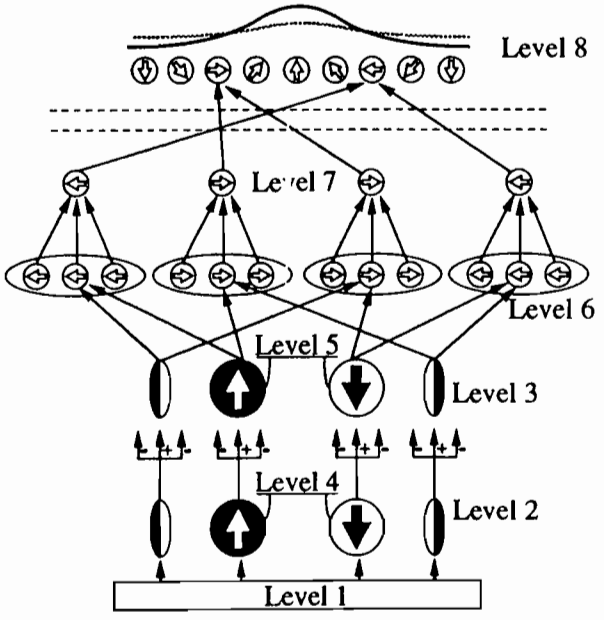


Figure 27. The 2-D MOC filter embodies a number of variations and extensions of the architecture of the 1-D MOC filter of Figure 14. See text for details.

opposite directions of contrast are merged at the short-range spatial filter stage (see Section 14). Other variations of the basic computational strategy also work, and they might be used by different species in vivo. The description that follows is keyed to Figure 27.

Level 1: Stimulus representation. Let $I_{pq}(t)$ denote the intensity of a time-varying image input at position (p,q) and time t.

Level 2: Oriented sustained receptive fields. Let the output J_{ijk} of a receptive field centered at position (i,j) with orientation k be defined by

$$J_{ijk} = \sum_{pq} A_{ijpq}^{(k)} I_{pq}, \tag{1}$$

where $A_{ijpq}^{(k)}$ defines the value of a Gabor kernel at position (p,q) that is centered at position (i,j) with orientation k. The Gabor kernel is

$$A_{ijpq}^{(k)} = \alpha_A \exp\left\{-\beta_A\left[\gamma_A\left(B_{ijpq}^{(k)}\right)^2 + \left(C_{ijpq}^{(k)}\right)^2\right]\right\} \sin\left(\delta_A C_{ijpq}^{(k)}\right), \tag{2}$$

where α_A is a constant that scales the Gabor amplitude, β_A scales the size of the kernel's Gaussian envelope, γ_A specifies the degree of elongation of the receptive field in the preferred orientation, and δ_A controls the frequency of the kernel's sinusoidal modulation. Terms

$$B_{ijpq}^{(k)} = (p-i)\cos\left(\frac{2\pi k}{K}\right) - (q-j)\sin\left(\frac{2\pi k}{K}\right) \tag{3}$$

and

$$C_{ijpq}^{(k)} = (p-i)\sin\left(\frac{2\pi k}{K}\right) + (q-j)\cos\left(\frac{2\pi k}{K}\right) \tag{4}$$

describe the effect of shifting a receptive field centered at position (0,0) to position (i,j), rotating it to orientation k, and evaluating it at position (p,q), given that K is the total number of orientations.

Level 3: Endstopped sustained cells. A competition of like-oriented Gabor receptive fields across neighboring positions and orientations leads to sustained cell activations x_{ijk} that are stronger at line ends and corners, as in the equation

$$\frac{d}{dt}x_{ijk} = -\alpha_x x_{ijk} + (\beta_x - x_{ijk})J_{ijk} - (\gamma_x + x_{ijk}) \sum_{pq \in D_{ij}} J_{pqk}, \tag{5}$$

where D_{ij} is the set of all positions within some radius E of (i,j); that is,

$$D_{ij} = [(p,q) : (p-i)^2 + (q-j)^2 \le E^2]. \tag{6}$$

Level 4: Unoriented transient receptive fields. The transient on-cell $y_{ij}^{(+)}$ responds to increments in the total input to a region F_{ij} surrounding (i,j), as in

$$y_{ij}^{(+)} = \left(\frac{d}{dt}\sum_{pq \in F_{ij}} I_{pq}\right)^+, \tag{7}$$

where F_{ij} is the set of all positions within some radius G of (i,j); that is,

$$F_{ij} = [(p,q) : (p-i)^2 + (q-j)^2 \le G^2] \tag{8}$$

for some radius G. Likewise, transient off-cell $y_{ij}^{(-)}$ responds to decrement in the total input to a region surrounding (i,j), as in

$$y_{ij}^{(-)} = \left(\frac{d}{dt}\sum_{pq \in F_{ij}} I_{pq}\right)^+, \tag{9}$$

where region F_{ij} is defined above.

Level 5: Center-surround transient cells. In this version of the model, lateral inhibition among transient cells, analogous to the endstopping operation among simple cells, is implemented. The center-surround transient cell activations $Y_{ij}^{(+)}$ and $Y_{ij}^{(-)}$ obey the same type of equation (Equation 5) as the endstopped sustained cells; namely,

$$\frac{d}{dt}Y_{ij}^{(+)} = -\alpha_y Y_{ij}^{(+)} + (\beta_y - Y_{ij}^{(+)})\, y_{ij}^{(+)} - (\gamma_y + Y_{ij}^{(+)}) \sum_{pq \in H_{ij}} y_{pq}^{(+)} \tag{10}$$

and

$$\frac{d}{dt}Y_{ij}^{(-)} = -\alpha_y Y_{ij}^{(-)} + (\beta_y - Y_{ij}^{(-)})\, y_{ij}^{(-)} - (\gamma_y + Y_{ij}^{(-)}) \sum_{pq \in H_{ij}} y_{pq}^{(-)}, \tag{11}$$

where

$$H_{ij} = [(p,q) : (p-i)^2 + (q-j)^2 \le L^2]. \tag{12}$$

The transformation from Level 4 signals to Level 5 signals results in enhanced activity at line ends and corners.

Level 6: Sustained-transient simple cells. As in Level 4 of the Grossberg-Rudd MOC filter (Section 6), multiplying outputs from sustained cells and transient cells starts to compute a local measure of motion direction. In particular, sustained cell activations that are sensitive to opposite directions of contrast give rise to activations that are sensitive to the same direction of motion by being multiplied with transient on-cells and off-cells, respectively. For example, the sustained-transient interaction

$$M_{ijm}^{(+)} = [x_{ijk}]^+ Y_{ij}^{(+)}, \tag{13}$$

where

$$m = \left(k+\frac{K}{4}\right)(mod\ K), \tag{14}$$

with K an integer multiple of 4, and the sustained-transient interaction

$$M_{ijm}^{(-)} = [x_{ij\hat{k}}]^+ Y_{ij}^{(-)}, \tag{15}$$

where

$$\hat{k} = \left(k+\frac{3K}{4}\right)(mod\ K), \tag{16}$$

are both maximally sensitive to oriented contours of orientation k moving in direction m, even though the activations x_{ijk} and $x_{ij\hat{k}}$ sense opposite directions of contrast in an orientation that is perpendicular to m.

These quantities are time-averaged by the cells at which the sustained-transient interactions occur. Let $z_{ijm}^{(+)}$ and $z_{ijm}^{(-)}$ compute the activities of the sustained-transient cells that receive the inputs $M_{ijm}^{(+)}$ and $M_{ijm}^{(-)}$, respectively. Then

$$\frac{d}{dt}z_{ijm}^{(+)} = -\alpha_z z_{ijm}^{(+)} + M_{ijm}^{(+)} \tag{17}$$

and

$$\frac{d}{dt}z_{ijm}^{(-)} = -\alpha_z z_{ijm}^{(-)} + M_{ijm}^{(-)}. \tag{18}$$

For simplicity, the time-averaging in Equations 17 and 18 was computed in discrete time steps using the equations

$$z_{ijm}^{(+)}(t) = \sum_{\tau=t-T}^{t} M_{ijm}^{(+)}(\tau)\ \zeta^{t-\tau}, \tag{19}$$

and

$$z_{ijm}^{(-)}(t) = \sum_{\tau=t-T}^{t} M_{ijm}^{(-)}(\tau)\zeta^{t-\tau}, \tag{20}$$

where $0<\zeta<1$ and the number of time steps T was chosen large enough to provide a good approximation to Equations 17 and 18, respectively. Comparison of the continuous and discrete time equations shows that $\zeta = \exp(-\alpha_z)$.

Level 7: Short-range spatial filter—Pooling of orientation detectors into local direction detectors. Each activation $z_{ijm}^{(+)}$ and $z_{ijm}^{(-)}$ is derived from an oriented receptive field whose orientation k is perpendicular to m. The next operation pools activations that are sensitive to the same direction of contrast using a short-range spatial filter. The spatial filter pools activations that lie along motion trajectories with different preferred directions of motion m. Activations $z_{ijn}^{(+)}$ or $z_{ijn}^{(-)}$ are accepted whose directions n are not too different from the trajectory direction m. This pooling operation exploits the relative advantage at line ends and corners achieved by endstopping to compute trajectory signals that begin to disambiguate a contour's direction of motion from its orientation.

The short-range spatial kernel $N_{ijpq}^{(m,n)}$ favors a motion trajectory with direction-of-motion m that is centered at position (i,j). It separately pools activations $z_{pqn}^{(+)}$ or $z_{pqn}^{(-)}$ with a weight that depends upon how close n is to m and (p,q) is to (i,j), as in

$$u_{ijm}^{(+)} = \sum_{pqn} N_{ijpq}^{(m,n)} z_{pqn}^{(+)} \tag{21}$$

and

$$u_{ijm}^{(-)} = \sum_{pqn} N_{ijpq}^{(m,n)} z_{pqn}^{(-)}, \tag{22}$$

where

$$N_{ijpq}^{(m,n)} = P_{ijpq}^{(m)}[\cos[2\pi(m-n)K^{-1}]]^+ \tag{23}$$

and

$$P_{ijpq}^{(m)} = \alpha_p \exp\{-\beta_p[\gamma_p(Q_{ijpq}^{(m)})^2 + (R_{ijpq}^{(m)})^2]\}\cos(\delta_p R_{ijpq}^{(m)}). \tag{24}$$

In Equation 24, just as in Equation 2, α_p is a constant that scales the kernel amplitude, β_p determines the size of the kernel envelope, γ_p specifies the degree of kernel elongation in the preferred direction m, and δ_p specifies the frequency of the kernel's cosine modulation. The values

$$Q_{ijpq}^{(m)} = (p-i)\cos\left(\frac{2\pi m}{K}\right) - (q-j)\sin\left(\frac{2\pi m}{K}\right) \tag{25}$$

and

$$R_{ijpq}^{(m)} = (p-i)\sin\left(\frac{2\pi m}{K}\right) + (q-j)\cos\left(\frac{2\pi m}{K}\right) \tag{26}$$

describe the effect of shifting a receptive field centered at position (0,0) to position (i,j), rotating it to point in direction m, and evaluating it at position (p,q).

The cell activations $u_{ijm}^{(+)}$ and $u_{ijm}^{(-)}$ are computed using the algebraic Equations 21 and 22 rather than the differential Equations 17 and 18 used to compute $z_{ijm}^{(+)}$ and $z_{ijm}^{(-)}$, because $u_{ijm}^{(+)}$ and $u_{ijm}^{(-)}$ are assumed to respond much more quickly to their inputs than $z_{ijm}^{(+)}$ and $z_{ijm}^{(-)}$. In particular, the kernel $N_{ijpq}^{(m,n)}$ accumulates evidence for motion in direction m along a trajectory through position (i,j). The persistence of $z_{pqn}^{(+)}$ and $z_{pqn}^{(-)}$ activations as they temporally decay, together with the spatial anisotropy of the kernel, begin to overcome the uncertainties that arise from the aperture problem. Kernel anisotropy is hypothesized to arise, at least in part, in response to experiences with the trajectories of moving contours during an early phase of brain development. Such trajectories are, with high probability, approximately straight over sufficiently small regions.

The cell activations $u_{ijm}^{(+)}$ and $u_{ijm}^{(-)}$ both detect an estimate of direction of motion m at position (i,j), but they are sensitive to opposite directions of contrast. The long-

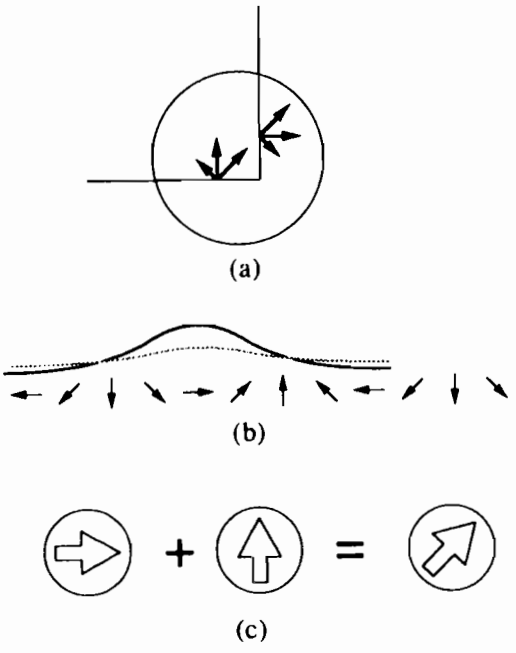


Figure 28. (a) At a corner, the Gaussian filter combines signals of several directions. (The circle indicates the spatial domain of the Gaussian filter.) (b) The resolution of motion signals of many directions can be accomplished by a shunting "center-surround" competition with normalization among direction signals at a given location. Such a competition chooses the globally most consistent direction at a location. The solid line indicates the "on-center," and the dashed line indicates the "off-surround" of interaction weights across directions. (c) An input distribution with peaks at the rightward and upward directions is transformed into a single peaked distribution pointing upward and rightward by such a network.

range spatial filter combines signals that are sensitive to opposite directions of contrast, but the same direction of motion, to compute a more accurate estimate of motion direction, as indicated in Figure 21. The output signals $U_{ijm}^{(+)}$ and $U_{ijm}^{(-)}$ from Level 7 to this long-range filter are rectified versions of the Level 7 activations, as in

$$U_{ijm}^{(+)} = [u_{ijm}^{(+)} - \eta]^+ \tag{27}$$

and

$$U_{ijm}^{(-)} = [u_{ijm}^{(-)} - \eta]^+. \tag{28}$$

Level 8: Long-range spatial filter and directional competition. The trajectory responses $U_{ijm}^{(+)}$ and $U_{ijm}^{(-)}$ are next combined across space and direction of contrast to generate a response v_{ijm} that is sensitive to the consensus within a region surrounding position (i, j) of how much evidence there exists for motion in direction m. Thus, we let

$$v_{ijm} = \sum_{pq \in W_{ij}} \left(U_{pqm}^{(+)} + U_{pqm}^{(-)}\right) \exp\{-\alpha_v[(p-i)^2 + (q-j)^2]\}, \tag{29}$$

where

$$W_{ij} = [(p,q) : (p-i)^2 + (q-j)^2 \leq X^2]. \tag{30}$$

The long-range filter is indicated by the parallel dashed lines near the top of Figure 27.

The output of the long-range filter undergoes a final competition to choose a consensual direction among all signals that have been grouped at a spatial location (Figure 28). This competition is in the form of a center-surround organization in "direction space," as given by

$$\frac{d}{dt} w_{ijm} = -\alpha_w w_{ijm} + (\beta_w - w_{ijm}) \sum_n v_{ijn} \gamma_{nm} - (\delta_w + w_{ijm}) \sum_n v_{ijn} \epsilon_{nm}, \tag{31}$$

where

$$\gamma_{nm} = \gamma \exp[-\mu_w (m-n)^2] \tag{32}$$

and

$$\epsilon_{nm} = \epsilon \exp[-\nu_w (m-n)^2]. \tag{33}$$

The competition stage at Level 8 is not necessary to generate consistent direction signals for simple, noise-free inputs. Snowden (1989, 1990) has reported psychophysical and physiological evidence for such shunting inhibition among directional signals, for situations in which conflicting signals must be resolved. Williams and Phillips (1987) and Watson and Ahumada (1985) have proposed a similar direction-averaging mechanism.

13. Computer Simulations: Distinguishing Motion Direction from Boundary Orientation

Computer simulations were carried out on simple moving forms to illustrate how the MOC filter can compute

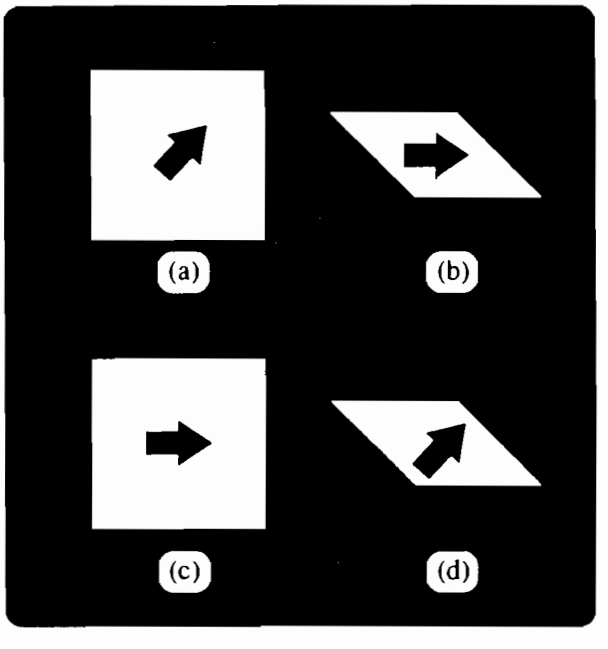


Figure 29. Simulations of the response of the 2-D MOC filter to the motions of two simple figures are shown in subsequent figures. These figures are (a) a square moving diagonally upward and to the right, (b) a parallelogram moving horizontally rightward, (c) a square moving horizontally rightward, and (d) a parallelogram moving diagonally up and to the right.

Table 1
Graphics Conventions

Symbol	Figures	Level	Cell Type
	31a, 33a	2	Oriented sustained
	31b, 33b	3	Endstopped oriented sustained
	31c, 33c	4	Unoriented transient (luminance decreasing)
	31c, 33c	4	Unoriented transient (luminance increasing)
	31d, 33d	5	Center-surround transient (luminance decreasing)
	31d, 33d	5	Center-surround transient (luminance increasing)
	31e, 33e	6	Sustained-transient simple (luminance decreasing)
	31e, 33e	6	Sustained-transient simple (luminance increasing)
	31f, 33f	6	Time-average of sustained-transient simple (luminance decreasing)
	31f, 33f	6	Time-average of sustained-transient simple (luminance increasing)
	31g, 33g	7	Short-range spatial filter (luminance decreasing)
	31g, 33g	7	Short-range spatial filter (luminance increasing)
	31h, 33h, 35, 37	8	Motion hypercomplex

an accurate measure of motion direction even if the bounding contours of the form have orientations that are not perpendicular to the direction of motion. Four illustrative moving forms are analyzed below: a square moving diagonally upward, a square moving to the right, a parallelogram moving diagonally upward, and a parallelogram moving to the right (Figure 29). For the square moving diagonally, the resultant activation pattern at each stage of the MOC filter is shown in order to aid the reader's intuition. The graphics conventions used are shown in Table 1.

Consider the case of a light square moving diagonally up and to the right on a dark background, as depicted in Figure 30. Figure 31a displays the response of the oriented sustained detectors of Level 2; the parameters controlling the scale of the detectors are such that only a single row or column is activated around the perimeter of the square, although several orientations are active at each position. Figure 31b displays the response of the endstopped, oriented sustained detectors of Level 3. Note the attenuated response along the interiors of the segments bounding the square and the enhanced response at corners. Figure 31c displays the response of increasing and decreasing transient detectors of Level 4. Note that the leading and trailing corners (upper right and lower left, respectively) have strong activity, because the axis they define is in the direction of motion, while the other two corners have attenuated activity. Figure 31d displays the result of contrast-enhancing competition of Level 5 transient detectors. Note that this competition occurs independently in the "increasing" and "decreasing" channels, permitting the greatest enhancement of activity at the corners where activity was weakest in Figure 31c.

Figure 31e displays the output of a field of Level 6 sustained-transient simple cells. Note that those signals along the top and right edges of the square are formed by gating sustained responses with luminance-increasing detectors, while those signals along the left and bottom contours are formed by gating with luminance-decreasing detectors. Figure 31f displays the temporal "smearing" of responses at Level 6. Note that at the upper left and lower right corners, signals formed by gating with luminance-increasing detectors (pointing upward or rightward, respectively) are superimposed upon signals formed by gating with luminance-decreasing detectors (pointing rightward or upward, respectively). Thresholded responses of the Level 7 short-

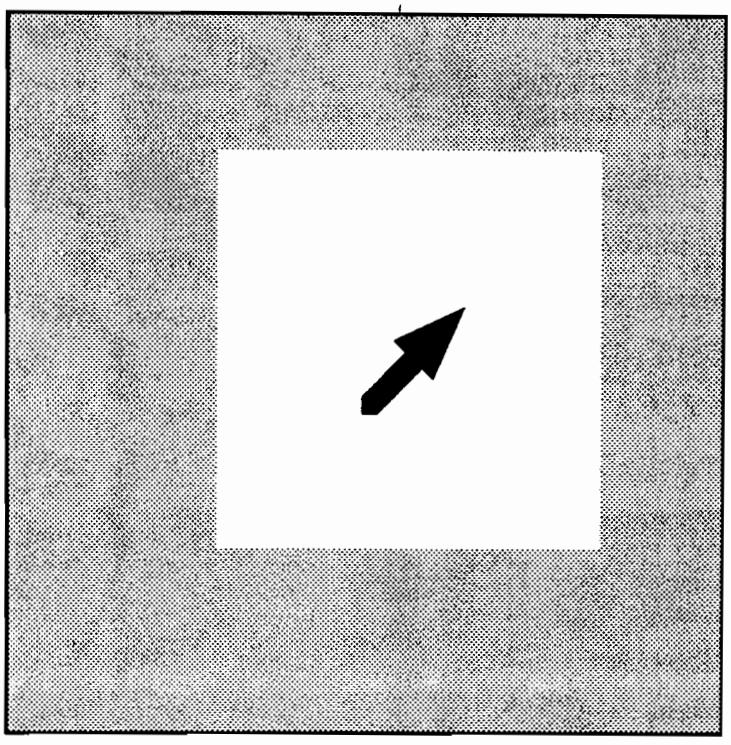


Figure 30. A representation of a frame from an input sequence presented to the 2-D MOC filter. A light square moves diagonally up and to the right against a dark background. The resolution of the image used in the simulation was 128 × 128 pixels; the distance between nearest receptive-field centers for network activities depicted in Figures 30-37 is 4 pixel units.

range spatial filters are displayed in Figure 31g. This filtering stage imposed a greater directional selectivity than did the previous stage, by virtue of the elongated form of its spatial filters. Level 8 is the first stage to combine information generated from both luminance-increasing and luminance-decreasing transient cells. The diagonal motion of the square is registered at all four corners, as shown in Figure 31h. The competition of Level 8 does not contribute materially to network function in this noise-free example, but it would be expected to play a more important role in noisier or more cluttered environments.

Figure 33 displays a sequence of transformations corresponding to those of Figure 31, but this time for the case of a parallelogram moving horizontally to the right, as depicted in Figure 32. The transformation from the output of oriented sustained cells of Level 2, shown in Figure 33a, to the output of endstopped cells of Level 3, shown in Figure 33b, changes the relative distribution of active orientations at corners. Likewise, the transformation from Level 4 transient cells to Level 5 contrast-enhanced transient cells (shown in Figures 33c and 33d, respectively) helps to create a stronger pool of activity at the ends of the diagonal lines, as was discussed with reference to the barberpole illusion in Section 2. The output of sustained-transient simple cells of Level 6, shown in Figure 33e, contains an assortment of active local direction signals. The spatial trail formed by the smearing of those signals in time, shown in Figure 33f, enables the short-range spatial filters of Level 7, shown in Figure 33g, to more accurately register the prevailing direction of motion. The output of Level 8, shown in Figure 33h, further smooths the signals over space, while sharpening their directional distribution. Note that the resulting pattern of activity is stronger at the 45° corners than at the 135° corners, and that the directional pattern at the latter is not symmetric about the horizontal. It may be that such effects underlie some of the anomalous motion percepts reported for motions of curves, such as those employed by Nakayama and Silverman (1988). Changing parameters in the simulation to include longer spatial filters at Level 7 would in any case reduce this effect and generate a more nearly horizontal pattern of responses.

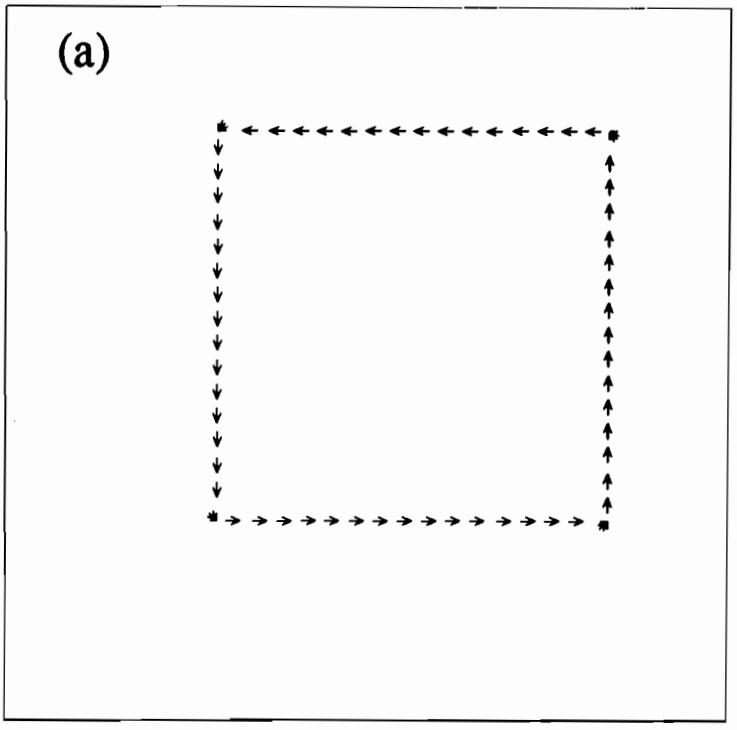
For comparison, we have included the response at Level 8 to the cases where a horizontal square moves rightward (Figures 34 and 35) and where a parallelogram moves diagonally (Figures 36 and 37).

14. A Related MOC Filter

Similar computational properties are found if the mask responses in Equation 1 are thresholded, as in

$$J_{ijk} = \left(\sum_{pq} A_{ijpq}^{(k)} I_{pq} - \epsilon_A\right)^+, \tag{34}$$

and the terms $M_{ijm}^{(+)}$ in Equation 13 and $M_{ijm}^{(-)}$ in Equa-



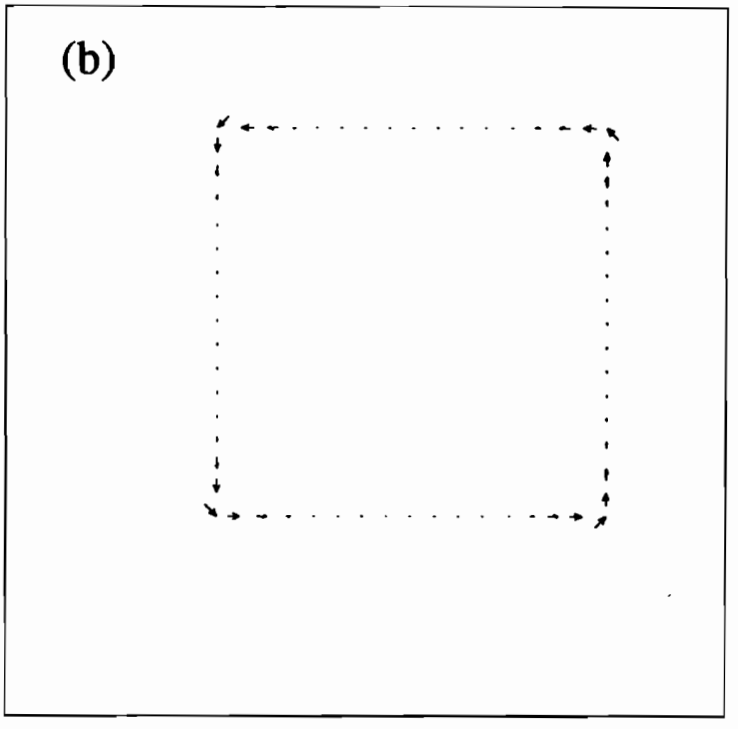
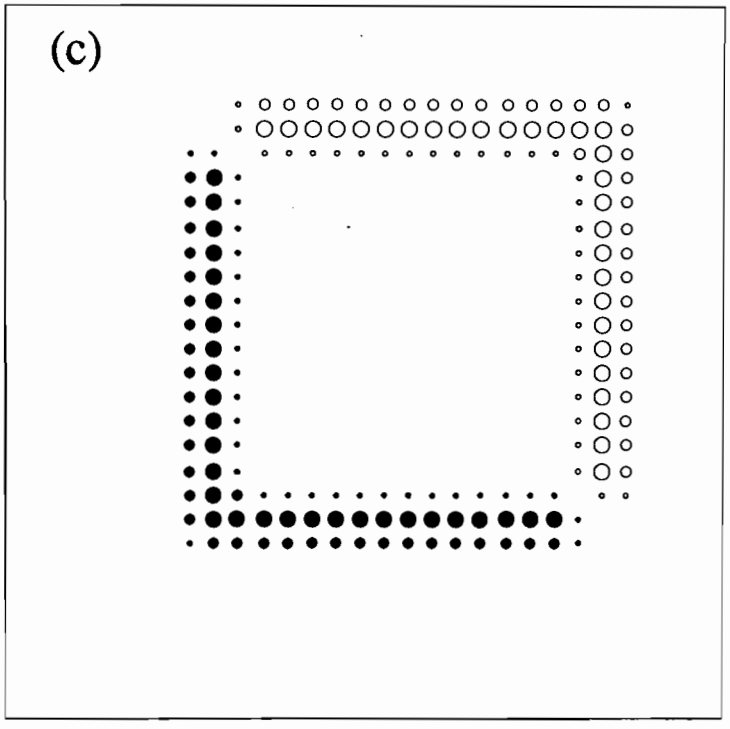
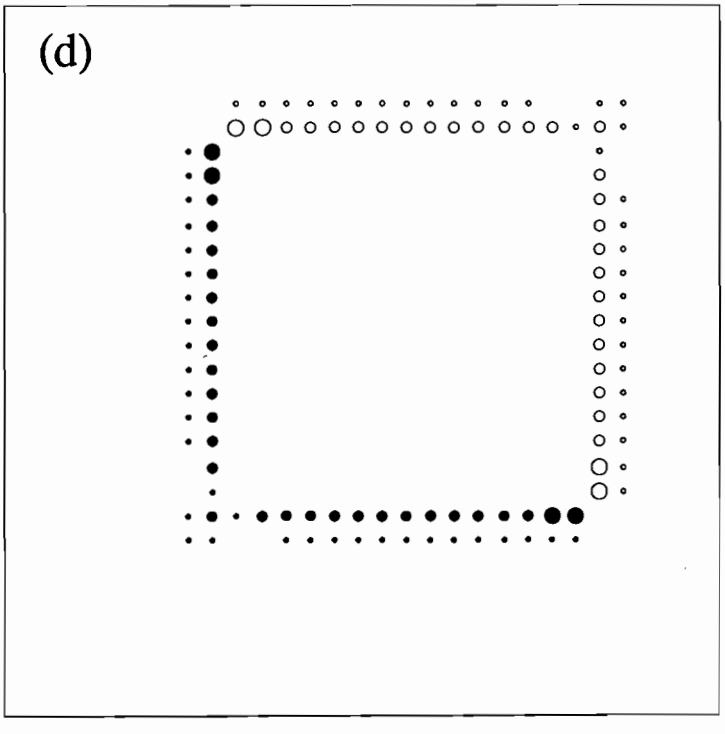
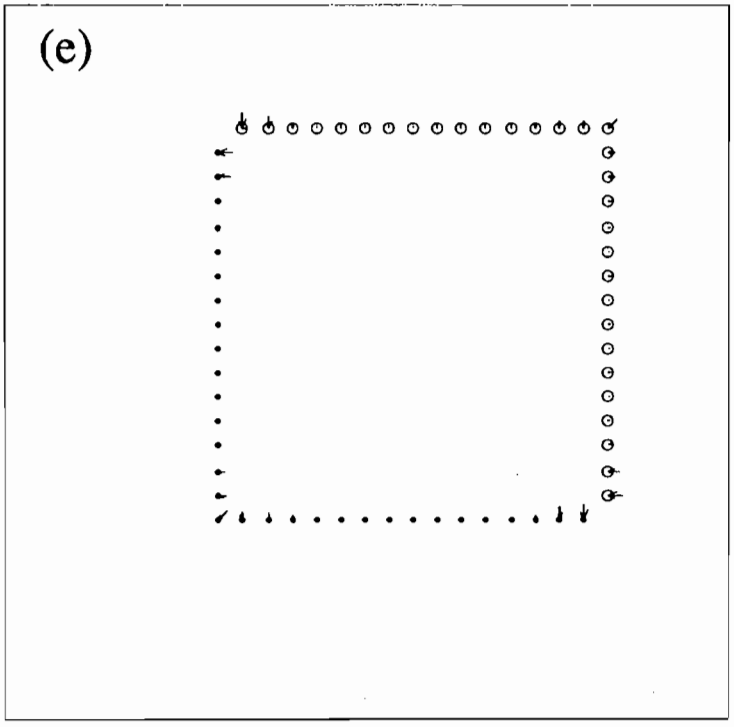
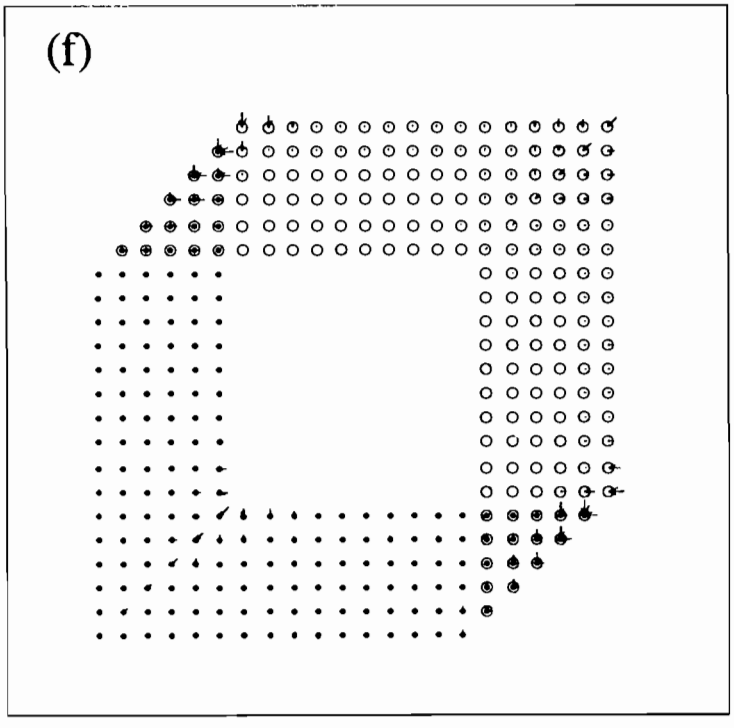


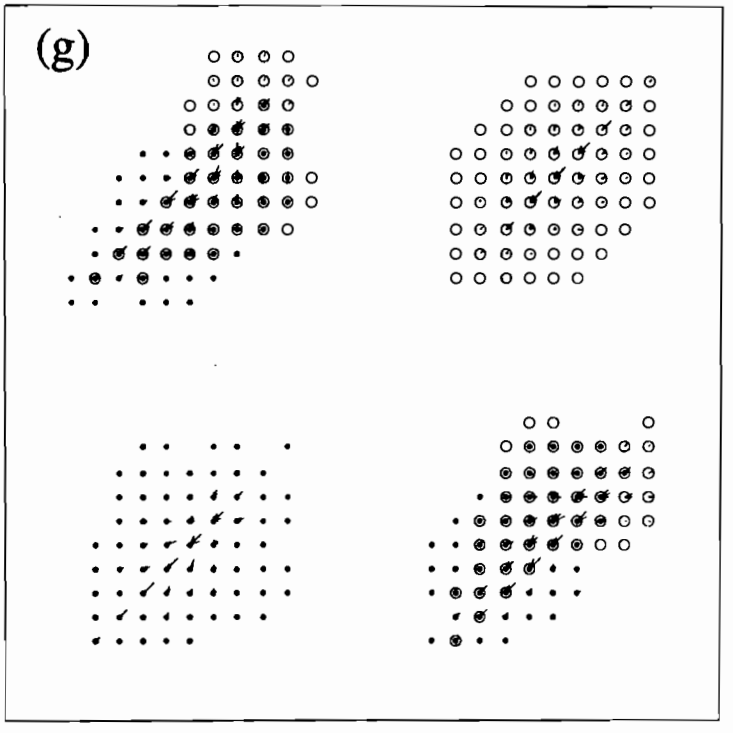
Figure 31. The output of the 2-D MOC filter of Figure 27 in response to the moving square of Figure 30. The parts of the figure (above and on the following three pages) show the responses of (a) the sustained cells of Level 2, (b) the endstopped sustained cells of Level 3, (c) the transient cells of Level 4, (d) the contrast-enhanced transient cells of Level 5, (e) the sustained-transient simple cells of Level 6, (f) the pattern formed by the temporal smearing of the responses of sustained-transient simple cells of Level 6, (g) the short-range spatial filter cells of Level 7, and (h) the competitive cells of Level 8.

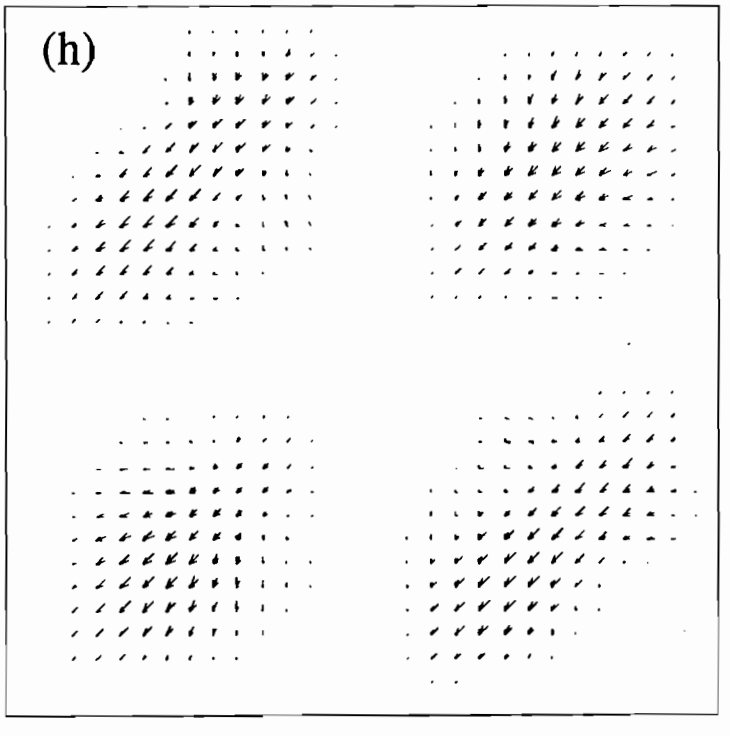

(c)


(d)


(e)


(f)


(g)


(h)

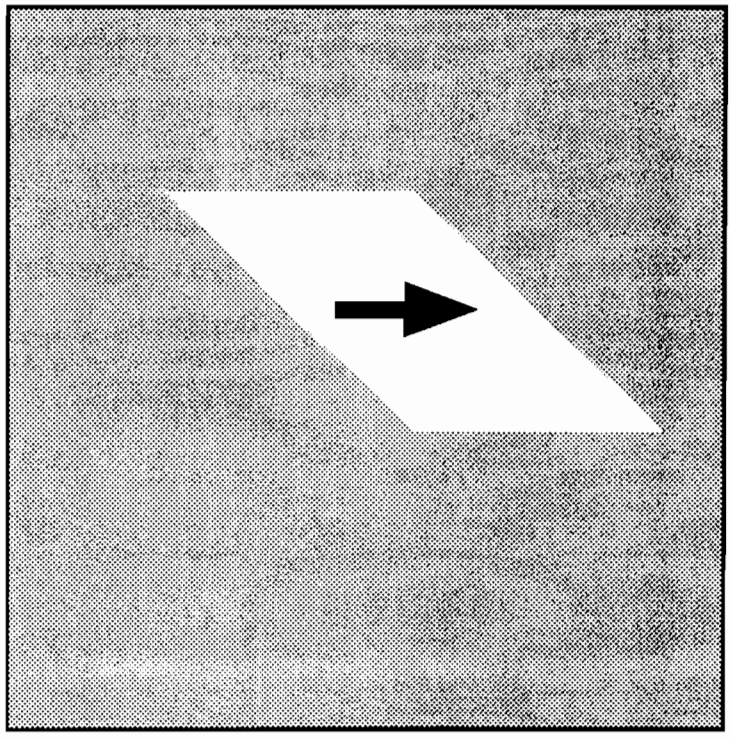


Figure 32. A representation of a frame from an input sequence presented to the 2-D MOC filter. A light parallelogram moves rightward against a dark background. The resolution of the image used in the simulation was 128 × 128 pixels; the distance between nearest receptive-field centers for network activities depicted in Figures 32-33 is 4 pixel units.

tion 15 are combined into a single time-averaged sustained- transient cell type z_{ijm}, instead of the separate streams Equations 19 and 20. Then

$$z_{ijm}(t) = \sum_{\tau=t-T}^{t} [M_{ijm}^{(+)}(\tau) + M_{ijm}^{(-)}(\tau)]\zeta^{t-\tau} \quad (35)$$

and

$$u_{ijm} = \sum_{pqn} \bar{N}_{ijpq}^{(m,n)} z_{pqn}, \quad (36)$$

where

$$\bar{N}_{ijpq}^{(m,n)} = P_{ijpq}^{(m)}\{[\cos(2\pi(m-n)K^{-1})]^+\}^p, \quad (37)$$

function $P_{ijpq}^{(m)}$ is defined in Equation 24, and parameter p in Equation 37 enables the smoothing across directions n to be made more or less broad within a 180° span. Using this approach, the threshold η in Equations 27 and 28 is not needed, because the masks (Equation 34) are already thresholded, and the functions $M_{ijm}^{(+)}$ and $M_{ijm}^{(-)}$ combine to generate a more central directional tendency at an early processing stage. Then

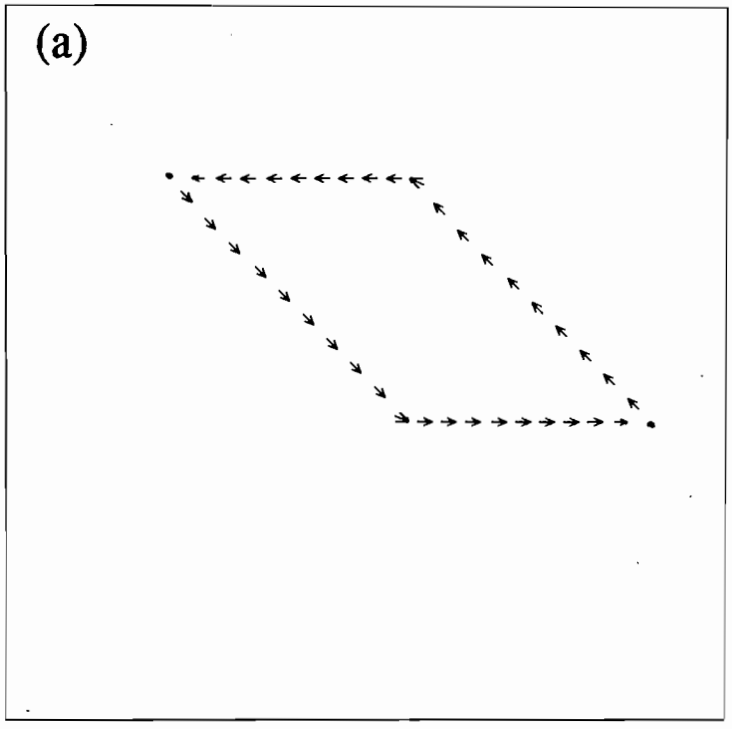
$$v_{ijm} = \sum_{pq \in W} u_{pqm} \exp\{-\alpha_v[(p-i)^2+(q-j)^2]\} \quad (38)$$

suffices instead of Equation 29.

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

The motion BCS circuitry suggests the types of processing stages, and their ordering, that are needed to effectively process moving images. As such, the circuit suggests experimental tests of the cell types and interactions in the V1→MT cortical processing stream. The theory also helps to understand (as described in Grossberg, 1991) 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). As noted in Section 7, 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 important in the selection of cortical cells that are tuned to the correct binocular disparities (von der Heydt, Hänny, & Dürsteler, 1981). The static BCS can carry out such precise oriented matches; the motion BCS cannot. As noted in Section 3, 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



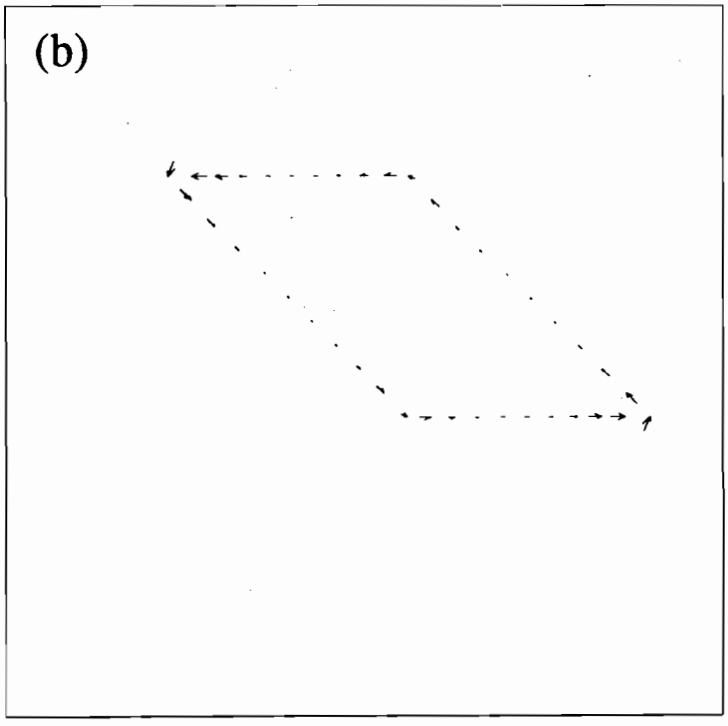
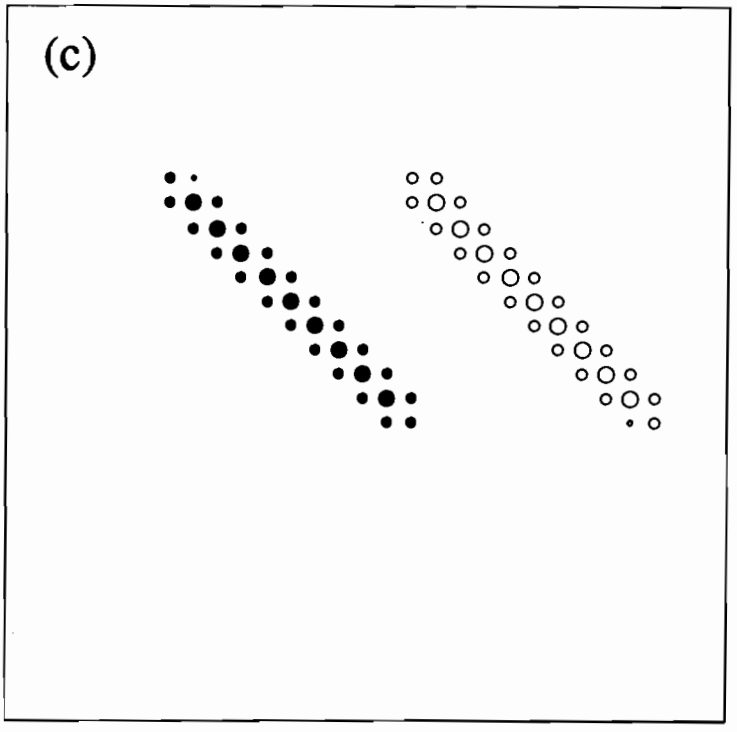
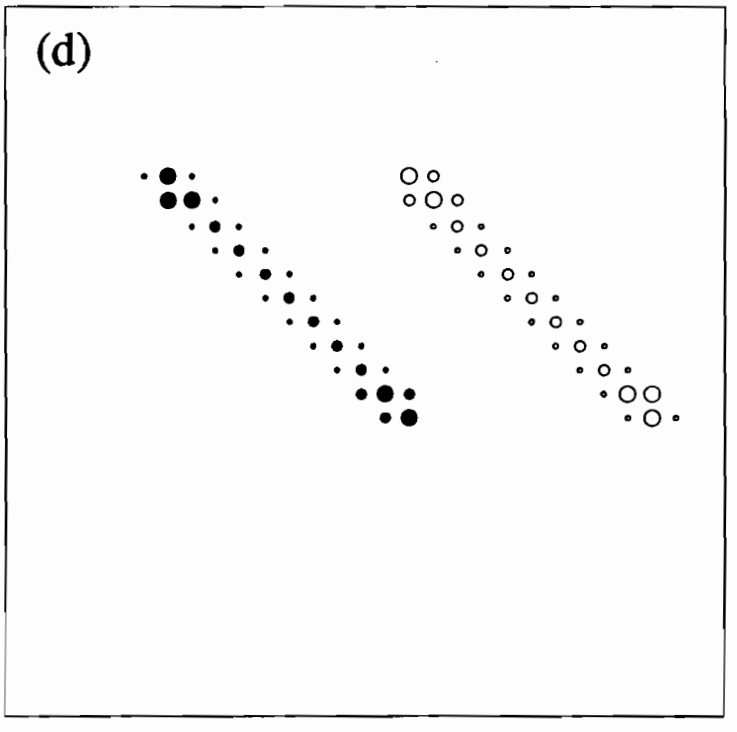
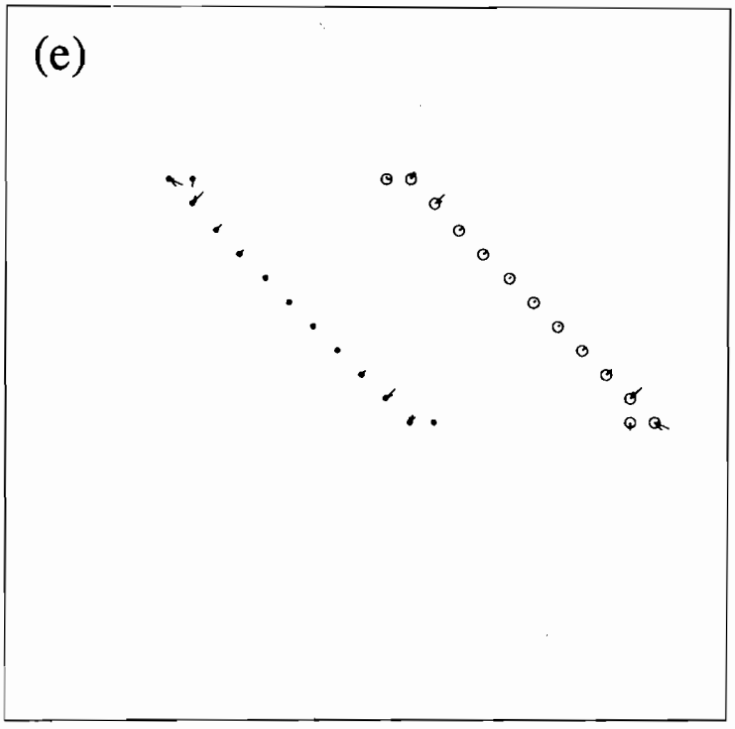
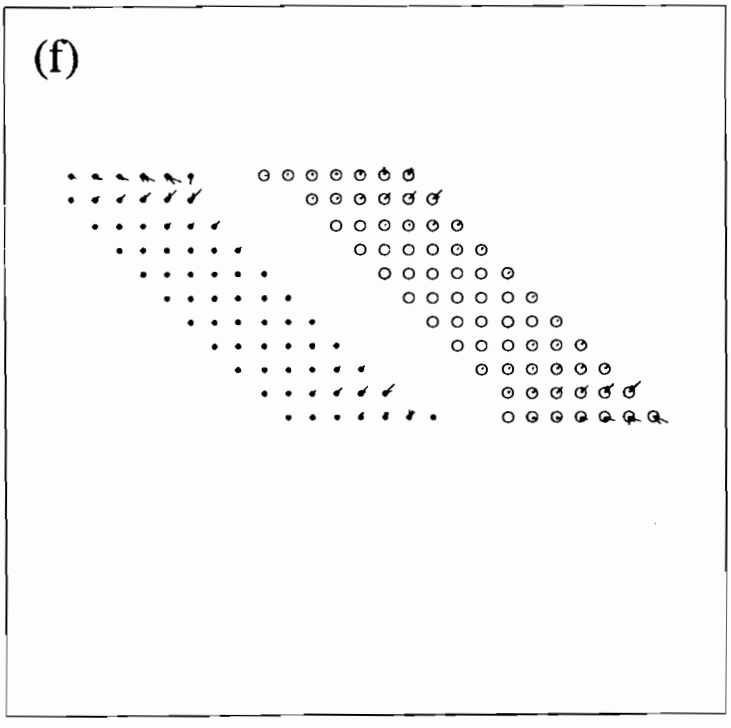


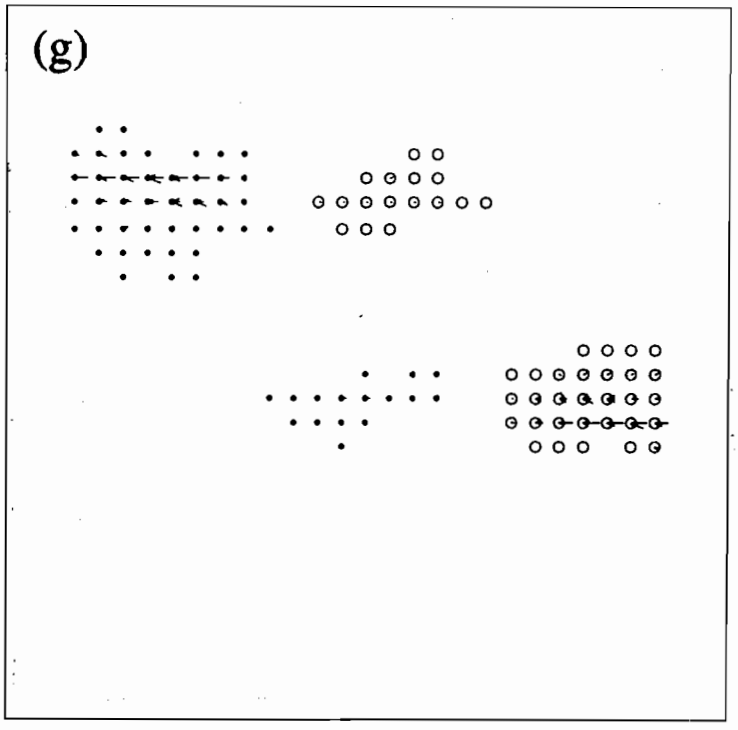
Figure 33. The output of the 2-D MOC filter of Figure 27 in response to the moving parallelogram of Figure 32. The parts of the figure (above and on the following three pages) show the responses of (a) the sustained cells of Level 2, (b) the endstopped sustained cells of Level 3, (c) the transient cells of Level 4, (d) the contrast-enhanced transient cells of Level 5, (e) the sustained-transient simple cells of Level 6, (f) the pattern formed by the temporal smearing of the responses of sustained-transient simple cells of Level 6, (g) the short-range spatial filter cells of Level 7, and (h) the competitive cells of Level 8.

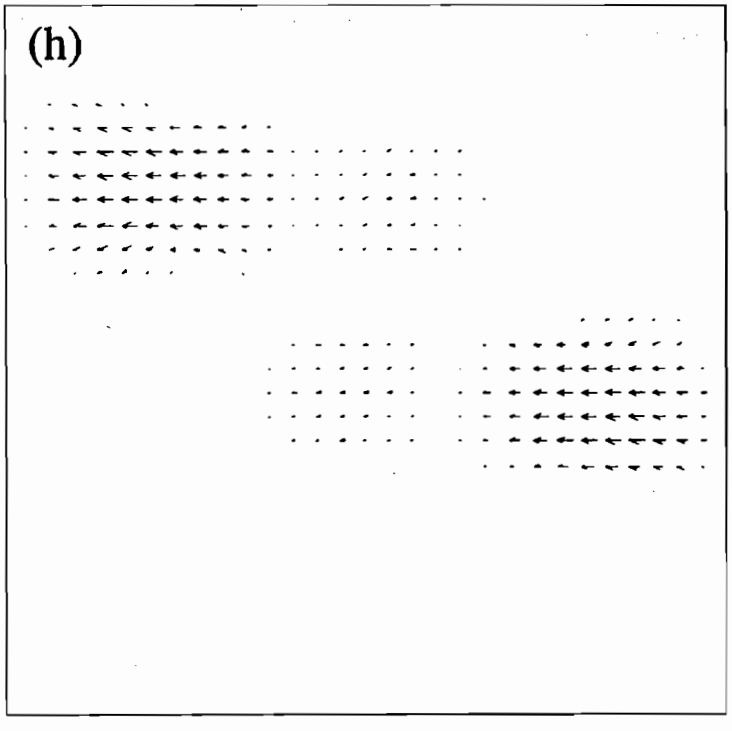

(c)


(d)


(e)


(f)


(g)


(h)

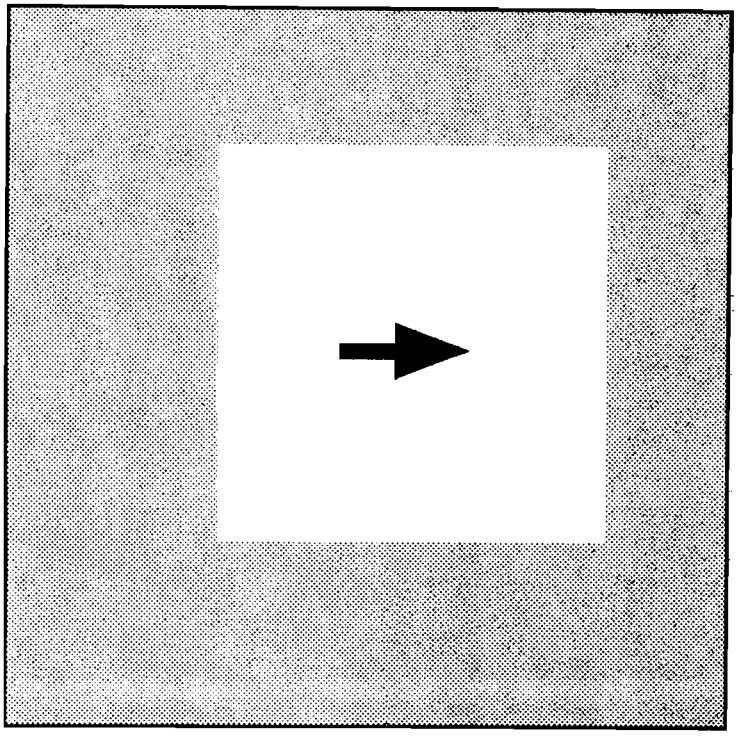


Figure 34. A representation of a frame from an input sequence presented to the 2-D MOC filter. A light square moves horizontally to the right against a dark background. The resolution of the image used in the simulation is 128×128 pixels; the distance between nearest receptive-field centers for network activities depicted in this figure is 4 pixel units.

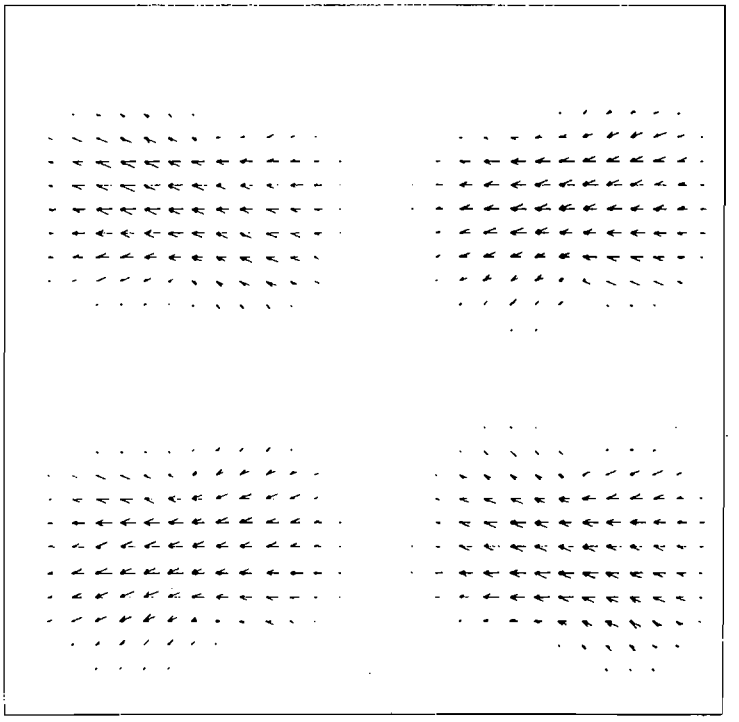


Figure 35. The output of the competitive cells of Level 8 of the 2-D MOC filter in response to the moving square of Figure 34.

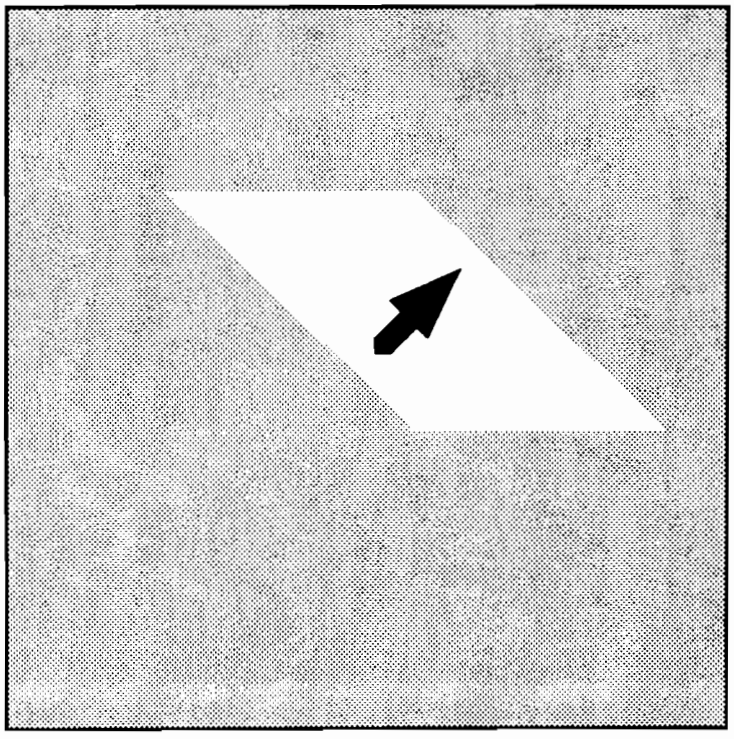


Figure 36. A representation of a frame from an input sequence presented to the 2-D MOC filter. A light parallelogram moves diagonally up and to the right against a dark background. The resolution of the image used in the simulation was 128×128 pixels; the distance between nearest receptive-field centers for network activities depicted in this figure is 4 pixel units.

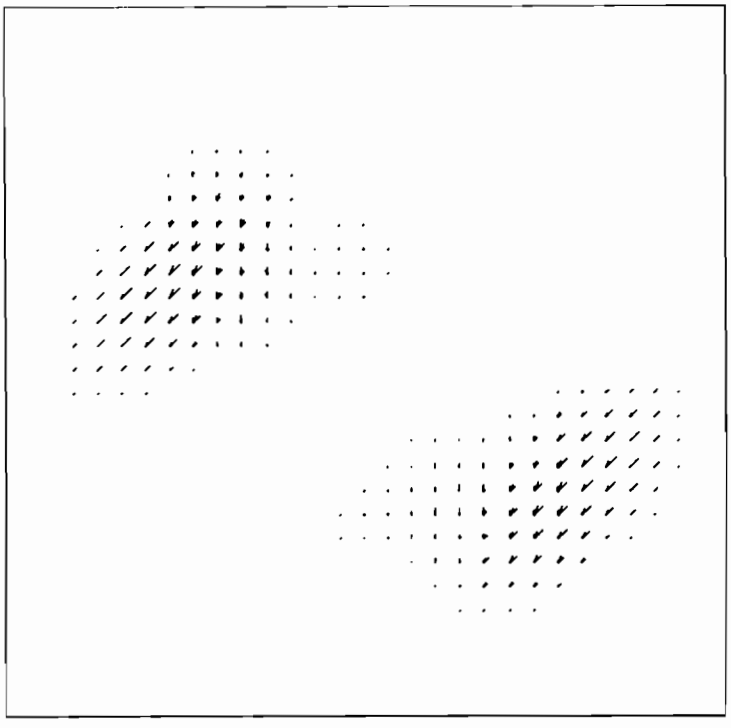


Figure 37. The output of the competitive cells of Level 8 of the 2-D MOC filter in response to the moving parallelogram of Figure 36.

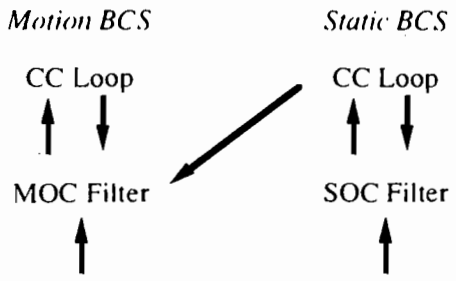


Figure 38. Model analog of V1 → V2 → MT pathway: Stereo-sensitive emergent segmentations from the static CC loop help to select the depthfully correct combinations of motion signals in the MOC filter. (Reprinted, with permission, from Grossberg, 1991.)

process is completed within the hypercomplex cells and bipole cells of the static BCS (Grossberg, 1987b; Grossberg & Marshall, 1989), and hence within the model analog of prestriate cortical area V2 (Figure 38). Thus, the existence of a pathway from V2 and/or V4 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 14, or correspondingly, Level 7 in Figure 27. Such a pathway would join like orientations within like spatial scales between the static BCS and the motion BCS. It could thereby enhance the activation within the motion BCS of those spatial scales and orientations that are binocularly fused within the static BCS.

16. Contextually Consistent and Coherent Motion Segmentation

In this article, we have described mechanisms for the processing of moving forms whereby the visual system frees itself from an excessive reliance on purely local computations that ignore the context of surrounding scenic data. Instead, competitive and cooperative interactions within a parallel and structured network with several scales of interaction help to choose and enhance those aspects of local data that contribute to consistent and coherent measures of object motion. Like its counterpart in the domain of quasi-static form perception, the motion BCS employs both unoriented and oriented computations, short-range and long-range interactions, and is in some places sensitive and at other places insensitive to direction of contrast. The precise configuration of levels, however, differs from that of the static BCS in order to efficiently generate coherent segmentations of moving forms. Indeed, the static BCS and motion BCS play dual roles in a larger architecture whose structure can be rationalized by principles of form-motion symmetry (Grossberg, 1991). Far from being independent modules or separate channels (Livingstone & Hubel, 1987), the principles and architectures of neither the static form perception system nor the motion form perception system can be fully understood without understanding the specializations of the complementary system and the interactions of the two. This complex interplay of overlapping computational properties puts to rest the comforting but simplistic idea that a primate vision system can be understood as a set of modules that compute separable properties of form, motion, stereo, and color. In the absence of such a convenient modular description, model neural network systems provide a natural language for analyzing the complex interactive properties that make vision so useful in our daily lives.

REFERENCES

ADELSON, E. H., & MOVSHON, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature*, **300**, 523-525.

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. M. (1980). The perception of apparent movement. *Philosophical Transactions of the Royal Society of London*, **290B**, 153-168.

ANSTIS, S. M. (1986). Motion perception in the frontal parallel plane: Sensory aspects. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance: Sensory processes and perception* (Ch. 16, pp. 16-1-16-27). New York: Wiley.

ANSTIS, S. M., & MATHER, G. (1985). Effects of luminance and contrast on direction of ambiguous apparent motion. *Perception*, **14**, 167-179.

ANSTIS, S. M., & RAMACHANDRAN, V. S. (1987). Visual inertia in apparent motion. *Vision Research*, **27**, 755-764.

BARLOW, H. B., & LEVICK, W. R. (1965). The mechanism of directional selective units in rabbit's retina. *Journal of Physiology*, **178**, 477-504.

BECK, J., GRAHAM, N., & SUTTER, A. (1991). Lightness differences and the perceived segregation of regions and populations. *Perception & Psychophysics*, **49**, 257-269.

BECK, J., ROSENFELD, A., & IVRY, R. (1990). Line segregation. *Spatial Vision*, 4, 75-101.

BISCHOF, W. F., & DILOLLO, V. (1990). Perception of directional sampled motion in relation to displacement and spatial frequency: Evidence for a unitary motion system. *Vision Research*, **30**, 1341-1362.

BRADDICK, O. J. (1974). A short range process in apparent motion. *Vision Research*, **14**, 519-527.

BRADDICK, O. J. (1980). Low-level and high-level processes in apparent motion. *Philosophical Transactions of the Royal Society of London*, **290B**, 137-151.

BROWN, J. M., & WEISSTEIN, N. (1988). A phantom context effect: Visual phantoms enhance target visibility. *Perception & Psychophysics*, **43**, 53-56.

BUCKLEY, D., FRISBY, J. P., & MAYHEW, J. E. W. (1989). Integration of stereo and texture cues in the formation of discontinuities during three-dimensional surface interpolation. *Perception*, **18**, 563-588.

BURR, D. C., ROSS, J., & MORRONE, M. C. (1986). Smooth and sampled motion. *Vision Research*, **26**, 643-652.

CAVANAGH, P., & MATHER, G. (1989). Motion: The long and short of it. *Spatial Vision*, **4**, 103-129.

CLEARY, R., & BRADDICK, O. J. (1990). Direction discrimination for band-pass filtered random dot kinematograms. *Vision Research*, **30**, 303-316.

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.

Dreher, B. (1972). Hypercomplex cells in the cat's striate cortex. *Investigative Ophthalmology*, **11**, 355-356.

Dresp, B., Lorenceau, J., & Bonnet, C. (1990). Apparent brightness enhancement in the Kanizsa square with and without illusory contour formation. *Perception*, **19**, 483-489.

Eckhorn, R., Bauer, R., Jordan, W., Brosch, 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.

Eskew, R. T., Jr. (1989). The gap effect revisited: Slow changes in chromatic sensitivity as affected by luminance and chromatic borders. *Vision Research*, **29**, 717-729.

Eskew, R. T., Jr., Stromeyer, C. F., III, Picotte, C. J., & Kronauer, R. E. (1991). Detection uncertainty and the facilitation of chromatic detection by luminance contours. *Journal of the Optical Society of America A*, **8**, 394-403.

Foster, K. H., Gaska, J. P., Nagler, M., & Pollen, D. A. (1985). Spatial and temporal frequency selectivity of neurons in visual cortical areas V1 and V2 of the macaque monkey. *Journal of Physiology*, **365**, 331-363.

Graham, N., Beck, J., & Sutter, A. (1992). Nonlinear processes in spatio-frequency channel models of perceived texture segregation: Effects of sign and amount of contrast. *Vision Research*, **32**, 719-744.

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. (1973). Contour enhancement, short-term memory, and constancies in reverberating neural networks. *Studies in Applied Mathematics*, **52**, 217-257.

Grossberg, S. (1982). *Studies of mind and brain: Neural principles of learning, perception, development, cognition, and motor control*. Boston: Reidel.

Grossberg, S. (1984). Outline of a theory of brightness, color, and form perception. In E. Degreef & J. van Buggenhaut (Eds.), *Trends in mathematical psychology* (pp. 59-86). Amsterdam: Elsevier/North-Holland.

Grossberg, S. (1987a). Cortical dynamics of three-dimensional form, color, and brightness perception: I. Monocular theory. *Perception & Psychophysics*, **41**, 87-116.

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

Grossberg, S. (Ed.) (1987c). *The adaptive brain: II. Vision, speech, language, and motor control*. Amsterdam: Elsevier/North-Holland.

Grossberg, S. (1990). Neural FACADES: Visual representations of static and moving form-and-color-and-depth. *Mind & Language*, **5**, 411-456.

Grossberg, S. (1991). Why do parallel cortical systems exist for the perception of static form and moving form? *Perception & Psychophysics*, **49**, 117-141.

Grossberg, S. (1992). *3-D vision and figure-ground separation by visual cortex* (Tech. Rep. CAS/CNS-TR-92-019). Boston: Boston University.

Grossberg, S. (in press). A solution of the figure-ground problem for biological vision. *Neural Networks*.

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. (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., & 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. *Neural Networks*, **2**, 421-450.

Grossberg, S., & Rudd, M. E. (1989b). Neural dynamics of visual motion perception: Group and element apparent motion. *Investigative Ophthalmology Supplement*, **30**, 73.

Grossberg, S., & Rudd, M. E. (1992). Cortical dynamics of visual motion perception: Short-range and long-range apparent motion. *Psychological Review*, **99**, 78-121.

Grossberg, S., & Somers, D. (1991). Synchronized oscillations during cooperative feature linking in a cortical model of visual perception. *Neural Networks*, **4**, 453-466.

Grossberg, S., & Somers, D. (1992). Synchronized oscillations for binding spatially distributed feature codes into coherent spatial patterns. In G. A. Carpenter & S. Grossberg (Eds.), *Neural networks for vision and image processing* (pp. 385-406). Cambridge, MA: MIT Press.

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.

Heggelund, P. (1981). Receptive field organization of simple cells in cat striate cortex. *Experimental Brain Research*, **42**, 89-107.

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. (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.

Humphreys, G. W., Quinlan, P. T., & Riddoch, M. J. (1989). Grouping processes in visual search: Effects with single- and combined-feature targets. *Journal of Experimental Psychology: General*, **118**, 258-279.

Julesz, B. (1971). *Foundations of cyclopean perception*. Chicago: University of Chicago Press.

Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology*, **23**, 141-221.

Kennedy, J. M. (1979). Subjective contours, contrast, and assimilation. In C. F. Nodine & D. F. Fisher (Eds.), *Perception and pictorial representation* (pp. 167-195). New York: Praeger.

Lappin, J. S., & Bell, H. H. (1976). The detection of coherence in moving random-dot patterns. *Vision Research*, **16**, 161-168.

Livingstone, M. S., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *Journal of Neuroscience*, **7**, 3416-3468.

Logothetis, N. K., Schiller, P. H., Charles, E. R., & Hurlbert, A. C. (1990). Perceptual deficits and the activity of the color-opponent and broad-band pathways at isoluminance. *Science*, **247**, 214-217.

Marr, D., & Ullman, S. (1981). Directional selectivity and its use in early visual processing. *Proceedings of the Royal Society of London B*, **211**, 151-180.

Marshall, J. A. (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.

Meyer, G. E., & Dougherty, T. (1987). Effects of flicker-induced depth on chromatic subjective contours. *Journal of Experimental Psychology: Human Perception & Performance*, **13**, 353-360.

Mikaelian, H. H., Linton, M. J., & Phillips, M. (1990). Orientation-specific luminance aftereffects. *Perception & Psychophysics*, **47**, 575-582.

Mingolla, E., Todd, J. T., & Norman, F. (1992). The perception

of globally coherent motion. *Vision Research*, **32**, 1015-1031.

MOVSHON, J. A., ADELSON, E. H., GIZZI, M. S., & NEWSOME, W. T. (1985). The analysis of moving visual patterns. In C. Chagas, R. Gattas, & C. Gross (Eds.), *Pattern recognition mechanisms* (pp. 117-152). New York: Springer-Verlag.

NAKAYAMA, K., SHIMOJO, S., & RAMACHANDRAN, V. S. (1990). Transparency: Relation to depth, subjective contours, luminance, and neon color spreading. *Perception*, **19**, 497-513.

NAKAYAMA, K., & 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. (1988). The aperture problem: II. Spatial integration of velocity information along contours. *Vision Research*, **28**, 747-753.

NEWSOME, W. T., GIZZI, M. S., & MOVSHON, J. A. (1983). Spatial and temporal properties of neurons in macaque MT. *Investigative Ophthalmology & Visual Sciences*, **24**, 106.

PARADISO, M., & NAKAYAMA, K. (1991). Brightness perception and filling-in. *Vision Research*, **31**, 1221-1236.

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.

POLLEN, D. A., GASKA, J. P., & JACOBSEN, L. D. (1989). Physiological constraints on models of visual cortical function. In R. M. J. Cotterill (Ed.), *Models of brain function* (pp. 115-135). Cambridge: Cambridge University Press.

PRAZDNY, K. (1984). On the perception of glass patterns. *Perception*, **13**, 469-478.

PRAZDNY, K. (1985). Stereopsis from kinetic and flicker edges. *Perception & Psychophysics*, **36**, 490-492.

PRAZDNY, K. (1986). Three-dimensional structure from long-range apparent motion. *Perception*, **15**, 619-625.

PRINZMETAL, W. (1990). Neon colors illuminate reading units. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 584-597.

PRINZMETAL, W., & BOAZ, K. (1989). Functional theory of illusory conjunctions and neon colors. *Journal of Experimental Psychology: General*, **118**, 165-190.

RAMACHANDRAN, V. S. (1981). *Perception of apparent movement: Studies in cognitive sciences* (Vol. 2). Irvine: University of California School of Social Sciences.

RAMACHANDRAN, V. S. (1992). Perception: A biological perspective. In G. A. Carpenter & S. Grossberg (Eds.), *Neural networks for vision and image processing* (pp. 45-92). Cambridge, MA: MIT Press.

RAMACHANDRAN, V. S., & INADA, V. (1985). Spatial phase and frequency in motion capture of random-dot patterns. *Spatial Vision*, **1**, 57-67.

REICHARDT, W. (1957). Autokorrelationsauswertung als Funktionsprinzip des Zentralnervensystems. *Zeitschrift für Naturforschung*, **12b**, 447-457.

SCHILLER, P. H., LOGOTHETIS, N. K., & CHARLES, E. R. (1990). Role of the color-opponent and broad-band channels in vision. *Visual Neuroscience*, **5**, 321-346.

SHAPLEY, R., & GORDON, J. (1985). Nonlinearity in the perception of form. *Perception & Psychophysics*, **37**, 84-88.

SHIPLEY, T. F., & KELLMAN, P. J. (1992). Perception of partly occluded objects and illusory figures: Evidence for an identity hypothesis. *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 106-120.

SNOWDEN, R. J. (1989). Motions in orthogonal directions are mutually suppressive. *Journal of the Optical Society of America*, **6**, 1096-1101.

SNOWDEN, R. J. (1990). Suppressive interactions between moving patterns: Role of velocity. *Perception & Psychophysics*, **47**, 74-78.

SPERLING, G. (1989). Three stages and two systems of visual processing. *Spatial Vision*, **4**, 183-207.

SUTTER, A., BECK, J., & GRAHAM, M. (1989). Contrast and spatial variables in texture segregation: Testing a simple spatial-frequency channels model. *Perception & Psychophysics*, **46**, 312-332.

TAKEICHI, H., SHIMOJO, S., & WATANABE, T. (1992). Neon flank and illusory contour: Interaction between the two processes leads to color filling-in. *Perception*, **21**, 313-324.

TAKEICHI, H., WATANABE, T., & SHIMOJO, S. (in press). Illusory occluding contours and surface formation by depth propagation. *Perception*.

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.

TODD, J. T., & AKERSTROM, R. A. (1987). Perception of three-dimensional form from patterns of optical texture. *Journal of Experimental Psychology: Human Perception & Performance*, **13**, 242-255.

VAN SANTEN, J. P. H., & SPERLING, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America*, **2**, 300-321.

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.

VON DER HEYDT, R., PETERHANS, E., & BAUMGARTNER, G. (1984). Illusory contours and cortical neuron responses. *Science*, **224**, 1260-1262.

WALLACH, H. (1976). *On perception*. New York: Quadrangle/New York Times Book Co.

WATANABE, T., & CAVANAGH, P. (1991). Texture and motion spreading, transparency and the aperture problem. *Perception & Psychophysics*, **50**, 459-464.

WATANABE, T., & SATO, T. (1989). Effects of luminance contrast on color spreading and illusory contour in the neon color spreading effect. *Perception & Psychophysics*, **45**, 427-430.

WATANABE, T., & TAKEICHI, H. (1990). The relation between color spreading and illusory contours. *Perception & Psychophysics*, **47**, 457-467.

WATSON, A. B., & AHUMADA, A. J. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America A*, **2**, 322-342.

WILLIAMS, D., & PHILLIPS, G. (1987). Cooperative phenomena in the perception of motion direction. *Journal of the Optical Society of America A*, **4**, 878-885.

WILLIAMS, D., & SEKULER, R. (1984). Coherent global motion percepts from stochastic local motions. *Vision Research*, **24**, 55-62.

WILSON, H. R. (1991). A psychophysically motivated model for two-dimensional motion perception. *Investigative Ophthalmology & Visual Science*, **32**, 893 (Abstract No. 1101).

ZEKI, S. M. (1974a). Cells responding to changing image size and disparity in the cortex of the rhesus monkey. *Journal of Physiology (London)*, **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 (London)*, **236**, 549-573.

(Manuscript received October 30, 1991;
revision accepted for publication July 28, 1992.)