Pergamon

0042-6989(95)00052-6

Vision Res. Vol. 36, No. 1, pp. 149–173, 1996

0042-6989/96 $9.50 + 0.00

Cortical Dynamics of Form and Motion Integration: Persistence, Apparent Motion, and Illusory Contours

GREGORY FRANCIS,* STEPHEN GROSSBERG†‡

Received 22 March 1994; in revised form 20 September 1994; in final form 12 January 1995

How does the visual system generate percepts of moving forms? How does this happen when the forms are emergent percepts, such as illusory contours or segregated textures, and the motion percept is apparent motion between the emergent forms? We develop a neural model of form–motion interactions to explain and simulate parametric properties of psychophysical motion data and to make predictions about how the parallel cortical processing streams V1 → MT and V1 → V2 → MT control form–motion interactions. The model explains how an illusory contour can move in apparent motion to another illusory contour or to a luminance-derived contour; how illusory contour persistence relates to the upper interstimulus interval (ISI) threshold for apparent motion; and how upper and lower ISI thresholds for seeing apparent motion between two flashes decrease with stimulus duration and narrow with spatial separation (Korte's laws). The model accounts for these data by suggesting how the persistence of a boundary segmentation in the V1 → V2 processing stream influences the quality of apparent motion in the V1 → MT stream through V2 → MT interactions. These data may all be explained by an analysis of how orientationally tuned form perception mechanisms and directionally tuned motion perception mechanisms interact.

Form perception Motion perception Neural networks Visual cortex Visual persistence Apparent motion Korte's laws

1. INTRODUCTION

How does the visual system integrate visual form and motion information to generate a coherent percept of moving forms? It is well known that many percepts of form are emergent properties of images and scenes, much as illusory contours help to group textured scenes into detectable objects (Beck, Prazdny, & Rosenfeld, 1983). It is also now known that many motion percepts depend on detection of form. Chubb and Sperling (1991) described motion percepts that are undetectable by some types of motion analysis and noted how detection of moving form might explain these percepts. Cavanagh and Mather (1989) argued that some properties of apparent motion require detection of moving forms. Wilson, Ferrera, and Yo (1993) described how integration of luminance-based and form-based motion could help to explain percepts of moving plaids.

There exists neurophysiological evidence of form and motion integration as well. Neurophysiological studies suggest that properties of motion (Maunsell & van Essen, 1983; Albright, Desimone & Gross, 1984; Mikami, Newsome & Wurtz, 1986a) and apparent motion (Mikami, Newsome & Wurtz, 1986b; Newsome, Mikami & Wurtz, 1986) are represented in the processing stream of visual cortex that includes area MT. von der Heydt, Peterhans, and Baumgartner (1984) and Peterhans and von der Heydt (1989) have reported evidence that the interblob cortical stream generates emergent properties of visual form, notably in area V2. The existence of these parallel cortical processing streams for form and motion processing raises the question of how form and motion processing can interact. One possible link is between cortical areas V2 and MT. Such a pathway does exist (DeYoe & van Essen, 1988).

Grossberg (1991) outlined a model that suggested how a V1 → V2 → MT link between form and motion processing mechanisms could play two related functional roles. First, such a connection would allow the motion detecting system to respond to diverse perceptual groupings, such as illusory contours and segmentations of textures. Second, it was explained how the model form processing mechanisms could make finer disparity matches than could its motion processing mechanisms.

*Purdue University, Department of Psychological Sciences, 1364 Psychological Sciences Building, West Lafayette, IN 47907, U.S.A.
†Center for Adaptive Systems and Department of Cognitive and Neural Systems, Boston University, 111 Cummington Street, Boston, MA 02115, U.S.A. [*Fax* 1 617 353 7755].
‡To whom reprint requests should be addressed.

These model mechanisms were used to clarify analogous neurophysiological evidence about depth sensitivity from cells in V2 (von der Heydt, Hänny & Dürsteler, 1981) and MT (Logothetis, Schiller, Charles & Hurlbert, 1990; Schiller, Logothetis & Charles, 1990). With this model analysis in hand, it was proposed how the form-to-motion pathway could help the motion system create motion signals at the computed depths of the perceptual groupings, thereby integrating form and motion data into consistent percepts of moving forms. This analysis suggested at what processing stages outputs from the form system should input to the motion system.

This article develops the Grossberg (1991) proposal by showing how to link neural models of emergent boundary segmentation (Grossberg, 1987a, b, 1994; Grossberg & Mingolla, 1985a, b, 1987) and motion perception (Grossberg & Mingolla, 1993; Grossberg & Rudd, 1989, 1992) to explain and simulate challenging data about form–motion interactions. Conceptually, the model building in this paper simply links two established models of form and motion processing in an appropriate way. The result is an analysis of form and motion percepts that mechanistically links together several types of data that heretofore have been treated separately. In particular, it links together data about the persistence of static images with data about the quality of apparent motion. The key idea is to relate the time taken to generate and reset a persistent boundary segmentation in the form cortical stream (through V2) with threshold properties of apparent motion in the motion cortical stream (through MT).

A key property distinguishing this paper from the earlier approaches to form and motion integration is thus that we use the *dynamic* characteristics of the form processing system to explain data about motion perception. As a first step in analyzing the form–motion interactions, we simulate visual displays that generate two dimensional percepts of apparent motion of moving forms. Our analysis simulates three sets of illustrative data using a fixed set of model parameters:

- Illusory contours move in apparent motion and do not obey the inverse relationship between upper interstimulus interval (ISI) thresholds and stimulus duration that is characteristic of luminance-based contours (von Grünau, 1979; Ramachandran, 1985; Mather, 1988).
- Apparent motion can occur between one stimulus defined by illusory contours and a second stimulus defined by luminance contrast (von Grünau, 1979; Cavanagh, Arguin & von Grünau, 1989).
- Korte's laws: for luminance-based stimuli, both upper and lower ISI thresholds are inversely related to flash duration. The range of ISIs capable of producing apparent motion narrows as the spatial separation between the flashes increases (Neuhaus, 1930; Kolers, 1972; Korte, 1915).

Before presenting the details of model mechanisms, we briefly describe these data and how the model addresses each of these data sets.

1.1. Apparent motion of illusory contours

Several authors have shown that illusory contours can move in apparent motion (von Grünau, 1979; Ramachandran, 1985; Mather, 1988). During the first time period of the experiment by Ramachandran (1985), subjects saw an illusory Kanizsa square on the left side and a jumbled set of lines on the right side. During the second time period, the pacman circles that induced the Kanizsa square filled up and the illusory Kanizsa square disappeared. At the same time, lines on the right within a region defined by an illusory square disappeared. Subjects reported seeing motion of the illusory square from the left to the right. Features in the two images cannot be matched, but the illusory Kanizsa squares which these features induce can be matched. Subjects in fact saw motion from one illusory square to the other. von Grünau (1979) reported similar results.

Mather (1988) investigated the temporal properties of illusory contour apparent motion. Figure 1(a) (from Mather) shows contour plots of reports of seeing apparent motion between two illusory Kanizsa squares as a function of the inducing stimuli duration for two subjects. [Some of these points are extrapolations from measured points as Mather (1988) did not provide the values of measured data points.] Of key interest is the inverted-U shape of the top curves that divide regions of 55% perceived motion. This curve can be considered as the upper ISI threshold for perceiving apparent motion of the illusory contour. Figure 1(c) replots the curves on non-logarithmic axes. The shape of these curves is unlike the data derived from apparent motion of luminance-based contours. For example, as described below, the upper ISI threshold values for illusory contour motion are larger than those of the corresponding luminance-based contour motion. The inverted-U shape of threshold ISIs as a function of stimulus duration is also unlike that of luminance-based contours, for which threshold ISIs are inversely related to stimulus duration, as described below.

1.2. Persistence of illusory contours

The model explains these properties of illusory contour apparent motion as a consequence of form–motion interactions. The illusory contours are generated in the form perception system and input to the motion perception system. Key properties of Mather's motion perception data are explained using properties of the stationary illusory contours that are computed in the form system. In particular, data on illusory contour persistence of Meyer and Ming (1988) are described by curves that are remarkably similar to the upper ISI thresholds found by Mather (1988). Francis, Grossberg, and Mingolla (1994) simulated the persistence of illusory and luminance-based contours using the Grossberg and Mingolla (1985b) Static Boundary Contour System (BCS) model of static boundary segmentation by the V1 → V2 cortical

stream. As in the data of Meyer and Ming (1988), the Static BCS representation of an illusory contour lasts longer than that of a real contour and exhibits an inverted U relationship between persistence and stimulus duration. These properties are traced in Francis *et al.* (1994) to an analysis of why the illusory contour boundary takes longer to form than a luminance-based boundary (increasing portion of the inverted-U curve) and has fewer reset signals to shut it off (decreasing portion of the inverted-U curve). This analysis is reviewed in Section 2 for completeness.

The Static BCS model is distinguished from the Motion BCS model of motion boundary segmentation by the V1 → MT cortical stream (Chey, Grossberg & Mingolla, 1994; Grossberg & Mingolla, 1993; Grossberg, Mingolla & Nogueira, 1993; Grossberg & Rudd, 1989, 1992). The current model of form and motion integration suggests how the Static BCS interacts with the Motion BCS to model the cortical V1 → V2 → MT interaction. In particular, the persistence of illusory contour inputs from the form model (Static BCS) to the motion model (Motion BCS) determines the upper ISI threshold of apparent motion of the contour. In this way the dynamic characteristics of form processing are used below to explain the data in Fig. 1(a). Figure 1(b) summarizes ISI thresholds for computer simulations of illusory contour motion in the model that qualitatively match the properties of the curves found by Mather (1988). In particular, Fig. 1(c) plots the ISI thresholds from the subjects in Mather's study and the model. The results are similar in magnitude and qualitative shape. This simulation used the same parameters for the Static BCS as that were used in Francis *et al.* (1994) to simulate data from Meyer and Ming (1988), who directly measured the inverted-U relationship between persistence and stimulus duration.

Although the simulated curve falls within the experimental curves, our goal in this article is to demonstrate key qualitative, rather than quantitative, relationships. One reason is that the specific shape of ISI curves found by Mather seems to be subject-dependent. In addition, the approximations that are necessary to make the simulations computationally feasible do not yet warrant a search for optimal parameters.

1.3. Interattribute motion

von Grünau (1979) observed that subjects sometimes can see apparent motion between an illusory contour and a contour defined by luminance edges. Cavanagh *et al.* (1989) generalized this result by showing that

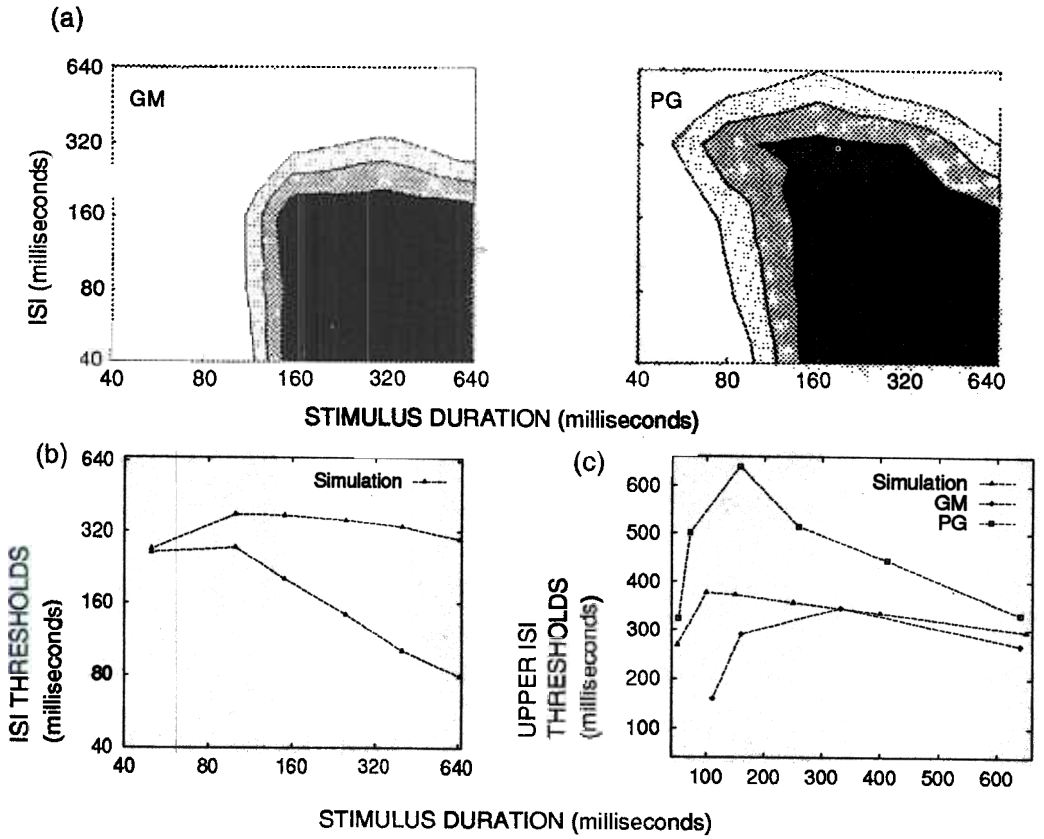


FIGURE 1. (a) Contour maps depicting the percentage of trails in which apparent motion of an illusory contour was reported as a function of stimulus duration and ISI. Black, 85–100%; mid-gray, 70–85%; light gray, 55–70%; white <55%. [Used with permission from Mather (1988).] (b) Computer simulation of upper and lower ISI thresholds as a function of stimulus duration. (c) Upper ISI thresholds for perceiving illusory contour apparent motion for subjects and the simulation. The simulation thresholds fall between the data of the subjects. For each subject and the simulation, maximal ISI takes a peak value at an intermediate stimulus duration.

subjects reported seeing motion between stimuli defined by any combination of attributes, including luminance, color, texture, relative motion, or stereopsis. They also noted that motion between stimuli of different attributes is weaker than motion between stimuli of the same attribute. Taken together, these studies suggest that interattribute motion is a result of form and motion integration. The Static BCS is capable of responding to multiple types of form-supporting cues, including luminance, color, texture, shading, and stereo cues (Cruthirds, Grossberg & Mingolla, 1993; Graham, Beck & Sutter, 1992; Grossberg, 1987a, b, 1994; Grossberg & Mingolla, 1985a, b, 1987; McLoughlin & Grossberg, 1994; Sutter, Beck, & Graham, 1989). The Motion BCS is capable of responding to a wide range of apparent motion, first-order motion, and second-order motion cues (Grossberg, Mingolla & Nogueira, 1993; Grossberg & Rudd, 1989, 1992). Thus many properties of interattribute motion could, in principle, be explained by interactions between the Static BCS and the Motion BCS.

Computer simulations described below show how the model generates interattribute motion between an illusory contour and a contour defined by luminance edges (von Grünau, 1979). Grossberg (1994) modeled how certain combinations of luminance, color, texture, size, and depth information are bound more closely together than others during three-dimensional perception, and thus may more easily activate motion percepts between themselves using mechanisms such as those in the present form-motion model.

1.4. Korte's laws, transient cells, and visual persistence

Figure 2(a) shows the upper and lower ISI threshold values for apparent motion of luminance-based stimuli (Kolers, 1972; after Neuhaus, 1930). This figure shows that as stimulus duration increases from 10 to 45 to 90 msec, each upper and lower ISI threshold curve decreases at every spatial separation. Moreover, as the distance between the two stimuli increases, the range of ISIs that produce apparent motion narrows, with the upper ISI decreasing and the lower ISI increasing for every stimulus duration. These properties are often collectively referred to as Korte's laws (Korte, 1915).

Grossberg and Rudd (1992) explained the characteristics of the lower ISI thresholds and the role of spatial separation using the Motion BCS. Both the Motion BCS and the Static BCS are broken up into a multistage filter followed by a grouping or segmentation network. Grossberg and Rudd focused their study on the motion filter, which is called the Motion Oriented Contrast (or MOC) Filter. The MOC Filter models how interactions between sustained cells and transient cells produce motion direction signals in response to input changes. In the Grossberg and Rudd simulations, the MOC Filter was sensitive only to changes in luminance. In the current form–motion model, the boundary segmentation outputs of the Static BCS input to the MOC Filter. The MOC Filter is thus sensitive to changes in form as well as changes in luminance. The fast responses of transient cells to luminance flashes allow the MOC Filter to respond to rapidly moving stimuli. In contrast, form processing by the BCS is substantially slower than motion processing, so that persisting form inputs from the Static BCS to the MOC Filter often outlast the effects of purely luminance based inputs. These form–motion interactions are shown below to be sufficient to explain the properties of the upper ISI thresholds in the classical Korte's laws.

In particular, Francis *et al.* (1994) showed that increasing the duration of a stationary form input decreases the persistence of the boundary representation in the Static BCS, much as experiments on visual persistence (e.g. Bowen, Pola & Matin, 1974) report an inverse relationship between persistence and stimulus duration. As in the case of illusory contours above, the persistence of form signals determines the upper ISI threshold of apparent motion. The remaining properties of Korte's laws—namely the lower ISI thresholds and influences of spatial separation—are explained below in terms of MOC Filter properties. Figure 2(b) summarizes computer simulations of how the form–motion model simulates Korte's laws. This figure demonstrates that the model reproduces all the qualitative properties of the classical Neuhaus (1930) data. The most important quantitative property is also explained, namely the 350 msec gap between the smallest lower ISI threshold and the largest upper ISI threshold.

2. MODEL FORM AND MOTION INTERACTIONS

2.1. Boundary segmentation

2.1.1. Spatial interactions. Grossberg (1984) and Cohen and Grossberg (1984) introduced the Static BCS model. Grossberg and Mingolla (1985a, b, 1987) developed the model to simulate how the visual system detects, completes, and regularizes boundary segmentations in response to a variety of retinal images. Such segmentations can be defined by regions of different luminance, color, texture, shading, or stereo signals. The Static BCS computations for single-scale monocular processing consist of a series of filtering, competitive, and cooperative stages as schematized in Fig. 3 and reviewed in several reports (e.g. Grossberg, 1987a, 1994; Grossberg, Mingolla & Todorović, 1989). The first stage, schematized as an unoriented annulus in Fig. 3, models in perhaps the simplest possible way the shunting on-center off-surround interactions at the retinal and LGN levels. These cells compensate for variable illumination and enhance regions of local contrast in the image. Interactions of on-center off-surround and off-center on-surround cells are not needed here, but their complementary responses to images are modeled elsewhere (Gove, Grossberg & Mingolla, 1994a, b; Grossberg, Mingolla & Williamson, 1994; Grossberg & Wyse, 1991; Pessoa, Mingolla & Neumann, 1994).

These model LGN cells input to pairs of like-oriented simple cells that are sensitive to opposite contrast polarity, or direction-of-contrast. The simple cell pairs, in

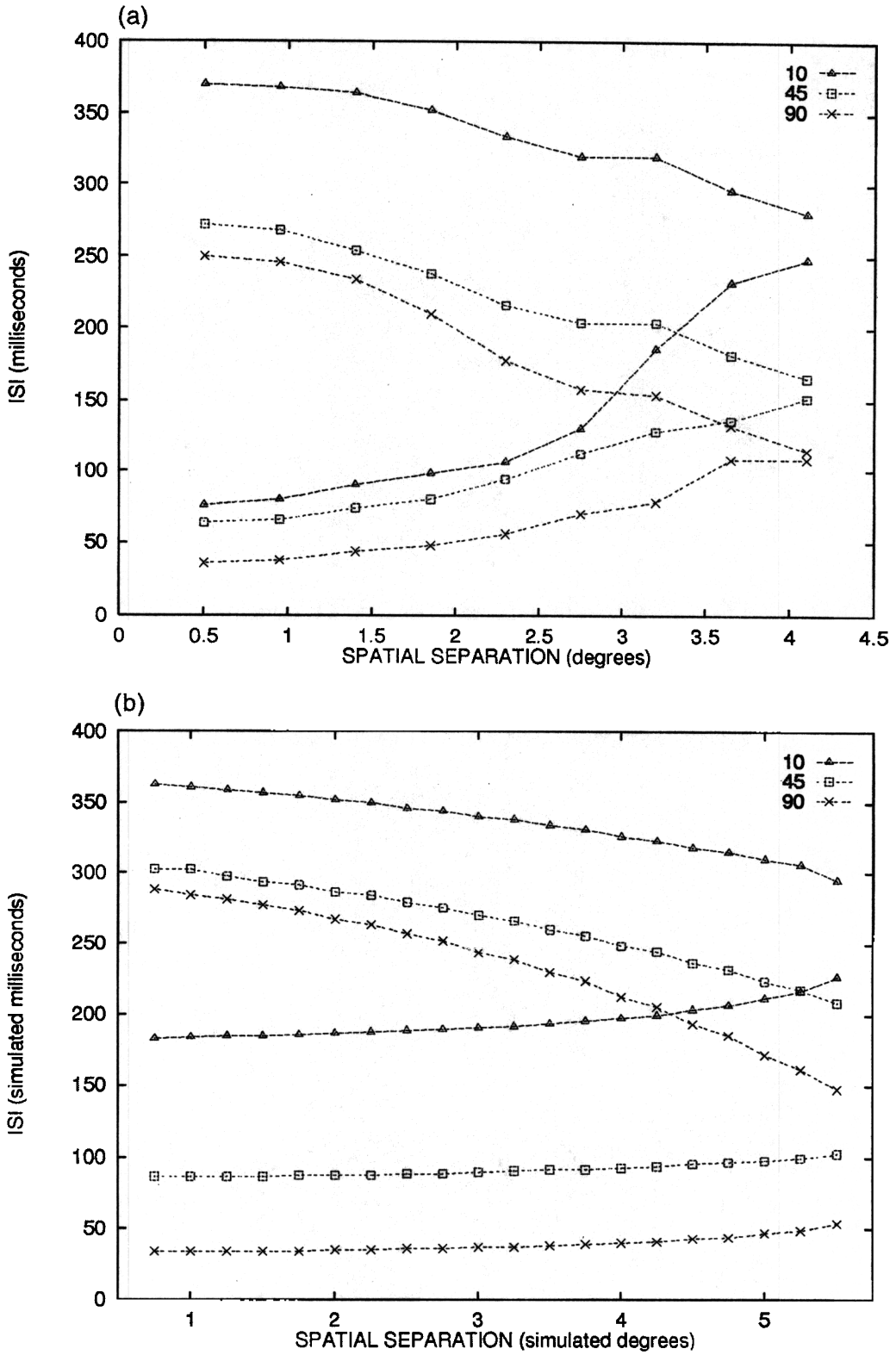


FIGURE 2. Upper and lower ISI thresholds as a function of spatial separation for three stimulus durations. Increasing flash duration decreases the threshold ISI values. The upper ISI threshold decreases with spatial separation and the lower ISI threshold increases with spatial separation. (a) Psychophysical data. [Redrawn from Kolers (1972) after data from Neuhaus (1930).] (b) Simulated ISI threshold values.

turn, send their rectified output signals to like-oriented complex cells. Complex cells are thus rendered insensitive to direction-of-contrast, as are all subsequent cell types in the model. Complex cells activate hypercomplex cells through an on-center off-surround network (first competitive stage) whose off-surround carries out an

endstopping operation. In this way, complex cells excite hypercomplex cells of the same orientation and position, while inhibiting hypercomplex cells of the same orientation at nearby positions. One role of this spatial competition is to spatially sharpen the neural responses to oriented luminance edges. Another role is to initiate the process, called *end cutting*, whereby boundaries are formed that abut a line end at orientations perpendicular or oblique to the orientation of the line itself (Grossberg, 1987a; Grossberg & Mingolla, 1985b).

The signals from complex cells to hypercomplex cells are multiplied, or gated, by habituative chemical transmitters. These habituative gates help to reset boundary segmentations in response to rapidly changing imagery, as discussed below. The hypercomplex cells input to a competition across orientations at each position (second competitive stage) among higher order hypercomplex cells. This competition acts to sharpen up orientational responses at each position, and to work with the habituative gates to reset boundary segmentations, as discussed below.

Output from the higher-order hypercomplex cells feed into cooperative bipole cells that initiate long-range boundary grouping and completion. Bipole cells fire only if both of their receptive fields are sufficiently activated by appropriately oriented hypercomplex cell inputs. Bipole cells hereby realize a type of long-range cooperation among the outputs of active hypercomplex cells. For example, a horizontal bipole cell, as in Fig. 3, is excited by activation of horizontal hypercomplex cells that input to its horizontally oriented receptive fields. A horizontal bipole cell is also inhibited by activation of vertical hypercomplex cells.

Bipole cells were predicted to exist in Cohen and Grossberg (1984) and Grossberg (1984) shortly before cortical cells in area V2 with similar properties were reported by von der Heydt *et al.* (1984). At around the time of the von der Heydt *et al.* report, Grossberg and Mingolla (1985a, b) used bipole cell properties to simulate and explain a variety of data about illusory contour formation, neon color spreading, and texture segregation. These same properties play a role in our

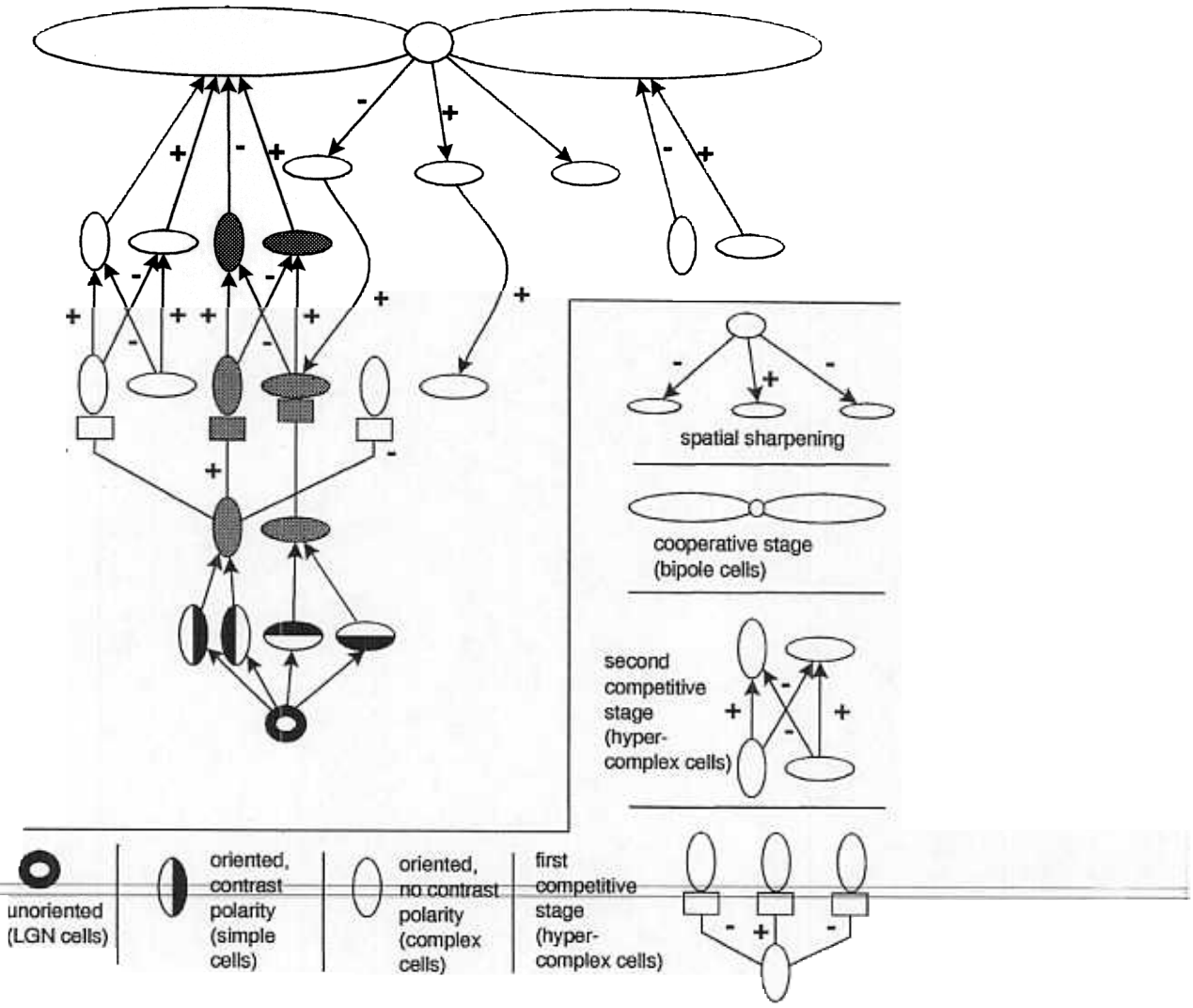


FIGURE 3. Boundary Contour System with embedded gated dipoles. See text for details

explanations of apparent motion of illusory contours and interattribute apparent motion.

Bipole cells generate feedback signals to like-oriented hypercomplex cells. These feedback signals help to create and enhance spatially and orientationally consistent boundary groupings, while inhibiting inconsistent ones. In particular, bipole cell feedback excites hypercomplex cells at the same orientation and position while inhibiting cells at nearby positions. Hypercomplex boundary signals with the most cooperative support from bipole grouping thereupon further excite the corresponding bipole cells. This cycle of bottom-up and top-down interaction between hypercomplex cells and bipole cells rapidly converges to a final boundary segmentation. Feedback among bipole cells and hypercomplex cells hereby drives a resonant cooperative-competitive decision process that completes the statistically most favored boundaries, suppresses less favored boundaries, and coherently binds together appropriate feature combinations in the image.

2.1.2. Temporal dynamics. The positive feedback within the hypercomplex-bipole feedback loop also creates hysteresis that could, if left unchecked, lead to undesirably long boundary persistence after stimulus offset, and thus to uncontrolled image smearing in response to image motion (Burr, 1980). In particular, each cell in the BCS has its own local dynamics involving activation by inputs and passive decay (of the order of 10 simulated msec). However, the excitatory feedback loop dominates the temporal aspects of the BCS. As shown in Francis *et al.* (1994), when inputs (luminance edges or illusory contour inducers) feed into the BCS, they trigger reverberatory interactions that, if left unchecked can last for hundreds of simulated msec. This is true because hypercomplex and bipole cell activities at a particular position and orientation decay away only when bipole cell output centered at the same position and orientation weakens. Since bipole cell activation depends on inputs to both receptive fields, bipole activation near the ends of contours weakens first after inputs shut off. As these bipole cells lose activation, so do all other cells of the same orientation and position. This decay causes more bipole cell activities to decay, which continues the process. The net effect of these spatial and temporal interactions is that boundary activities erode from contour ends to the contour middle. This erosion is observable in the simulations in Figs 11(c), 13(c) and 14(c).

Outward-to-inward boundary erosion makes predictions about how masking stimuli may influence the perception of illusory contours such as Kanizsa squares. Masking the pacmen that generate a Kanizsa square may not immediately obliterate the illusory contours between the pacmen because of persistent resonance at these locations. A second masking stimulus at these locations can thus influence the persistence of these illusory fragments, as Shapley and his colleagues have recently shown (Shapley, personal communication).

The problem for the Static BCS is to accelerate this boundary erosion in response to rapidly changing imagery. More generally, the BCS needs to use resonant feedback to maintain segmentations of unmoving scenic objects, even as it actively resets segmentations corresponding to rapidly changing scenic objects. The net effect is to control image smearing in a form-sensitive way. Remarkably, the same BCS mechanisms that create resonant boundaries can also be used to reset them. Two types of mechanism maintain the desired tradeoff between resonance and reset. The first mechanism uses the orientational competition that converts model hypercomplex cells into higher-order hypercomplex cells. Consider how this competition works between pairs of mutually perpendicular cells. Pairs of mutually perpendicular complex, hypercomplex, and higher-order hypercomplex cells, designated in gray within Fig. 3, define a specialized type of opponent processing circuit that Grossberg (1972) has called a gated dipole. The gates in the dipole are the habituative transmitters that multiply signals in the pathways from complex to hypercomplex cells (square synapses in Fig. 3). Such a gated dipole can rapidly inhibit a bipole cell when its activating image features shut off or are removed due to image motion.

To see how this works, suppose that a horizontal edge turns on horizontally oriented complex, hypercomplex, and bipole cells, thereby generating a horizontal boundary segmentation. Offset of the horizontal edge can cause an antagonistic rebound of activity in the corresponding gated dipoles, leading to activation of vertically oriented hypercomplex cells and inhibition of horizontal bipole cells. The rebound is generated as follows. When the horizontal input is on, horizontal transmitter gates habituate. The net result is an overshoot of input to horizontal bipole cells, followed by a steady input level after habituation takes place. When the input subsequently shuts off, the altered balance of transmitter between the horizontal and vertical channels favors the vertical channel and permits vertical cell activity to rebound in response to an internally generated tonic input that equally activates both channels. When this happens, an inhibitory input to the bipole cell occurs. The rebound is transient because transmitters in both channels then gradually equilibrate to equal levels. In summary, rebound-driven inhibition of the bipole cells selectively limits persistence and smearing at those locations where the image is changing. [See Francis *et al.* (1994) for further details and simulations.]

Several conceptual and data-related properties of reset by a transient antagonistic rebound are worth noting here. The first is that, in more complex versions of the BCS, both ON cells that are turned on by an input and OFF cells that are turned off by an input are modeled (Gove *et al.*, 1994a, b; Grossberg, 1991). In such a network, offset of a horizontal ON cell mediates transient activation of a horizontal OFF cell, as well as the type of onset of a vertical ON cell that is here simulated. Inhibition of horizontal bipole cells may thus be mediated by horizontal OFF cells, rather than by vertical ON cells, as here assumed, for simplicity. In this more

general model, any influence of vertical cells on horizontal bipoles could be mediated by horizontal ON or OFF cells.

In support of such opponent interactions, whether mediated by ON cells or OFF cells, orientationally opponent aftereffects are well known to occur psychophysically (MacKay, 1957; Taylor, 1958). From a physiological perspective, several components of the gated dipole circuit have known cellular correlates in visual cortex, including tonically active cells (such as the cells that feed the habituative transmitters) and polarization from opposite orientations (Creutzfeldt, Kuhnt & Benevento, 1974; Levitt, Kiper & Movshon, 1994). Further neurophysiological experiments are needed to test the cellular substrate of this predicted boundary reset mechanism and, by implication, of orientationally opponent aftereffects.

The rebound-driven reset mechanism shuts off boundary segmentations at locations that lose input support due to image offset or motion. The second reset mechanism helps to prevent image smearing across space. It uses the spatial endstopping competition among like-oriented hypercomplex cells at the first competitive stage (Fig. 3). Castet (1994) has reported experiments that are consistent with this model prediction. Francis *et al.* (1994) showed that these two mechanisms of the BCS model are sufficient to explain the key parametric properties of visual persistence experiments.

To explain properties of apparent motion, two characteristics of visual persistence are particularly important. First, psychophysical studies of visual persistence demonstrate that persistence duration decreases in response to image edges as stimulus duration increases (Bowen *et al.*, 1974). Francis *et al.* (1994) provide simulations of boundary signal persistence that agrees with these findings. The strength of the inhibitory rebound in the gated dipole mechanism explains the model's results. As stimulus duration increases, the gate habituates more and the strength of the subsequent rebound covaries with the amount of habituation. A longer stimulus generates stronger inhibition at stimulus offset, thereby hastening the erosion of boundary signals, and reducing measured persistence.

Second, psychophysical studies of illusory contour persistence (Meyer & Ming, 1988) show that persistence duration increases with stimulus duration up to about 200 msec and then decreases as stimulus duration grows still longer. These findings, too, have been simulated using BCS interactions (Francis *et al.*, 1994). Since illusory contour inducers have a smaller proportion of luminance edges than an image edge of equal length, they take longer to establish a strong reverberation in the feedback loop of the BCS. As stimulus duration increases, the reverberation grows stonger, up to some maximum. Without reset signals, stronger reverberations lead to longer persistence. But the reset signals also grow stronger as stimulus duration increases. As stimulus duration increases up to about 200 msec, the increase in reverberation strength leads the increase in reset signal strength. Beyond stimulus durations of 200 msec, the reverberation strength does not change greatly, but the strength of the reset signals continues to grow due to the slowly habituating transmitters, thereby causing persistence to decrease. Since the illusory contour has shorter luminance contour inducers than an equal length edge, it produces fewer reset signals, thereby allowing greater persistence of illusory contours than luminance-defined stimuli, in agreement with the psychophysical data of Meyer and Ming (1988).

We will show below how the persistence of BCS output signals establishes the upper ISI thresholds for apparent motion. This hypothesis links psychophysical data on visual persistence of non-moving stimuli to the data on moving stimuli.

2.2. Early motion processing

Grossberg and Rudd (1989, 1992) and Grossberg and Mingolla (1993) developed the Motion BCS and its front end, the MOC Filter, to explain a broad range of motion and apparent motion data. First and foremost, the Motion BCS embodies the idea that motion processing can generate boundary segmentations of moving objects. Whereas boundary segmentations of the Static BCS compute properties based on static image *orientations*, boundary segmentations of the Motion BCS compute properties based on moving image *directions*. Both BCS systems generate segmentations whose outputs are insensitive to direction-of-contrast, so that their boundaries can interpolate textured and shaded image regions where contrast polarity reverses.

The MOC Filter may be conceptually described in several ways. It is a minimal filter that produces output signals that are insensitive to direction-of-contrast but sensitive to direction-of-motion. It pools information from multiple orientations and unoriented input signals into directionally selective output signals. To accomplish the transformation from multiple orientations to prescribed motion directions, the MOC Filter uses a hierarchy of short-range and long-range spatial interactions that help to explain data about short-range and long-range motion within a single system. The qualitative properties of the five MOC Filter processing levels of Fig. 4 are summarized below. Representative equations are listed in the Appendix.

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," and to thereby process ratio contrasts in the image (Grossberg & Todorović, 1988).

Level 2: sustained cell short-range filter

Four operations occur here, as illustrated in Fig. 5.

Space average. Inputs are processed by individual oriented receptive fields, or simple cells, which add excitatory and inhibitory contributions from two halves of the receptive field.

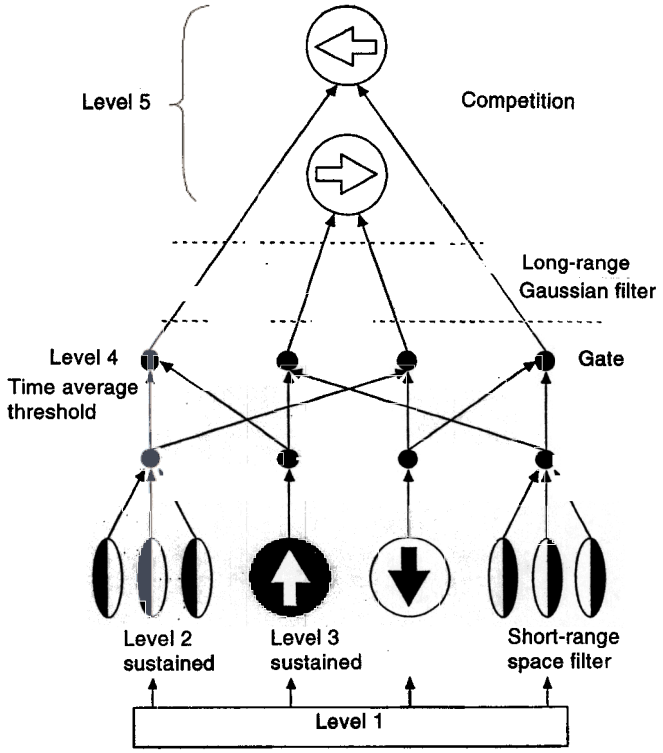


FIGURE 4. The MOC Filter. The input pattern (Level 1) is spatially and temporally filtered parallel by both sustained response cells with oriented receptive fields that are sensitive to direction-of-contrast (Level 2) and transient response cells with unoriented receptive fields that are sensitive to the direction-of-contrast change in the cell input (Level 3). Level 4 cells combine sustained cell and transient cell signals multiplicatively and are thus rendered sensitive to both direction-of-motion and direction-of-contrast. Level 5 cells sum across space, orientation, and oppositely polarized Level 4 cells to become sensitive to direction-of-motion but insensitive to direction-of-contrast.)

Rectify. The output signal from a simple cell grows with its activity above a signal threshold. Thus, the output is half-wave rectified.

Short-range spatial filter. A spatially aligned array of simple cells with like orientation and direction-of-contrast pool their output signals to activate the next cell level. As shown in Fig. 5, the target cells are pooled in a movement direction that is not necessarily perpendicular to the simple cell's preferred orientation. This spatial pooling plays the role of the short-range motion limit D_{max} (Braddick, 1974). The breadth of spatial pooling scales with the size of the simple cell receptive fields [Fig. 5(a, b)]. Correspondingly, D_{max} depends on the spatial frequency content of the image (Anderson & Burr, 1987; Burr, Ross & Morrone, 1986; Nakayama & Silverman, 1984, 1985; Petersik, Pufahl & Krasnoff, 1983) and is not a universal constant.

Time average. The target cell time averages the inputs that it receives from its short-range spatial filter. This operation has properties akin to the "visual inertia" during apparent motion that was reported by Anstis and Ramachandran (1987); see Fig. 5(c).

Level 3: transient cell filter

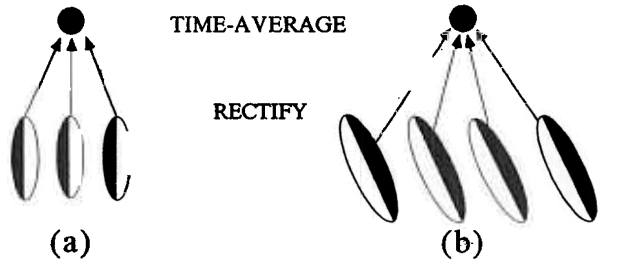
In parallel with the sustained cell filter, a transient cell filter reacts to input increments (on-cells) or decrements (off-cells) with positive outputs (Fig. 6). These filters use five operations.

Space average. This is accomplished by a receptive field that sums inputs over its entire range, unlike the receptive field of a sustained cell. This receptive field is assumed to be unoriented, or circularly symmetric, for simplicity.

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

Change detector. The on-cells are activated when the time average increases [Fig. 6(a)]. The off-cells are activated when the time average decreases [Fig. 6(b)].

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



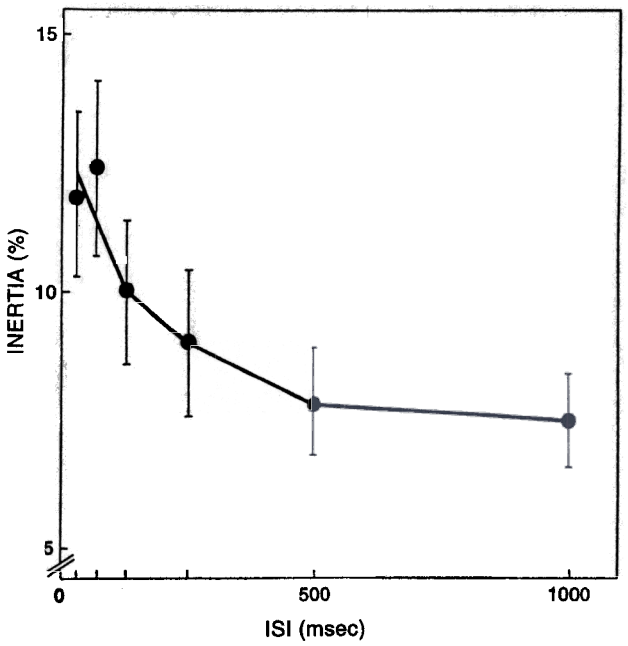


FIGURE 5. The sustained cell short-range filter. Inputs are spatiotemporally filtered by sustained cells with individual oriented receptive fields and temporal filtering characteristics that are determined by the dynamics of a shunting membrane equation. The ouput of each sustained cell is rectified and thresholded. The outputs of a spatially aligned array of cells with like orientation, direction-of-contrast, and direction-of-motion are pooled. The breadth of the spatial pooling scales with the size of the simple cell receptive fields, as in (a) and (b). [Reprinted with permission from Grossberg and Rudd (1992).] (c). Visual inertia in apparent motion measured by Anstis and Ramachandran (1987). [Ambiguous apparent motion was biased by priming dots, and the degree of bias (inertia) was measured as a function of the ISI between the priming dot and test. The bias induced by the priming dots was about 12% at short ISIs and fell monotonically to about 7% for ISIs exceeding 500 msec.] Reprinted with permission from Anstis and Ramachandran (1987).

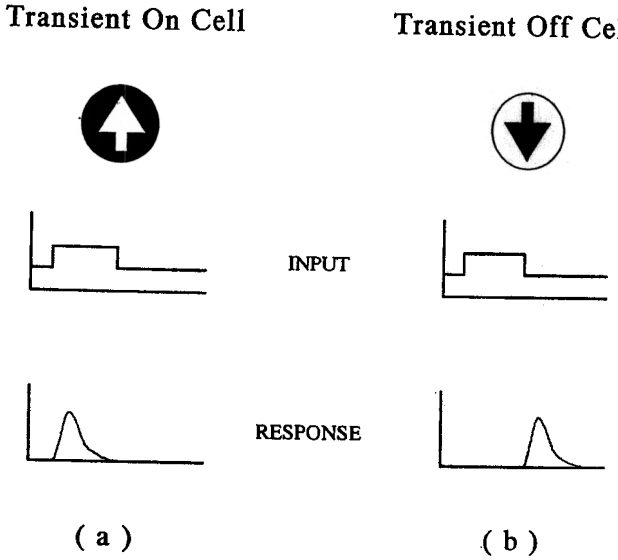


FIGURE 6. Responses over time of transient on- and off-cells. (a) On-cell responses are formed from the positive-rectified and thresholded time derivative of a spatiotemporally filtered image. The spatial filter has an unoriented on-center, off-surround receptive field. The temporal filter is based on the dynamics of a shunting membrane equation that time averages the spatially filtered input. The on-cell thus produces a time-averaged response to an increment in the input. (b) Off-cells are formed from the negative-rectified and thresholded time-averaged response to a decrement in the input. [Reprinted with permission from Grossberg and Rudd (1992).]

Habituative transmitter. The rectified signals are multiplied by a habituative transmitter that limits their duration even in response to prolonged monotonic inputs.

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 deg. Multiplicative gating of each Level 2 sustained cell output with a Level 3 transient cell on-cell or off-cell removes this ambiguity (Fig. 7). For example, consider a sustained cell output from vertically oriented dark-light simple cell receptive fields that are joined together in the horizontal direction by the short-range spatial filter [Fig. 5(a)]. Such a sustained cell output is maximized by a dark–light image contrast moving to the right or to the left. Multiplying this Level 2 output with a Level 3 transient on-cell output generates a Level 4 cell that responds maximally to motion to the left. Multiplying it with a Level 3 off-cell output generates a Level 4 cell that responds maximally to motion to the right.

Multiplying a sustained cell with a transient cell is the main operation of the Marr and Ullman (1981) motion detector. Despite this similarity, Grossberg and Rudd (1989) described six basic differences between the MOC Filter and the Marr–Ullman model: 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 step requires long-range spatial filtering and competitive sharpening, described below, that are also not part of the Marr–Ullman model. This difference in design rationale is fundamental. The Marr–Ullman model espouses an "independent modules" perspective. In contrast, the MOC Filter generates an output that is independent of direction-of-contrast, and thus is perceptually invisible. Its boundary segmentations help to form compartments in which a complementary "seeing" system, called the Feature Contour System (FCS), fills-in surface representations of brightness, color, depth, and form (Arrington, 1994; Cohen & Grossberg, 1984; Grossberg, 1987b, 1994; Grossberg & Todorović, 1988; Paradiso & Nakayama, 1991). The BCS and FCS are thus not independent modules. Rather, they have been shown to obey computationally complementary rules whose individual insufficiencies are overcome via BCS ⇆ FCS interactions.

Level 5: long-range spatial filter and competition

Outputs from Level 4 cells that are sensitive to the same direction-of-motion but opposite directions-of-contrast activate individual Level 5 cells by a long-range spatial filter that has a Gaussian profile across space (Fig. 8). This long-range filter also groups together Level 4 cell outputs that are derived from Level 3 short-range filters with the same directional preference but different simple cell orientations. Thus the long-range filter provides the extra degree of freedom that enables Level 5 cells to function as *direction* cells, rather than as *orientation* cells. Cells in cortical area MT can also respond to a range of orientations that are not perpendicular to their preferred direction-of-motion (Albright, 1984;

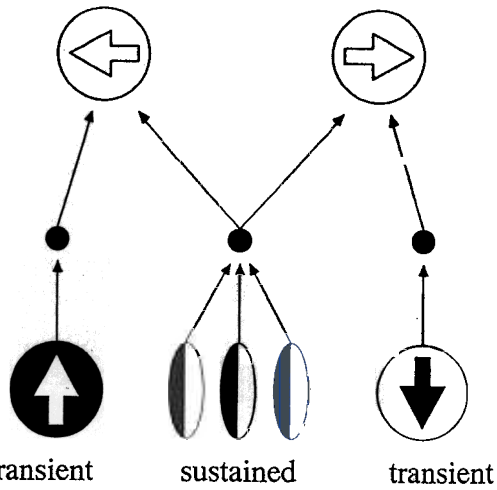


FIGURE 7. Transient cell gating of sustained cell activities to produce directionally sensitive responses. The short-range filter, which is constructed from like-oriented simple cells, responds ambiguously to a contrast pattern (dark–light in the illustration) moving either to the right or to the left. This ambiguity of motion direction is eliminated by gating the short-range filter response with either a transient on-cell response (to produce a left-motion signal) or a transient off-cell response (to produce a right-motion signal). [Reprinted with permission from Grossberg and Rudd (1992).]

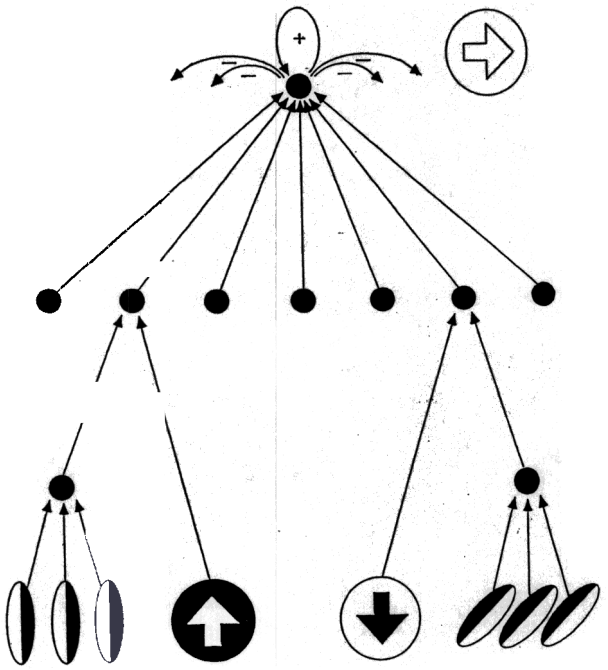


FIGURE 8. Combination of like direction-of-motion activities across space by a long-range Gaussian filter. Local direction-sensitive responses of opposite direction-of-contrast, over a range of orientations, are gated by transient cells of opposite types to produce like direction-of-motion signals. These local signals are combined by a long-range Gaussian spatial kernel to produce a spatially broad pattern of activity across the Level 5 network. This broad pattern is then contrast enhanced by a competitive, or lateral inhibitory, interaction. The contrast enhancement restores positional information. [Reprinted with permission from Grossberg and Rudd (1992).]

Albright, Desimone & Gross, 1984; Maunsell & van Essen, 1983; Newsome, Gizzi & Movshon, 1983).

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 may only partially contrast-enhance its input pattern to generate a motion signal whose breadth across space increases with the breadth of its inducing pattern.

These model interactions can generate a continuously moving signal, called a G-wave (after the long-range Gaussian), in response to the discrete flashes of an apparent motion display. Upon offset of the first flash, its Level 4 cell activations began to decay. If the ISI between the flashes is not too long, these decaying signals can summate through the long-range Gaussian filter with growing signals from the second flash that code the same direction. As the local motion signals at the first flash weaken and the local motion signals at the second flash strengthen, the peak value of activity among the Level 5 cells, which adds all these signals through the Gaussian filter, continuously shifts from the location of the first flash to the location of the second flash. Grossberg and Rudd (1989, 1992) correlated properties of this traveling peak of activity with properties of many apparent motion phenomena, including beta motion, gamma motion, delta motion, split motion, Ternus motion, and reverse-contrast Ternus motion.

Grossberg and Rudd (1992) also suggested that the inverse relationship between lower ISI thresholds and stimulus duration, as in Fig. 2, is due to a lag in the response time of the Level 3 transient cells. MOC Filter model transient cells respond more quickly to the offset of a long duration flash than to the offset of a short duration flash. Breitmeyer (1984) reviewed studies of transient cells that are consistent with this property. Moreover, keeping the ISI constant and increasing the spatial separation of apparent motion stimuli produces weaker G-waves at Level 5. This property clarifies why the lower ISI threshold increases and the upper ISI threshold decreases as a function of spatial separation.

This paper shows how to overcome a processing limitation of the MOC filter pathways that model luminance-based V1 → MT interactions. In particular, the MOC Filter, by itself, cannot generate motion between stimuli defined by other stimulus characteristics, such as the emergent boundary segmentations that help to define many visual form percepts in response to textured and shaded images. Apparent motion of illusory contours nicely illustrates this human competence because the illusory contours are emergent boundaries and their apparent motion depends critically upon the motion system. Figure 9 shows simulation results of the MOC Filter to one-dimensional inputs [Fig. 9(a)] that mimic illusory contour inducers. Level 1, 2, and 3 cells respond to the spatiotemporal changes in luminance and combine to form local motion signals at Level 4 cells [Fig. 9(b)] at the onset and offset of each inducer pair. These signals converge on Level 5 cells to generate a weak G-wave from the first pair of inducers to the second pair of inducers. Compared to other apparent motion displays (described below), this G-wave is so weak as to be perceptually undetectable.

2.3. *Integration of form and motion processing*

Since psychophysical studies indicate that people can see strong apparent motion between illusory contours, some additional inputs, sensitive to illusory contours, must contribute to the MOC Filter, thereby allowing it to generate a G-wave in response to changing illusory contours and, more generally, to emergent percepts of form. The Static BCS, which does respond to illusory contours, provides these inputs. Moreover, as the following sections indicate, the persistence properties of the Static BCS signals account for many of the upper ISI threshold properties of apparent motion.

Figure 10 schematizes the Static BCS-to-MOC Filter connections that are simulated in this article. Oriented boundary signals from the Static BCS feed into like-oriented sustained cells and unoriented transient cells in the MOC filter that correspond to the same retinal

location. This BCS–MOC Filter pathway, which models a V2 → MT pathway *in vivo*, renders the MOC Filter sensitive to spatiotemporal changes in form as well as to spatiotemporal changes in luminance. This additional sensitivity allows it to generate apparent motion signals in response to illusory contours and other boundary segmentations.

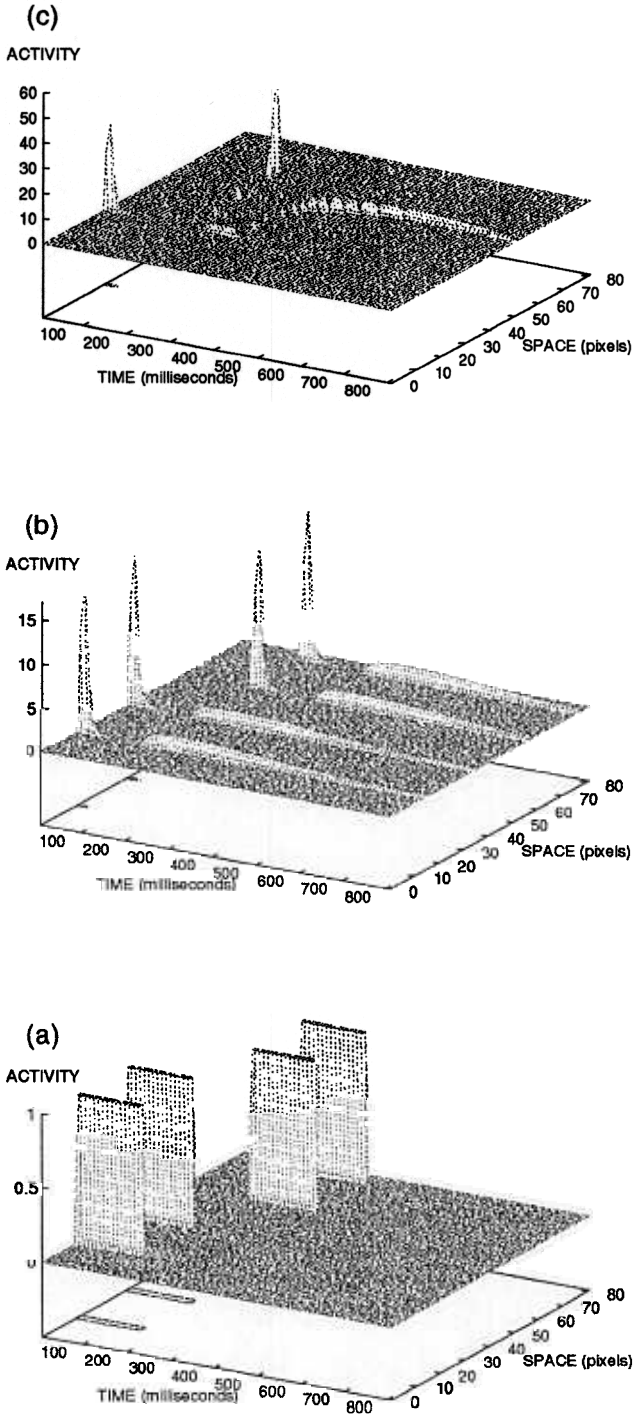


FIGURE 9. Computer simulation of how the MOC Filter without BCS input responds to an illusory contour apparent motion display. (a) Input stimuli. (b) Local motion signals at Level 4 cells. The only responses are at the location of the inducing stimuli. There is no response along the perceived illusory contour. (c) Response of Level 5 global motion cells. A weak (subthreshold) motion signal travels from the location of the first set of inducers to the second set of inducers. This corresponds to perceived motion of the inducing stimuli and not of the illusory contour.

More precisely, signals resulting from the second competitive stage feed into the Level 2 (sustained) and Level 3 (transient) cells of the MOC Filter at the same position. The MOC Filter Level 2 equations are adjusted so that the sustained cells respond to direct luminance inputs and to inputs from the Static BCS. Similarly the Level 3 equations are adjusted so that the transient cells respond to changes in luminance and to changes in the inputs from the Static BCS. A luminance edge or an illusory contour could produce the Static BCS inputs. Inputs from illusory contours persist longer than luminance-defined inputs. In either case, the MOC Filter combines the sustained and transient cell outputs to produce local motion signals at Level 4. The local motion signals then contribute to Level 5 cells and can, given the correct image parameters, generate a G-wave between a pair of temporally displaced illusory contours.

3. SIMULATIONS OF FORM–MOTION INTERACTIONS

3.1. Simulation of illusory contour apparent motion

Figure 11 shows the results of simulating the Static BCS and MOC Filter interactions with illusory contour inducers. Figure 11(a) shows the inputs for a display presenting two sets of illusory contour inducers in sequence. Figure 11(b) shows the responses of Level 1 cells in the MOC Filter. These activities respond only at the location of the luminance increments. Figure 11(c) shows the responses of the BCS hypercomplex cells at the second competitive stage (see Figs 3 and 10), notably the illusory contour between the two luminance increments. The activities in Fig. 11(b, c) feed into the sustained cells at Level 2 and the transient cells at Level 3 of the MOC Filter, whose outputs are multiplied to generate local motion signals at Level 4. Figure 11(d) shows the responses of the Level 4 cells. The tall spikes indicate the onset of the luminous inducers. The smaller hills mark the offset of different parts of the illusory contour. These responses are pooled by the long-range filter to generate the Level 6 activities that are shown in Fig. 11(e). Due to the strong spatial competition between these cells, only one cell is active at a time. The location of the active cell shifts continuously from the first stimulus to the second stimulus during the apparent motion display. This demonstrates apparent motion of the illusory contour.

For fixed spatial separation, the strength of the G-wave depends on the stimulus duration and ISI of the display. A strong G-wave requires overlap between the BCS inputs to the first stimulus and the BCS inputs to the second stimulus. Thus, the strength of the G-wave, as a function of ISI, depends on the persistence of the BCS signals to the first stimulus and, as was noted above, persistence of BCS signals in response to illusory contours depends upon the duration of the illusory contour inducers.

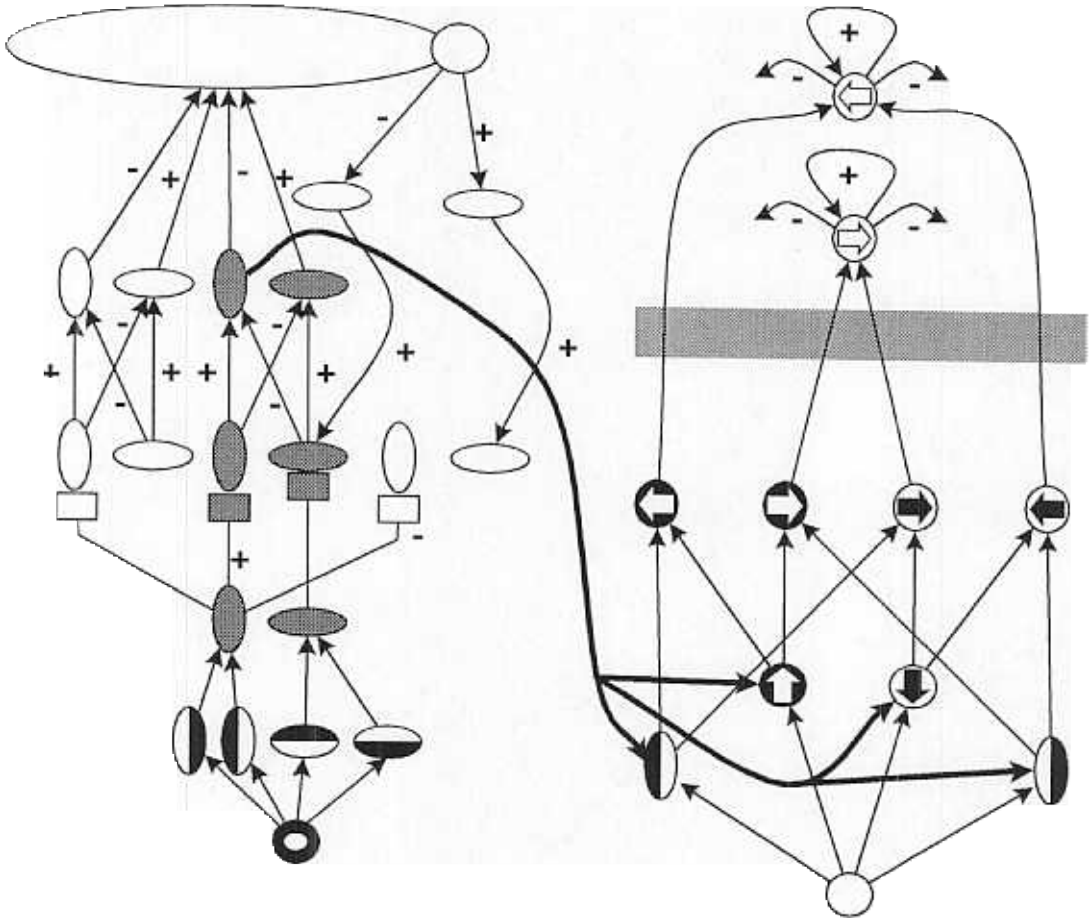


FIGURE 10. Model of form and motion integration. Oriented boundary signals in the BCS feed into like-oriented sustained cells and unoriented transient cells in the MOC Filter.

Figure 12 shows the strength of the G-wave for different combinations of stimulus duration and ISI. Also plotted is a threshold value. We assume that when the combinations of stimulus parameters create a G-wave with a strength above threshold, then the motion is observable. When the G-wave strength is below threshold, then subjects are assumed not to see it. The measure of G-wave strength we use is specific to these simulations. Other measures would produce similar results. See Appendix B for details. A notable property of Fig. 12 is that, as the stimulus duration increases from 50 to 100 msec, the intersection between the G-wave strength curve and the threshold shifts to a longer ISI; but as the stimulus duration increases still further, the intersection between the G-wave strength curve and the threshold shifts to shorter ISIs. This inverted U is qualitatively explained in Section 2.1.2. The ISIs that produce intersections in the strength and threshold curves identify the upper and lower ISI values for perceiving apparent motion. Figure 1(b) plots those threshold ISI values.

As noted in Section 1.2, the persistence of illusory contours shows a shape qualitatively similar to the ISI thresholds in Fig. 1. Contours that persist longer supply strong inputs to the motion system for longer durations so that greater ISIs continue to generate strong motion percepts. Figure 1(c) plots the upper ISI thresholds from the data and the simulation. The simulation thresholds fall between the thresholds of the two subjects. Moreover, the simulation thresholds are an inverted-U function of stimulus duration. The simulation of the BCS boundary signals to explain these apparent motion thresholds of Mather (1988) used *exactly* the same parameters and equations as in our previous study (Francis *et al.*, 1994) to explain the visual persistence data of Meyer and Ming (1988). In summary, the model properties responsible for integrating form and motion information explain percepts of illusory contour apparent motion (Ramachandran, 1985) by linking dynamic persistence properties of illusory contour form perception (Meyer & Ming, 1988) to the dynamic properties of apparent motion (Mather, 1988).

3.2. Simulation of interattribute apparent motion

A similar analysis explains the generation of percepts of interattribute apparent motion. Figure 13(a) shows inputs for an interattribute simulation. The first stimulus is a pair of illusory contour inducers and the second stimulus is a luminance edge. Figure 13(b) shows the activation of Level 1 cells at the MOC Filter. Figure 13(c) shows the boundary signals produced at BCS hypercomplex cells to the inducers and the luminous

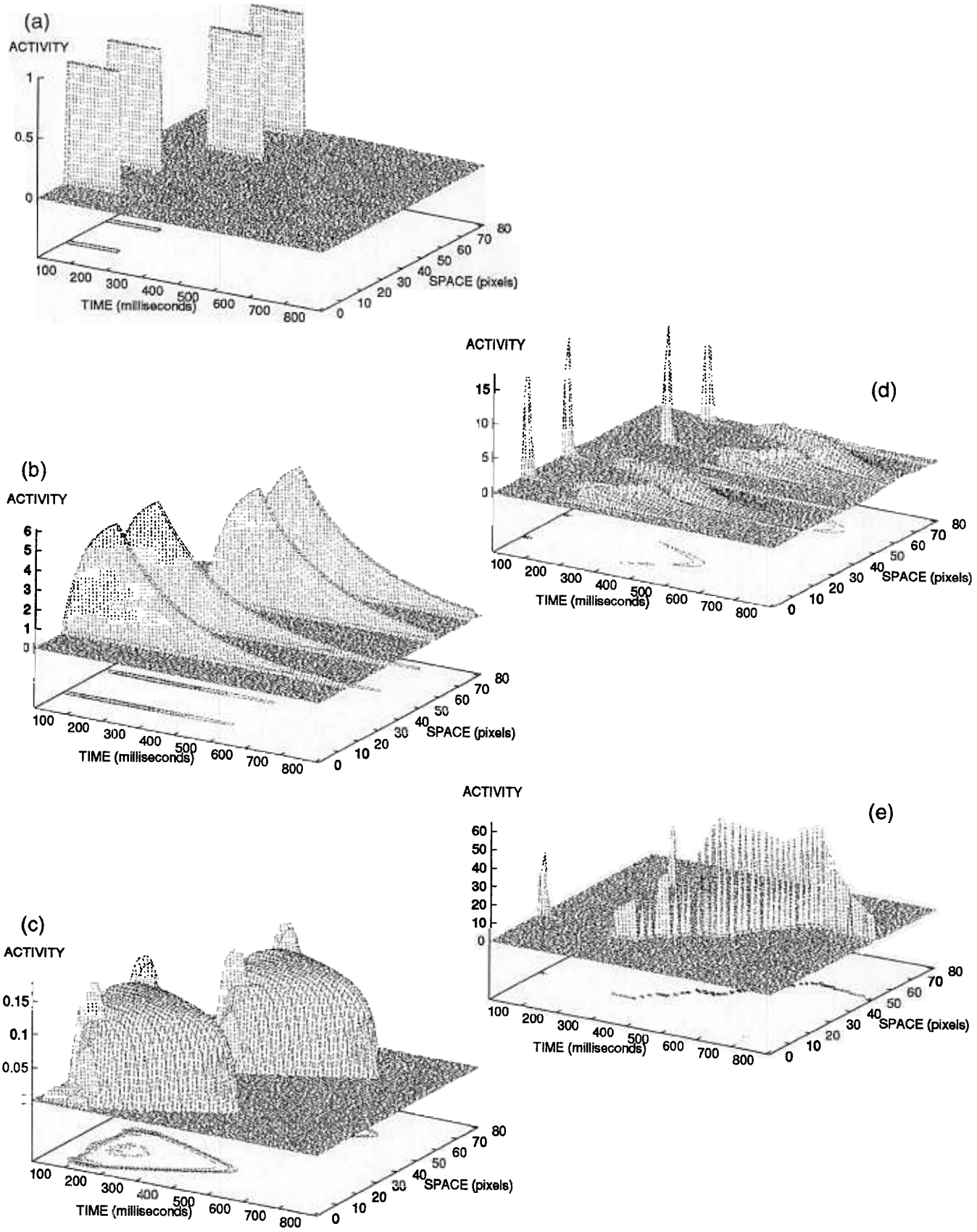


FIGURE 11. Computer simulation of how the MOC Filter with BCS input responds to an illusory contour apparent motion display. (a) The stimuli consists of two sets of illusory contour inducers. (b) Level 1 activities in the MOC Filter do not create illusory contours. (c) Boundary signals generated by the BCS create an illusory contour between the two inducers. (d) Local motion signals at Level 4 cells. The tall spikes are responses to the luminous inducers, while the smaller curves are produced by the illusory contours generated by the BCS. (e) Response of Level 5 global motion cells. The activity shifts continuously via a G-wave from the location of the first illusory contour to the location of the second illusory contour.

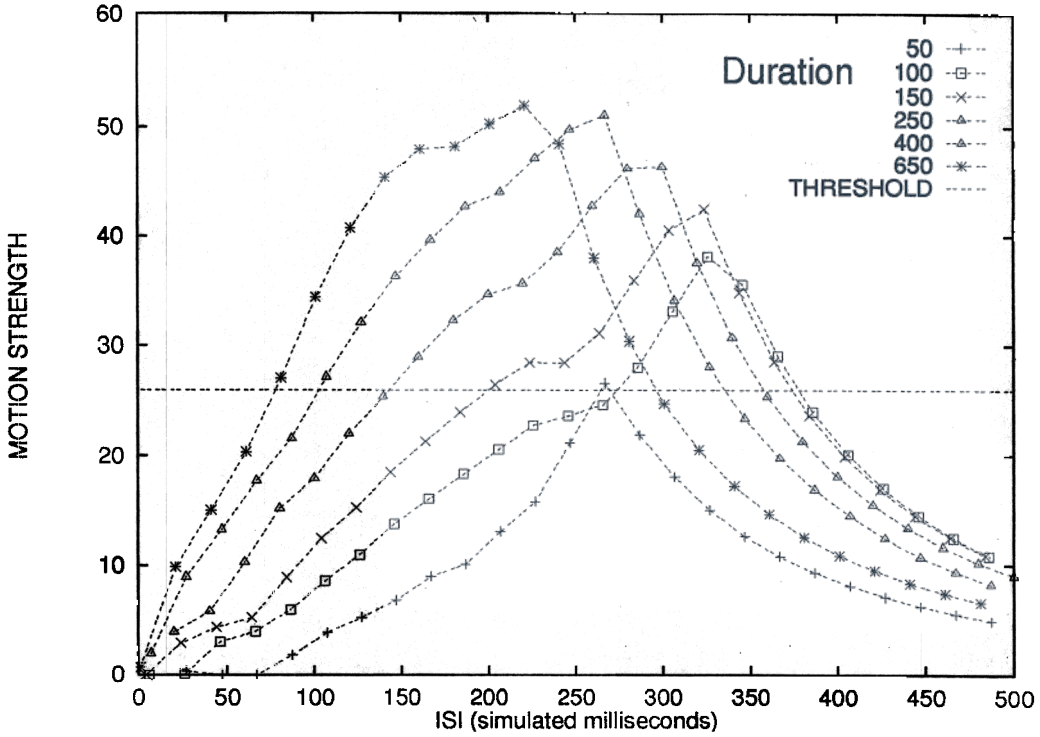


FIGURE 12. G-wave strength for illusory contours as stimulus duration and ISI varies. Intersections between G-wave strength curves and the threshold mark upper and lower ISI thresholds.

contour. The BCS generates an illusory contour between the inducers and a contour along the luminous inputs. Figure 13(d) shows the pooled response of Level 4 local motion cells that respond to rightward motion. [To better show the response at offset of the first stimulus, the larger activities at stimulus onset are sometimes beyond the range of the plot.] With form–motion interactions, the Level 4 cells respond to both the illusory contour and the luminance-based contour. These signals feed into the Level 5 global motion cells [Fig. 13(e)]. This plot, sampled with greater frequency than Fig. 11(e), shows that the activity of global motion cells shifts continuously from the location of the illusory contour to the location of the luminance-based contour. Figure 13(e) thus demonstrates apparent motion between stimuli of different attributes.

While not simulated here, the BCS model segments stimuli of many different attributes, including illusory contours (Gove *et al.*, 1994b; Grossberg & Mingolla, 1985a, b, 1987), textures (Cruthirds *et al.*, 1993; Grossberg & Mingolla, 1985b, 1987), surface brightness and color (Grossberg & Mingolla, 1985a; Grossberg & Todorović, 1988), and stereopsis (Grossberg & Marshall, 1989; Grossberg, 1994; McLoughlin & Grossberg, 1994). The integration of form and motion offers a consistent explanation of many types of interattribute apparent motion by suggesting that these segmentations feed into the MOC Filter, which generates the apparent motion percept.

3.3. Simulation of Korte's laws

The previous two sections demonstrate how integration of form and motion information helps to explain dynamic properties of apparent motion that depend on visual form. This section shows that the dynamic properties of form perception are also relevant to stimuli that do not obviously require form processing for motion detection.

Figure 14(a) shows the inputs for a standard apparent motion display, the stimuli are luminance edges separated in space and time. Figure 14(b) shows the MOC Filter Level 1 activations, and Fig. 14(c) shows the boundary segmentation generated in the BCS. Figure 14(d) shows the pooled responses of local rightward motion cells at Level 4 in the MOC Filter. The response to the first stimulus weakens as the response to the second stimulus grows. Recall from Section 2.2 that our apparent motion signal, or G-wave, adds the Gaussianly filtered decay of the first response during the growth of the second response. Figure 14(e) plots the activity of Level 5 global motion cells in the MOC Filter. [To better show the apparent motion signals in the contour plot, any Level 5 activities >100 are set equal to 100.] The activity among these cells shifts continuously from the location of the first stimulus to the location of the second stimulus, indicating a percept of apparent motion.

Korte's laws, summarized in Fig. 2(a), describe how upper and lower ISI thresholds vary inversely with duration, and how the range of ISIs that produce

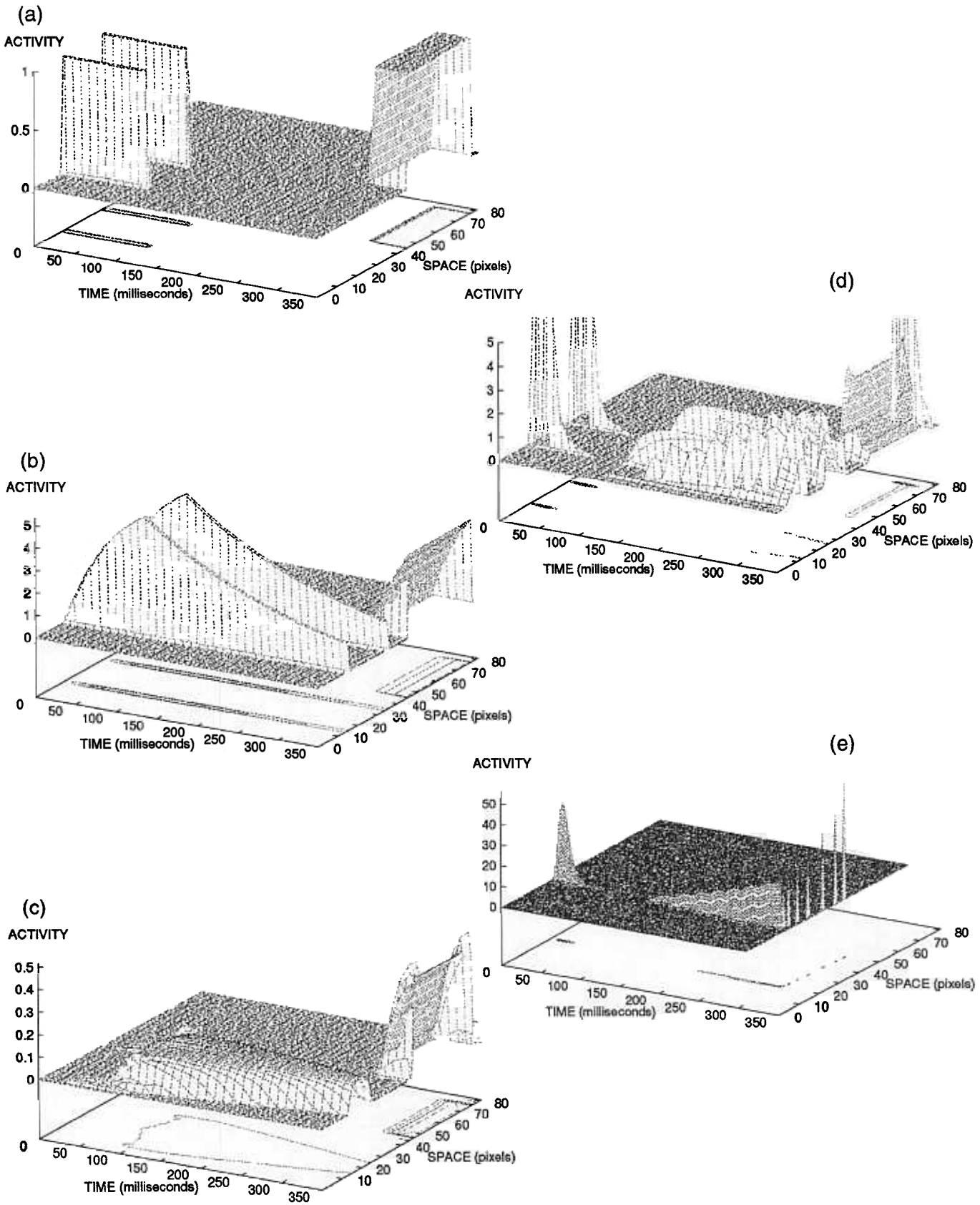


FIGURE 13. Computer simulation of interattribate apparent motion. (a) Stimulus input consists of illusory inducers followed by a luminous stimulus. (b) Level 1 MOC Filter responses. (c) Boundary signals generated by the BCS create an illusory contour between the two inducers. Boundary signals also respond to the luminance edge. (d) Level 4 local motion signals generated by illusory inducers and the luminous input. Offset of the boundary signals generated by the illusory inducers produces local motion signals cells that overlap with the local motion signals produced by the onset of the luminous stimulus. (e) Activity of Level 5 global motion cells shifts continuously via a G-wave from the location of the illusory contour to the location of the luminous stimulus.

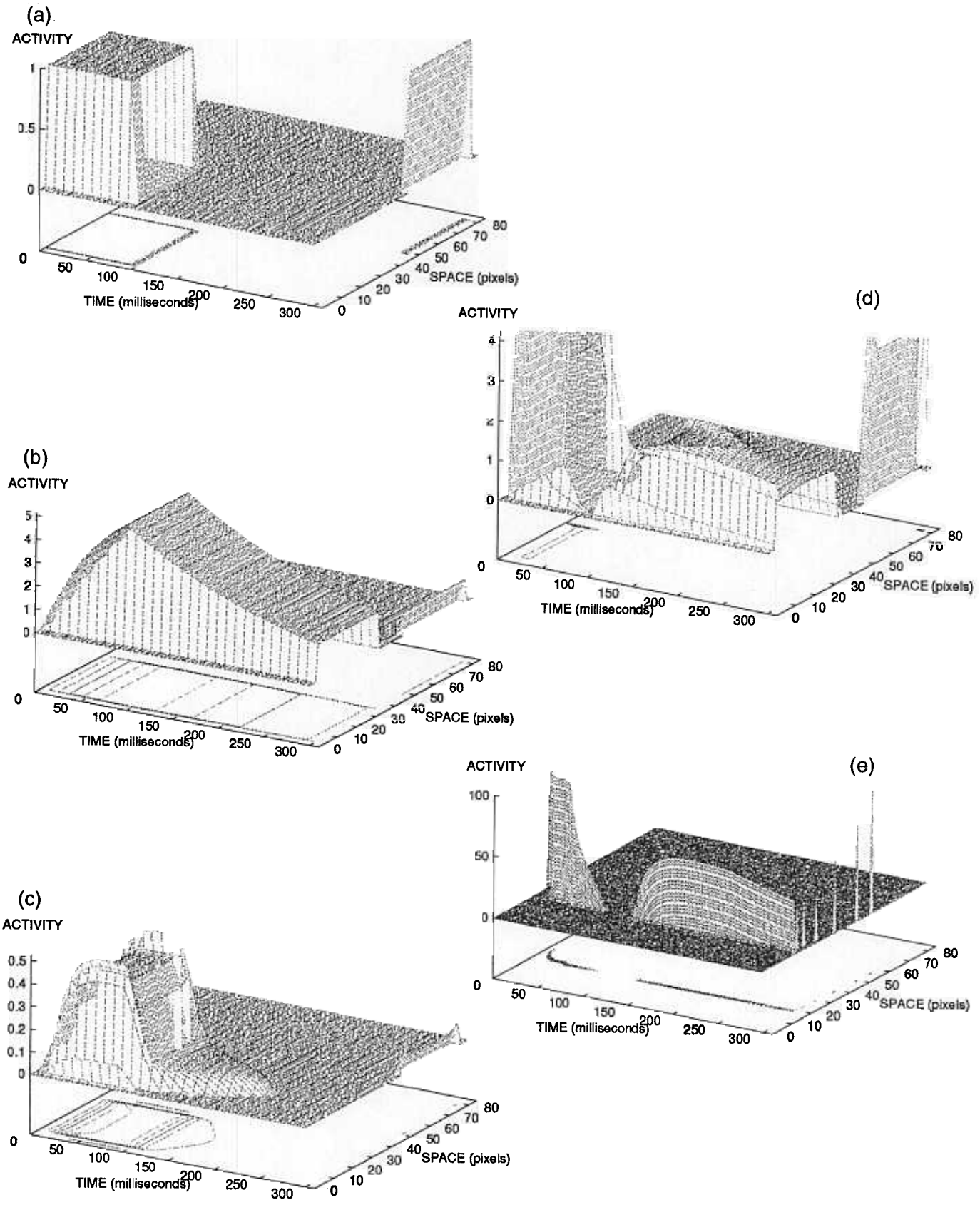


FIGURE 14. Computer simulation of apparent motion between luminous stimuli. (a) Two luminous inputs separated in space and time. (b) Level 1 MOC Filter responses. (c) Boundary signals generated by the BCS. (d) Level 4 local motion signals. Offset of the luminance and BCS inputs produces local motion signals for the first stimulus. These local motion signals temporally overlap with the local motion signals generated by the onset of the second luminous stimulus. (e) Activity of Level 5 global motion cells shifts continuously via a G-wave from the location of the first stimulus to the location of the second stimulus.

apparent motion narrows as a function of the distance between the stimuli. Grossberg and Rudd (1992) related the inverse dependence of ISI threshold on duration to a model circuit, called a *gated shunting cascade*, whose off-cells respond sooner after offset of long duration stimuli than of short duration stimuli. Such a circuit is naturally embedded in the MOC Filter design, as well as the Static BCS design. All that is required is an opponent process in which there are at least two stages of cell processing by a membrane equation (also called a shunting equation) followed by habituative transmitter gates. In the MOC Filter, these processing stages occur during transient on-cell and off-cell cell processing; see the Appendix. In the Static BCS, these stages are embedded in the gated dipoles in Fig. 3 that are used to reset boundary segmentations.

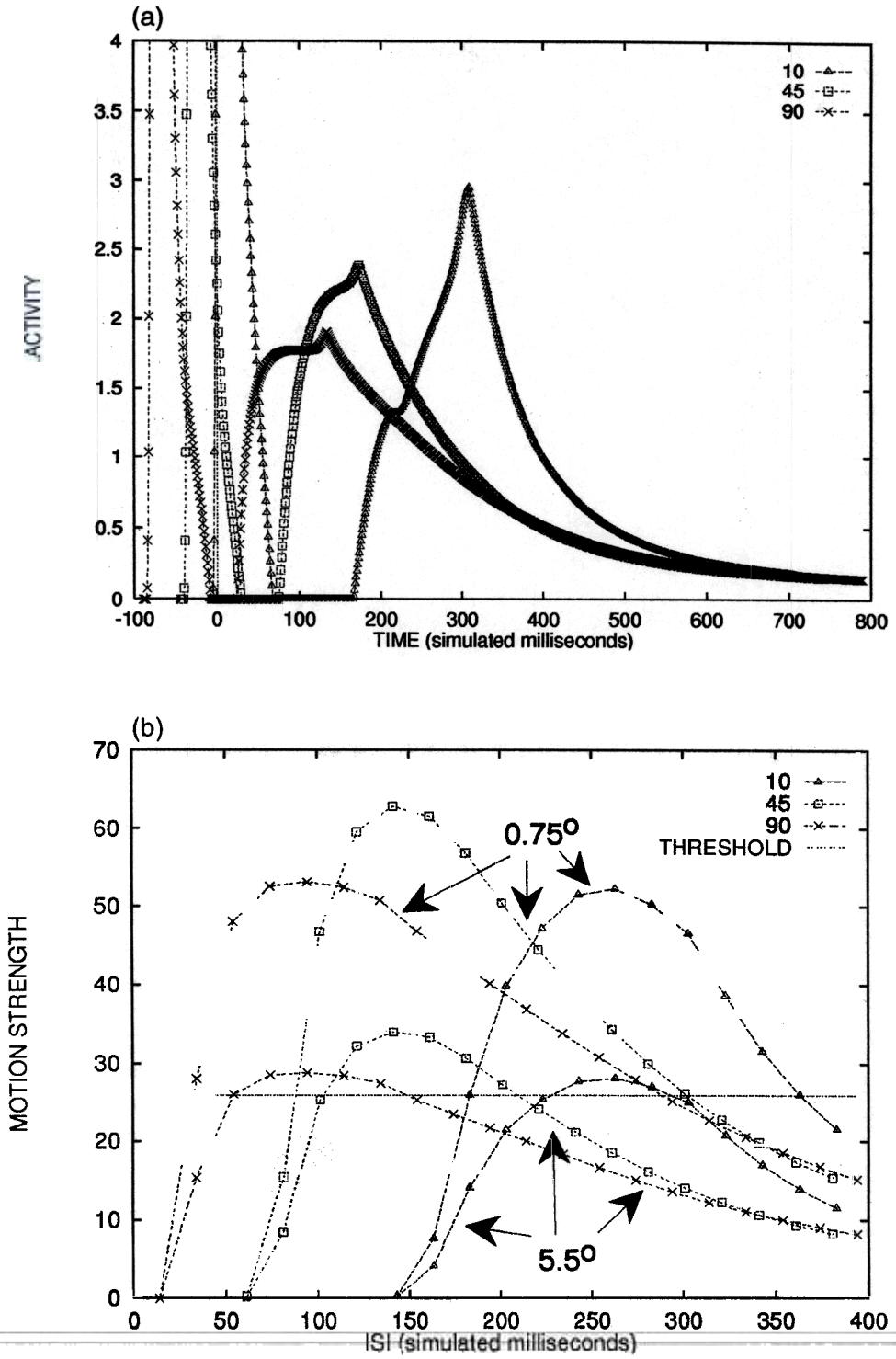


FIGURE 15. (a) Net activity of the rightward local motion cells at pixel 20 (center of the first stimulus) for stimulus durations of 10, 45, and 90 msec. Local motion cells respond more quickly and disappear more quickly as stimulus duration increases. (b) G-wave strength for apparent motion displays as a function of ISI, stimulus duration, and spatial separation. Intersection of each G-wave strength curve and the threshold line indicates the threshold ISI value for reporting apparent motion.

The main idea of the gated shunted cascade is that the *amplitude* of activation in the first stage of processing is larger in response to longer inputs; hence, due to shunting dynamics, both the amplitude and the *rate* of activation in the second stage of processing are larger in response to longer inputs. When second stage activities are gated by habituative transmitters, the *rate* of transmitter habituation is faster in response to longer inputs, so both onset and offset of gated responses are faster to longer inputs. Our main task herein is to convert this insight into a computer simulation of Korte's laws. Since Grossberg and Rudd (1992) did not analyze the effects of boundary persistence on ISI thresholds, part of this task is to show how the combined effects of form–motion interactions can generate both upper and lower ISI threshold curves.

Figure 15(a) plots the response of the rightward local motion cell at pixel 20, which responds to the onset and offset of the first stimulus, with time translated so that each stimulus offset occurs at time zero. To better display the responses produced by stimulus offset, the onset responses sometimes extend beyond the range of the Y-axis. The response of this cell to stimulus offset shifts to the right (greater lag) as stimulus duration decreases. These simulation results show that the properties of the gated shunting cascade described in Grossberg and Rudd (1992) also exist in the model of form–motion interactions described here.

The fastest response to the offset of the first stimulus establishes the lower ISI threshold. Likewise, the longest duration response after offset of the first stimulus establishes the upper ISI threshold. Figure 15(a) shows that shorter duration stimuli have longer persisting responses for this cell. This property is due to the persistence of boundary signals in the BCS, which is inversely related to stimulus duration, as discussed in Section 2. For many stimuli, persisting BCS boundary signals last much longer than the inputs from the direct MOC Filter luminance pathway, so the dynamic properties of form processing establish the upper ISI thresholds of apparent motion.

The lag time of the local motion responses of Level 4 cells to the first stimulus offset, the duration of these responses, and the spatial separation between the two stimuli all contribute to the strength of the G-wave generated in the MOC Filter. Figure 15(b) plots the strength of the G-wave, generated from offset signals of the first stimulus to onset signals of the second stimulus, as a function of ISI, spatial separation, and flash duration. As Fig. 15(a) predicts, the G-wave strength curve shifts toward smaller ISI values as stimulus duration increases. Also, for fixed stimulus duration and ISI, the G-wave strength decreases as spatial separation increases from 15 to 110 pixels (corresponding to 0.75 and 5.5 visual deg), since the overlap between off-responses and on-responses, as registered through the long-range Gaussian filter, decreases with distance. Also plotted is the G-wave strength threshold. When the strength of the G-wave is below threshold, we assume that subjects do not perceive motion.

The intersections of the motion strength threshold with a motion strength curve identify the ISI thresholds. Figure 2(b) plots the points of intersection and shows that the model captures the key qualitative properties of the classical Neuhaus (1930) data. Both upper and lower ISI thresholds are inversely related to stimulus duration and the range of ISIs that produce apparent motion percepts narrows as spatial separation increases. Lower ISI thresholds increase with spatial separation, upper ISI thresholds decrease with spatial separation, and the maximum separation of the upper and lower ISI thresholds is quantitatively correct (350 msec). In summary, the classical but paradoxical parametric properties of Korte's laws may be derived from form-motion interactions that we hypothesize to be mediated by interactions between $V1 \rightarrow MT$ and $V1 \rightarrow V2 \rightarrow MT$ pathways.

4. PREDICTIONS

4.1. Psychophysical prediction

In all cases where the duration of the Static BCS response to a stimulus establishes the upper ISI threshold, then the stimulus properties that favor longer visual persistence should also favor longer upper ISI thresholds for apparent motion. This relationship predicts a result that can support the role of the BCS–MOC Filter pathway in establishing upper ISI thresholds. Meyer, Lawson and Cohen (1975) showed that adaptation to an oriented grating influences the persistence of a subsequent test grating in an orientation-specific manner. When the orientation of the adaptation and test grating are orthogonal, persistence of the test grating increases relative to the no-adaptation case. Francis *et al.* (1994) simulated this property with the BCS model.

If persistence of boundary signals sets the upper ISI threshold of apparent motion, then, other things being equal, adaptation to a grating should increase the upper ISI threshold for apparent motion of an orthogonally oriented grating. In the BCS, the increase in persistence is due to habituation of an oriented channel by the adaptation stimulus, followed by habituation of the orthogonal channel by the test stimulus. When the test stimulus shuts off, a weaker-than-usual reset rebound occurs in the habituated pathways after competition takes place between the corresponding orthogonally oriented hypercomplex cells at the second competitive stage (Fig. 3). The MOC Filter does not include a stage of competition between orthogonal orientations. Instead, competition occurs between opposite *directions* of motion, which differ by 180 deg, not 90 deg, thereby creating motion contrast cells (Grossberg, 1991). Thus, the luminance-based pathway of the MOC Filter should not contribute to a change in the upper ISI threshold after adaptation. However, BCS–MOC Filter interactions could explain the predicted change in the upper ISI threshold. More generally, stimulus features that change the duration of visual persistence should similarly affect upper ISI thresholds of apparent motion.

4.2. *Neurophysiological prediction*

Grossberg (1991) suggested that the luminance-based pathways of the MOC Filter exist in the brain as connections from area V1 to MT, and that signals from the Static BCS to the MOC Filter exist in the brain as a pathway from area V2 to MT (or area V4 to MT). Grossberg (1991) also described a method of testing whether the V2 to MT pathway plays the role suggested in this article. An experimenter could train a monkey to respond when it sees apparent motion of illusory contours. A (reversible) lesion of area V2 or the V2 → MT pathway should abolish the percept and the response.

Additional data links may be derived from the Grossberg (1991) analysis of how the fine stereo computations of the parvocellular cortical stream could be used to sharpen the coarser stereo computations of the magnocellular cortical stream, and thereby achieve more precise depth estimates for moving forms. This analysis suggested that the same V2 → MT pathway that is modeled above to simulate form–motion interactions in response to planar stimuli may also play a key role in segmenting moving forms in depth. Several psychophysical experiments are consistent with this suggestion. For example, experiments by Corbin (1942) and Attneave and Block (1973) indicated that three-dimensional information can influence the quality of apparent motion. Subsequent experiments have supported the hypothesis that structure-from-motion can be influenced by stereopsis constraints (Dosher, Sperling & Wurst, 1986; Green & Odom, 1986; Mowafy, 1990; Nawrot & Blake, 1989). Demonstrations in which illusory contours influence percepts of motion are particularly informative (Nawrot & Blake, 1989), since they emphasize that the perceptual units that define the forms undergoing motion need to be actively constructed before the movement directions and speeds of their illusory contours can be determined. These experiments suggest a host of experiments for linking properties of the V1 → MT and V1 → V2 → MT pathways to interacting percepts of depth, persistence, and motion.

5. CONCLUDING REMARKS

Table 1 delineates the role of the Static BCS, the MOC Filter and their interactions in explaining the data discussed in this paper. The temporal characteristics of BCS form processing simulate the persistence characteristics of illusory and luminous contours, while BCS spatial properties allow it to segment a wide variety of stimulus attributes. The MOC Filter is insensitive to many types of form-based motion, but its dynamics do help to explain the characteristics of lower ISI thresholds to variable durations and spatial separations. The form–motion interactions link the properties of the BCS and the MOC Filter to data on apparent motion of illusory contours, interattribute motion, Korte's laws, and neurophysiological structures between the parvocellular and magnocellular cortical streams.

The model hereby clarifies the need (Chubb & Sperling, 1991; Cavanagh & Mather, 1989) for motion detection based both upon direct luminance inputs and upon direct form-based inputs. Moreover—and here the model differs significantly from other theories of these interactions—the dynamic aspects of the form-based inputs determine many temporal aspects of perceived apparent motion. Boundary signals generated by luminance-based stimuli tend to persist for shorter lengths of time than boundary signals generated by illusory contours. The latter stimuli lead to fewer reset signals to inhibit the reverberating circuits of the BCS. Since the MOC Filter depends on the persistence of BCS inputs to establish the upper temporal limit of motion, the upper ISI threshold follows the persistence of the first stimulus. The persistence of a luminance-based first flash is inversely related to flash duration, while the persistence of an illusory contour first rises and then falls with increases in stimulus duration. As Francis *et al.* (1994) showed, these are robust properties of the BCS dynamics within a broad parameter range. Of particular interest are the new relationships proposed by the model between the persistence of static form percepts and the quality of apparent motion percepts. The model hereby links long

TABLE 1. Summary of the roles each model and their interactions play in explaining the data simulated in this paper

	Model		
Data	BCS	MOC Filter	Form-motion interactions
	Generation; persistence	Apparent motion of dynamic inputs	Apparent motion of illusory contours; threshold data
Interattribute motion	Boundary segmentation across many attributes	Apparent motion of dynamic inputs	Apparent motion across attributes
Korte's laws	Persistence of form	Lower ISI thresholds, role of spatial separation	Upper ISI thresholds
Neurophysiology	V1 → V2	V1 → MT	V2 → MT

known, but poorly understood, parametric psychophysical properties such as Korte's laws to directly measurable neural mechanisms. These neural mechanisms, in turn, may be understood in terms of concepts about how the static form perception system compensates for variable illumination and fills-in surface representations within ecologically useful boundary segmentations (Grossberg, 1987a, 1994; Grossberg *et al.*, 1989) and how the motion perception system generates motion segmentations that compute unambiguous directional signals that overcome aperture ambiguities and are independent of direction-of-contrast (Grossberg & Mingolla, 1993; Grossberg & Rudd, 1992).

While the current simulations are restricted to illusory contours and luminance-based stimuli, more powerful versions of the Static BCS have been shown to generate boundaries for a much larger class of perceptual forms, including forms defined by texture gradients (Cruthirds *et al.*, 1993; Grossberg & Mingolla, 1985b) and stereo gradients (Grossberg, 1994; McLoughlin & Grossberg, 1994). While the current discussion has emphasized the role of static form inputs to motion inputs, Grossberg and Mingolla (1993) suggested how the MOC Filter inputs to a grouping network that is analogous to the hypercomplex–bipole cell network of the Static BCS. This extended Motion BCS model can account for various percepts of form derived from motion, including percepts of motion capture that help to solve the global aperture problem. It is also of interest that the BCS model passes a test that every plausible model of biological vision needs to face: it is not a toy model. The BCS is being used to segment complex imagery derived from a variety of artificial sensors, including synthetic aperture radar, laser radar, multispectral infrared, and magnetic resonance sensors (Cruthirds, Gove, Grossberg, Mingolla, Nowak & Williamson, 1992; Grossberg *et al.*, 1994; Lehar, Worth, and Kennedy, 1990; Waxman, Seibert, Bernardon & Fay, 1993).

REFERENCES

Albright, T. D. (1984). Direction and orientation selectivity of neurons in visual area MT of the Macaque. *Journal of Neurophysiology, 52,* 1106–1130.

Albright, T., Desimone, R. & Gross, C. (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. & Ramachandran, V. S. (1987). Visual inertia in apparent motion. *Vision Research, 27,* 755–764.

Arrington, K. (1994). The temporal dynamics of brightness filling-in. *Vision Research, 34,* 3371–3387.

Attneave, F. & Block, G. (1973). Apparent movement in tridimensional space. *Perception & Psychophysics, 13,* 301–307.

Beck, J., Prazdny, K. & Rosenfeld, A. (1983). A theory of textural segmentation. In Beck, J., Hope, B. & Rosenfeld, A. (Eds), *Human and machine vision.* New York: Academic Press.

Bowen, R., Pola, J. & Matin, L. (1974). Visual persistence: Effects of flash luminance, duration and energy. *Vision Research, 14,* 295–303.

Braddick, O. (1974). A short range process in apparent motion. *Vision Research, 14,* 519–527.

Breitmeyer, B. (1984). *Visual masking: An integrative approach.* New York: Oxford University Press.

Burr, D. C. (1980). Motion smear. *Nature, 284,* 164–165.

Burr, D. C., Ross, J. & Morrone, M. C. (1986). Smooth and sampled motion. *Vision Research, 26,* 643–652.

Carpenter, G. A. & Grossberg, S. (1981). Adaptation and transmitter gating in vertebrate photoreceptors. *Journal of Theoretical Neurobiology, 1,* 1–42.

Castet, E. (1994). Effect of the ISI on the visible persistence of a stimulus in apparent motion. *Vision Research, 34,* 2103–2114.

Cavanagh, P. & Mather, G. (1989). Motion: The long and short of it. *Spatial Vision, 4,* 103–129.

Cavanagh, P., Arguin, M. & von Grünau, M. (1989). Interattribute apparent motion. *Vision Research, 29,* 1197–1204.

Chey, J., Grossberg, S. & Mingolla, E. (1994). Neural dynamics of motion processing and speed discrimination. Technical Report CAS/CNS-TR-94-030. Boston University, Boston, Mass.

Chubb, C. & Sperling, G. (1991). Texture quilts: Basic tools for studying motion-from-texture. *Journal of Mathematical Psychology, 35,* 411–442.

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

Corbin, H. H. (1942). The visual perception of grouping and apparent motion in visual space. *Archives of Psychology, 273,* 5–50.

Creutzfeldt, O., Kuhnt, U. & Benevento, L. (1974). An intracellular analysis of visual cortical neurons to moving stimuli: Responses in a co-operative neuronal network. *Experimental Brain Research, 21,* 251–274.

Cruthirds, D. R., Grossberg, S. & Mingolla, E. (1993). Emergent groupings and texture segregation. *Investigative Ophthalmology and Visual Science, 34,* 1237.

Cruthirds, D. R., Gove, A., Grossberg, S., Mingolla, E., Nowak, N. & Williamson, J. (1992). Processing of synthetic aperture radar images by the Boundary Contour System and Feature Contour System. *Proceedings of the International Joint Conference on Neural Networks, IV* (pp. 414–419). Piscataway, N.J.: IEEE Service Center.

DeYoe, E. & van Essen, D. (1988). Concurrent processing streams in monkey visual cortex. *Trends in Neuroscience, 11,* 219–226.

Dosher, D. A., Sperling, G. & Wurst, S. A. (1986). Tradeoffs between stereopsis and proximity luminance covariance as determinants of perceived 3D structure. *Vision Research, 26,* 793–990.

Francis, G., Grossberg, S. & Mingolla, E. (1994). Cortical dynamics of feature binding and reset: Control of visual persistence. *Vision Research, 34,* 1089–1104.

Gaudiano, P. (1992a). A unified neural model of spatiotemporal processing in X and Y retinal ganglion cells. I. Analytical results. *Biological Cybernetics, 67,* 11–21.

Gaudiano, P. (1992b). A unified neural model of spatiotemporal processing in X and Y retinal ganglion cells. II. Temporal adaptation and simulation of experimental data. *Biological Cybernetics, 67,* 23–34.

Gove, A. N., Grossberg, S. & Mingolla, E. (1994a). A link between brightness perception, illusory contours, and binocular corticogeniculate feedback. *Investigative Ophthalmology and Visual Science, 35,* 1437.

Gove, A. N., Grossberg, S. & Mingolla, E. (1994b). Brightness perception, illusory contours, and corticogeniculate feedback. Technical Report CAS/CNS-TR-94-033. Boston University, Boston, Mass. *Visual Neuroscience.* In press.

Graham, N., Beck, J. & Sutter, A. (1992). Nonlinear processes in spatial-frequency channel models of perceived texture segregation: Effects of sign and amount of contrast. *Vision Research, 32,* 719–743.

Green, M. & Odom, J. V. (1986). Correspondence matching in apparent motion: Evidence for three-dimensional spatial representation. *Science, 233,* 1427–1429.

Grossberg, S. (1972). A neural theory of punishment and avoidance: II. Quantitative theory. *Mathematical Biosciences, 15,* 253–285.

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

Grossberg, S. (1976). Adaptive pattern classification and universal recoding, II: Feedback, expectation, olfaction, and illusions. *Biological Cybernetics, 23*, 187–202. Reprinted in *Studies of mind and brain*. Boston, Mass.: Reidel Press.

Grossberg, S. (1982). *Studies of mind and brain*. Dordrecht: Kluwer.

Grossberg, S. (1984). Outline of theory of brightness, color, and form perception. In Degreef, E. & van Buggenhaut, J. (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*, 97–116.

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

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. (1994). 3-D vision and figure–ground separation by visual cortex. *Perception & Psychophysics, 55*, 48–120.

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, and Image Processing, 37*, 116–165.

Grossberg, S. & Mingolla, E. (1993). Neural dynamics of motion perception: Direction fields, apertures, and resonant grouping. *Perception & Psychophysics, 53*, 243–278.

Grossberg, S. & Rudd, M. (1989). A neural architecture for visual motion perception: Group and element apparent motion. *Neural Networks, 2*, 421–450.

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

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.

Grossberg, S. & Wyse, L. (1991). Invariant recognition of cluttered scenes by a self-organizing ART architecture: Figure–ground separation. *Neural Networks, 4*, 723–742.

Grossberg, S., Mingolla, E. & Nogueira, C. A. M. (1993). Computation of first order and second order motion by a model of magnocellular dynamics. *Investigative Ophthalmology and Visual Science, 34*, 1029.

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

Grossberg, S., Mingolla, E. & Williamson, J. (1994). Synthetic aperture radar processing by a multiple scale neural system for boundary and surface representation. Technical Report CAS/CNS-TR-94-001. Boston University, Boston, Mass. *Neural Networks*. In press.

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 Srastyán, E. & Molnár, P. (Eds), *Advanced physiological science, sensory functions* (Vol. 16). New York: Pergamon Press.

von der Heydt, R., Peterhans, E. & Baumgartner, G. (1984). Illusory contours and cortical neuron responses. *Science, 244*, 1260–1262.

von Grünau, M. (1979). The involvement of illusory contours in stroboscopic motion. *Perception & Psychophysics, 25*, 205–208.

Kolers, P. (1972). *Aspects of motion perception*. Oxford: Pergamon Press.

Korte, A. (1915). Kinematoskopische Untersuchungen. *Zeitschrift für Psychologies*, 194–296.

Lchar, S., Worth, A. J. & Kennedy, D. N. (1990). Application of the boundary contour/feature contour system to magnetic resonance brain scan imagery. In *Proceedings of the International Joint Conference on Neural Networks*, San Diego, Calif. Vol. 1, pp. 435–440.

Levitt, J., Kiper, D. & Movshon, J. (1994). Receptive fields and functional architecture of macaque V2. *Journal of Neurophysiology, 71*, 2517–2542.

Logothetis, N. K., Shiller, 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.

MacKay, D. M. (1957). Moving visual images produced by regular stationary patterns. *Nature, 180*, 849–850.

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.

Mather, G. (1988). Temporal properties of apparent motion in subjective figures. *Perception, 17*, 729–736.

Maunsell, J. & van Essen, D. (1983). Response properties of single units in middle temporal visual area of the macaque. *Journal of Neurophysiology, 49*, 1127–1147.

McLoughlin, N. & Grossberg, S. (1994). How are monocularly viewed surfaces perceived in depth during DaVinci Stereopsis. *Investigative Ophthalmology and Visual Science, 35*, 2110.

Meyer, G. & Ming, C. (1988). The visible persistence of illusory contours. *Canadian Journal of Psychology, 42*, 479–488.

Meyer, G., Lawson, R. & Cohen, W. (1975). The effects of orientation-specific adaptation on the duration of short-term visual storage. *Vision Research, 15*, 569–572.

Mikami, A., Newsome, W. & Wurtz, R. (1986a). Motion selectivity in Macaque visual cortex. I. Mechanisms of direction and speed selectivity in extrastriate area MT. *Journal of Neurophysiology, 55*, 1308–1327.

Mikami, A., Newsome, W. & Wurtz, R. (1986b). Motion selectivity in macaque visual cortex. II. Spatiotemporal range of directional interactions in MT and V1. *Journal of Neurophysiology, 55*, 1328–1339.

Mowafy, L. (1990). Static depth cues do affect the perceived direction of motion. *Perception, 19*, 595–609.

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 A, 2*, 267–273.

Nawrot, M. & Blake, R. (1989). Neural integration of information specifying structure from stereopsis and motion. *Science, 224*, 716–718.

Neuhaus, W. (1930). Experimentelle untersuchung der scheinbewegung. *Archiv für die gesamte Psychologie, 75*, 315–458.

Newsome, W. T., Gizzi, M. S. & Movshon, J. A. (1983). Spatial and temporal properties of neurons in macaque MT. *Investigative Ophthalmology and Visual Science, 24*, 106.

Newsome, W., Mikami, A. & Wurtz, R. (1986). Motion selectivity in macaque visual cortex. III. Psychophysics and physiology of apparent motion. *Journal of Neurophysiology, 55*, 1340–1351.

Ögmen, H. (1993). A neural theory of retino-cortical dynamics. *Neural Networks, 6*, 245–273.

Ögmen, H. & Gagné, S. (1990). Neural network architectures for motion perception and elementary motion detection in the fly visual system. *Neural Networks, 3*, 487–506.

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

Pessoa, L., Mingolla, E. & Neumann, H. (1994). A multi-scale network model of brightness perception. Technical Report CAS/CNS-TR-94-017. Boston University, Boston, Mass. *Vision Research*. In press.

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.

Petersik, J. T., Pufahl, R., & Krasnoff, E. (1983). Failure to find an absolute retinal limit of a putative short-range process in apparent motion. *Vision Research, 23*, 1663–1670.

Ramachandran, V. (1985). Apparent motion of subjective surfaces. *Perception, 14*, 127–134.

Schiller, P. H., Logothetis, N. K. & Charles, E. R. (1990). Role of the color-opponent and broad-band channels in vision. *Visual Neuroscience, 5*, 321–326.

Sutter, A., Beck, J. & Graham, N. (1989). Contrast and spatial variables in texture segregation: Testing in a simple spatial-frequency channels model. *Perception & Psychophysics, 46*, 312–332.

Taylor, W. (1958). Visual organization. *Nature, 182*, 29–31.

Waxman, A. M., Seibert, M., Bernardon, A. M. & Fay, D. A. (1993). Neural systems for automatic target learning and recognition. *The Lincoln Laboratory Journal, 6*, 77–116.

Wilson, H., Ferrera, V. & Yo, C. (1993). A psychophysically motivated model for two-dimensional motion perception. *Visual Neuroscience, 9*, 79–97.

Acknowledgements—This material is based upon work by GF supported under a Natural Science Foundation Graduate Fellowship, the Air Force Office of Scientific Research (AFOSR 90-0175) and the Office of Naval Research (ONR N00014-91-J-4100). SG was partially supported by the Advanced Research Projects Agency (AFOSR 90-0083) and the Office of Naval Research (ONR N00014-91-J-4100). The authors wish to thank Diana Meyers and Robin Locke for their valuable assistance in the preparation of the manuscript.

APPENDIX A

This appendix describes the equations and methodology used to produce the simulation results. All the simulations used a single set of equations and parameters. The equations and parameters used to simulate the BCS processing of form are identical to those used in Francis *et al.* (1994).

BCS: form processing

The BCS model and parameters used in Francis *et al.* (1994) were used here too. Since these BCS simulations processed two-dimensional images and the MOC Filter equations (described below) process one-dimensional images, a one-dimensional cross-section of boundary signals (Level 6 in Francis *et al.*, 1994) corresponding to one stimulus edge were sampled every simulated msec and stored in a data file. This process was repeated for each stimulus type (real or illusory) and for each stimulus duration. These activities were used in the MOC Filter equations as described below and are plotted in Figs 11(c), 13(c), and 14(c).

MOC filter: motion processing

This section describes the equations used to simulate a version of the one-dimensional MOC Filter of Grossberg and Rudd (1989, 1992). Grossberg and Mingolla (1993) provide equations of the MOC Filter for two-dimensional simulations.

*Level 1: shunting response to input pattern.*The cells of the first array in the MOC Filter obey shunting equations and receive luminance inputs from the "retina". The activity x_i^1 of a Level 1 cell at position i obeys the differential equation:

$$\frac{dx_i^1}{dt} = -Ax_i^1 + (B - x_i^1)I_i, \tag{A1}$$

where parameter A sets the rate of passive decay, parameter B sets the maximum activity of the cell, and I_i is the input to the cell. More generally, the equation might include inhibitory input from nearby positions (Grossberg & Todorović, 1988) that would allow the cell to compensate for variable illumination. Since the simulations described below use a constant level of illumination, these inhibitory interactions are not necessary. The activities of this level are shown in Figs 11(b), 13(b), and 14(b).

*Level 2: oriented sustained cells.*The cells of Level 2 respond to spatial contrasts among the activities of Level 1 cells. The receptive field of each Level 2 cell is divided into left and right sides. One side receives excitatory inputs from Level 1 cells and the other side receives inhibitory inputs. The activity x_i^{2BD} of a Level 2 cell sensitive to a bright–dark (BD) change in luminance (from left to right) obeys an equation of the form:

$$\frac{dx_i^{2BD}}{dt} = -Cx_i^{2BD} + (D - x_i^{2BD})E(b_i(t) + [x_i^1 - x_{i+1}^1]^+), \tag{A2}$$

where parameter C sets the passive decay rate of the cell, parameter D sets the maximum activity of the cell, parameter E scales the value of oriented inputs, $[w]^+ = \max(w, 0)$ defines a threshold-linear half wave rectification, and $b_i(t)$ is the contribution of boundary signals from the BCS. A Level 2 cell at position i receives excitatory input from Level 1 if the Level 1 cell at the same position is active and the Level 1 cell one position to the right $(i + 1)$ is not as active. It also receives excitatory input if the BCS generates a boundary signal of the same orientation at the same location. Similarly, a cell sensitive to a dark–bright (DB) luminance change obeys the equation:

$$\frac{dx_i^{2DB}}{dt} = -Cx_i^{2DB} + (D - x_i^{2DB})E(b_i(t) + [x_i^1 - x_{i-1}^1]^+). \tag{A3}$$

These oriented sustained cells respond at spatial luminance edges of a flash and at locations of a BCS segmentation.

*Level 3: transient cells.*The cells of Level 3 respond to dynamic changes in Level 1 cell activities and BCS boundary signal inputs. At each spatial location, there exists an on-cell (responsive to increases in activity) and an off-cell (responsive to decreases in activity). The activities of these transient on-cells and off-cells are modeled with a series of stages. First the system computes the rectified time-derivative given by the shunting equation:

$$\frac{dy_i}{dt} = -Fy_i + (G - y_i)(b_i(t) + x_i^1), \tag{A4}$$

where parameter F sets the rate of passive decay and parameter G sets the maximum activity of y_i. Here y_i is a time average of the Level 1 input x_i^1. More generally, the equation could include excitatory inputs from a spatial range of Level 1 cells, thereby providing both a space and time average of the input [Grossberg & Rudd, 1992; equation (A5)]. The sequential levels of shunting equations from equations (1) to (4) create a *shunting cascade* (Grossberg & Rudd, 1992).

To generate transient responses, positive and negative half-wave rectifications of the time derivative are first performed independently by defining

$$w_i^+ = \left[\quad H \right]^+ \tag{A5}$$

and

$$\frac{dy_i}{dt}\Big] \tag{A6}$$

where parameters H and J are constant thresholds. The activity w_i^+ produces a non-zero response at input onset, and the activity w_i^- responds at input offset. Each w_i^+ could maintain its activity as long as its y_i input continues to grow. Likewise, w_i^- could remain active as long as y_i decreases. The w_i^+ and w_i^- are converted into responses that are transient under all conditions by being modulated with an activity-dependent habituative process. Several authors have applied the Grossberg (1976) model of early vision habituation to explain their data (Carpenter & Grossberg, 1981; Gaudiano, 1992a, b; Öğmen, 1993; Öğmen & Gagné, 1990). In this model, each input signal is multiplied by a transmitter gate that habituates, or is inactivated, at a rate proportional to the strength of the signal and accumulates at a constant rate to a finite target level. The strength of the transmitter gate for the transient on-cell at position i thus obeys the equation:

$$\frac{dz_i^+}{dt} - K(L - z_i^+), \quad Mw_i^+ z^+ \tag{A7}$$

Term $K(L - z_i^+)$ says that the transmitter accumulates to a maximum value of L at a rate K. Term $-Mz_i^+ w_i^+$ says that the transmitter habituates in proportion to the strength of the signal passing through the gate with parameter M scaling the interaction. A similar equation (replacing superscript $-$ for superscript $+$) exists for the transient off-cell.

The final transient on-cell response is the gated signal. The transient on-cell response is

$$x_i^{3+} = w_i^+ z_i^+, \tag{A8}$$

and the transient off-cell response is

$$x_i^{3-} = w_i^- z_i^- \tag{A9}$$

The outputs (8) and (9) represent an opponent process (on vs off) with gated shunting cascade properties.

*Level 4: sustained-transient gating.*Image contrasts moving in either of two directions that differ by 180 deg can cause maximal activation of a Level 2 sustained cell filter. Multiplicative gating of each Level 2 sustained output with a Level 3 transient on-cell or off-cell removes this ambiguity (see Fig. 4). For example, consider the output of a dark–light sustained cell. A dark-light image contrast moving to the right or to the left maximizes such a sustained cell's output. Multiplying this Level 2 output with a Level 3 transient on-cell output generates a Level 4 cell that responds maximally to motion to the left, as in the model of Marr and Ullman (1981). Multiplying it with a Level 3 off-cell output generates a Level 4 cell that responds maximally to motion to the right.

In the one-dimensional MOC Filter described here, there are two types of sustained cells (corresponding to the two directions-of-contrast) and two types of transient cells (on-cells and off-cells). Consequently, the system computes four types of gated responses. Two of these produce cells that are sensitive to local rightward motion: the (BD, +) cells that respond to $x_i^{2BD}x_i^{3+}$, and the (DB, −) cells that respond to $x_i^{2DB}x_i^{3-}$. The other two produce cells that are sensitive to local leftward motion: the (BD, −) cells that respond to $x_i^{2BD}x_i^{3-}$ and the (DB, +) cells that respond to $x_i^{2DB}x_i^{3+}$.

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

*Level 5: local and global motion signals.*To create local motion signals that are insensitive to direction of contrast, define a local right motion response by

$$r_i = Nx_i^{2BD}x_i^{3+} + Px_i^{2DB}x_i^{3-} \tag{A10}$$

and a local left motion response by

$$l_i = Px_i^{2BD}x_i^{3-} + Nx_i^{2DB}x_i^{3+}, \tag{A11}$$

where parameter N scales the contribution of local motion signals created by transient on-cells and parameter P scales the contribution of local motion signals created by transient off-cells. These responses are sensitive to direction-of-motion, but are insensitive to the direction-of-contrast of a moving luminance edge. Boundary signal inputs are already insensitive to direction-of-contrast (Grossberg & Mingolla, 1985a, b), so their influence does not change the insensitivity of local motion cells to contrast polarity. The value r_i is plotted in Figs 9(b), 11(d), 13(d), and 14(d).

These local motion responses are pooled by a long-range spatial filter that has a Gaussian profile across space. 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) transforms even a very broad input pattern into a focal activation at the position that receives the maximal input.

The outputs of Level 4 are assumed to be filtered by a long-range operator with a Gaussian kernel. The Gaussian weight of a pathway from a Level 4 cell at position j to a Level 5 cell at position i is

$$G_{ji} = \exp[-(j-i)^2/2Q^2]. \tag{A12}$$

Parameter Q establishes the spread of the Gaussian kernel with larger values of Q creating broader kernels. Thus, a rightward motion sensitive cell at Level 5 receives input of the form

$$\bar{R}_i = \sum_j r_j G_{ji}. \tag{A13}$$

Similarly, a Level 5 cell sensitive to leftward motion receives input of the form

$$\bar{L}_i = \sum_j l_j G_{ji}. \tag{A14}$$

The Gaussian kernel generates a spatially distributed input to Level 5 in response to even a focal input to Level 1. The contrast-enhancing competitive interactions within Level 5 generate the activities that encode a local measure of motion information. In the simulations reported here, the competition selects that population whose input is maximal. Thus, the activity of a rightward Level 5 cell is

$$R_i = \begin{cases} \bar{R}_i & \text{if } \bar{R}_i = \max_j \bar{R}_j \\ 0 & \text{otherwise} \end{cases}. \tag{A15}$$

The values of R_i are computed each time step and are plotted in Figs 11(e), 13(e), and 14(e).

G-wave motion. There are several methods of measuring strength of the G-wave. One method is to measure the average value of Level 5 global motion cells over the course of the motion. Such an approach proves impractical in our simulations because the activities of Level 5 cells are sampled once every (simulated) msec. Simulations with different combinations of ISI, stimulus duration, and spatial separation sample Level 5 cell activities at different spatial and temporal locations during the movement. If most of the samples are near a weak stimulus, the average G-wave strength is less than if the samples are near a strong stimulus. The discrete sampling of the simulation can warp the calculated strength when the motion is faster than the sampling rate.

As an alternative, we measure the input to the global motion cell centered between the two stimuli at a time that is sampled in every simulation. Thus, whenever the stimulus display produces apparent motion, we calculate motion strength as $\bar{R}_v(\tau)$, where v is the pixel position centered between the two stimulus and τ is the time just before the global motion signal moves away from the first stimulus. Although $\bar{R}_v(\tau)$ does not survive the spatial competition among Level 5 global motion cells, it nonetheless acts as a sensitive measure of G-wave strength, as the simulations shown here attest.

APPENDIX B

Parameters

All simulations use one set of parameters. These include: $A = 0.5$, $B = 10.0$, $C = 0.1$, $D = 10.0$, $E = 10.0$, $F = 0.4$, $G = 2.0$, $H = 0.5$, $J = -0.001$, $K = 0.06$, $L = 3.0$, $M = 5.0$, $N = 5.0$, $P = 1.0$, $Q = 60.0$. With these parameters, 0.01 time units correspond to 1 msec. All differential equations defining the MOC Filter were integrated using Euler's method with a step size of 0.001 time units.

These parameters were not "tweaked" to provide the best quantitative fit to the data, but were chosen somewhat loosely to generate the qualitative properties. The key properties of the model that needed to be generated by the parameters were a lag in response time to the offset of the luminance input (parameters A, B, F, and G) and a transient response from Level 3 cells (parameters F, G, H, J, K, L and M). It is probably possible to generate the particular characteristics needed in this paper with fewer parameters, but we have chosen to remain consistent with other versions of the MOC Filter (Grossberg & Rudd, 1989, 1992; Grossberg & Mingolla, 1993) where the inclusion of those parameters is more important. The remaining parameters were chosen to put the quantitative values of the simulation results in the general range of the psychophysical data.

Each illusory contour inducer consists of two luminous increments. The inducers within each stimulus are separated by 26 pixels. In Fig. 11 the edge-to-edge distance between stimulus inducers is 15 pixels. Each pair of inducers is presented for 150 simulated msec with an ISI of 10 simulated msec. The calculation of illusory contour ISI thresholds in Figs 1(b) and 12 use an edge-to-edge distance of 12 pixels between the inducer pairs.

For the interattribute apparent motion simulation in Fig. 13, the illusory inducers remain unchanged. The luminous contour is a row of luminance increments 30 pixels long ranging from pixel 48 to pixel 77.

The illusory inducers are presented for 100 simulated msec and the luminous contour presented for 90 simulated msec. The stimuli are separated by 12 pixels and an ISI of 200 simulated msec.

The luminous contours for the studies of Korte's laws are the same as above (with the first stimulus now a luminous contour). For both the illusory and luminous stimuli, the inputs to the BCS simulations described in Francis *et al.* (1994) keep the same spatial and emporal properties, but differ in magnitude from the retinal inputs to the MOC Filter. This parameter change achieves consistency with the additional preprocessing of luminous inputs in the BCS simulations (Level 1 in Francis *et al.*, 1994). Including such preprocessing in the MOC Filter would not change any of the results described herein, but would force the use of substantially more computer simulation time.

The G-wave strength threshold plotted in Figs 12 and 15 is 26 units. G-wave strength was measured at unit intervals (not every 20 units, as plotted in Figs 12 and 15). The translation of pixel units to visual degrees [Fig. 2(b)] is the same used by Francis *et al.* (1994).

The simulations calculating ISI thresholds were performed on a multi-user Iris 8/280 Silicon Graphics workstation. The data for Fig. 15 take approximately 2 weeks to calculate. Data showing the time–course of cell activations for one apparent motion display were computed on a Gateway 486 4DX2-66V personal computer and take approximately 1 min.