A NEURAL MODEL OF HIGH-LEVEL MOTION PROCESSING: LINE MOTION AND FORMOTION DYNAMICS

Aijaz A. Baloch[1]   Stephen Grossberg[2]
Department of Cognitive and Neural Systems
and
Center for Adaptive Systems
677 Beacon Street
Boston, MA 02215

July 1996
Revised: January 1997
Revised: March 1997

Technical Report CAS/CNS-TR-96-020
Boston, MA: Boston University

Key Words: visual cortex, line motion, motion induction, attention,
transformational apparent motion, formotion, neural network

Running Head: Neural Dynamics of Formotion Perception

Requests for reprints should be sent to:
Stephen Grossberg
Department of Cognitive and Neural Systems
Boston University
677 Beacon Street
Boston, MA 02215

ABSTRACT

The percepts known variously as the line motion illusion, motion induction, and transformational apparent motion have attracted a great deal of experimental interest, since they sensitively probe interactions between preattentive and attentive vision processes. The present article develops a neural model that qualitatively explains essentially all the data reported thus far, and quantitatively simulates key illustrative percepts. The model suggests how these data arise from neural mechanisms of preattentive boundary and surface formation, long-range apparent motion, form-motion interactions, and spatial attention. The boundary and surface formation processes model aspects of the interblob V1 $\rightarrow$ interstripe V2 $\rightarrow$ V4 and blob V1 $\rightarrow$ thin stripe V2 $\rightarrow$ V4 cortical processing streams, respectively. The long-range apparent motion process models aspects of the V1 $\rightarrow$ MT $\rightarrow$ MST processing stream. An interstream V2 $\rightarrow$ MT form-motion interaction is proposed to allow the motion processing stream to track transient properties of emergent boundaries and filled-in surface colors from the form processing stream. It does so by generating motion waves using the long-range apparent motion process. This interstream interaction controls the formation of form-motion percepts, which are herein called formotion percepts. Other transients directly cause motion waves within the motion processing stream. All the data are attributed to properties of such motion waves. It is also suggested how bottom-up motion mechanisms can engage top-down attention as part of the motion capture process that solves the aperture problem. This interaction is proposed to occur between areas MT and MST. The model hereby explains how attention can be engaged even in percepts whose explanation can be derived from preattentive mechanisms.

1 Introduction

A number of authors have reported motion percepts that occur when two spatially overlapping shapes that are presented discretely in time appear to transform smoothly from one shape to another. Hikosaka, Miyauchi, and Shimojo (1993a, 1993b) reported experiments in which a line or bar that is presented next to a previously presented spot appear to smoothly grow out of the spot. This line motion illusion was attributed to an attentional gradient such that regions of the line closest to the attended spot are processed faster and thereby activate higher-level motion detectors earlier. Speed-up of information processing by attention has also been reported by Stelmach *et al* (1991, 1994), who showed that attending to one of two stimuli in a long-range apparent motion display altered the perceived direction, pattern, and quality of motion. Steinman, Steinman, and Lehmkuhle (1996) showed that cues that preferentially excite the magnocellular cortical pathway predominantly capture this type of visual attention. Shimojo, Miyauchi, and Hikosaka (1992) showed that motion could also be attentionally primed in response to an auditory or somatosensory stimulus.

Faubert and von Grünau (1992, 1995) and von Grünau and Faubert (1994) extended research on line motion illusion to include a larger class of phenomena that they called motion induction. For example, when the line is shut off, motion appears to reverse, and the line seems to be sucked back into the spot. In split priming experiments, when a line is presented between two spatially separated priming spots, motion emerges from both spots and collides in the middle of the line. When the spots are not turned on simultaneously, the collision point occurs closer to the first spot. In attribute priming experiments, the contribution of low-level features, such as color or luminance, to the direction of perceived motion was assessed. In some experiments, the color (or luminance) of two spots was different and the color (or luminance) of the line matched only one of the spots. Motion was always perceived as emerging from the spot that matched the color (or luminance) of the line. Varying the delay between the spots did not have a major effect on the direction of perceived motion. Apparent motion studies of Kenkel (1913) on gamma motion and of Kanizsa (1951, 1979) on polarized gamma motion had earlier noted some of these effects.

These studies demonstrate that top-down attention cannot be the only mechanism mediating motion induction percepts. Bottom-up processes must also be at work in attribute priming experiments because the line always appears to grow out of the spot that matches its color or luminance, even if both spots are equally salient as attentional primes. Likewise, it is not clear why attention should cause split priming effects or reverse motion at line offset. Faubert and von Grünau (1995) argued that, in experiments with single spot primes, top-down factors are more dominant, but that bottom-up factors are more important in two spot experiments. How these factors might work, from a functional and mechanistic viewpoint, was not disclosed by the various experiments. We provide such a functional and mechanistic account below in which bottom-up factors may also play a key role even in single spot experiments.

Tse and Cavanagh (1995) and Tse, Cavanagh, and Nakayama (1996) have also contested the claim that line motion and motion induction percepts are due to a gradient of attention. They call these phenomena transformational apparent motion, and claim that they "obey different properties than those obeyed by standard apparent motion". This type of motion is asserted to involve a parsing and matching stage that occurs before the motion system is activated. "Determining that something moved requires that something be identified at the first instance and then paired off with what is presumed to be the same thing in the next instant. The first component of this process is to identify candidates at both instants and the second is to match them". As in the work of Faubert and von Grünau, they manipulate featural factors such as contour continuity, color, and shape, but also explore effects of figure-ground separation and amodal completion on motion percepts. Unlike

classical apparent motion percepts wherein parsing is resolved due to spatially non-overlapping stimuli, transformational apparent motion percepts parse spatiotemporal data wherein overlapping forms that are ambiguous in one time frame may be disambiguated by configural factors that operate across time.

Taken together, these phenomena invoke processes of low-level and high-level motion, interactions between motion and other vision processes such as "parsing", and interactions between bottom-up stimuli-driven processes and top-down attention. Here is a data base wherein small changes in stimulus properties such as luminance, color, and shape can substantially shift the balance between several contributing processes. As a result, the Tse *et al* notion of "parsing" leaves open large areas of ambiguity. As Tse *et al* (1996) note, "geometry-based parsing principles ... are not sufficient ... geometry supplies no basis for attributing new image data to one cue rather than another". To offset these deficiencies, auxiliary concepts such as "minimal mapping" and "minimal cover property" were invoked by these authors. Tse *et al* (1996) also suggest that parsing occurs "before" motion processing, that there is "a stage of form extraction in the high-level motion processing stream" and that the "high-level signal ... serves as a solution to the aperture problem".

We argue below that these claims mix up processing streams and stages of before and after because the language in which they are framed is not powerful enough to describe the underlying mechanisms. Line motion, motion induction, and transformational apparent motion phenomena are thus ripe for a modeling approach in which the contributing visual processing mechanisms and their interactions can be rigorously defined and simulated, the functional design principles that are realized by these mechanisms can be articulated, and an explanation can be given of how these principles handle more ecologically natural phenomena.

2   A Framework for Explaining Formotion Percepts

In this study, we show that all these motion percepts can be explained by available vision models whose functional principles and neural mechanisms were originally derived to explain other data bases, notably data about boundary segmentation, surface filling-in, apparent motion, form-motion interactions, and spatial attention (e.g., Francis and Grossberg, 1996a, 1996b; Grossberg, 1991, 1994; Grossberg, Mingolla and Ross, 1994; Grossberg and Rudd, 1992; Grossberg and Todorović, 1988). Our analysis develops two central themes. The first theme suggests that many of these motion data can be explained without recourse to spatial attention mechanisms. The second theme suggests how spatial attention mechanisms may get involved.

The first theme explores the hypothesis that form and motion processing take place in parallel streams of visual cortex, but interact across streams in order to compensate for deficiencies of each stream towards generating percepts of moving-form-in-depth (Francis and Grossberg, 1996a; Grossberg, 1991). The form processing uses *orientationally* tuned operations to group edges, textures, and shading into 3-D boundary and surface representations of objects separated from their backgrounds (Figure 1). It has been proposed (Grossberg, 1994; Grossberg and Mingolla, 1985a, 1985b) that boundary representations are generated and separated in the interblob V1 → interstripe V2 → V4 processing stream of visual cortex, and that surface representations are generated in the blob V1 → thin stripe V2 → V4 processing stream. This view differs from that, say, of Livingstone and Hubel (1984) who identify these two streams with form and color processing, respectively. The present view proposes that both streams are devoted to form processing; in particular, to the complementary properties of boundary form and surface form.

The motion processing stream sacrifices orientational precision to generate estimates of motion *direction* and speed (Albright, Desimone, and Gross, 1984; Allman, Miezin, and McGuinness, 1985;
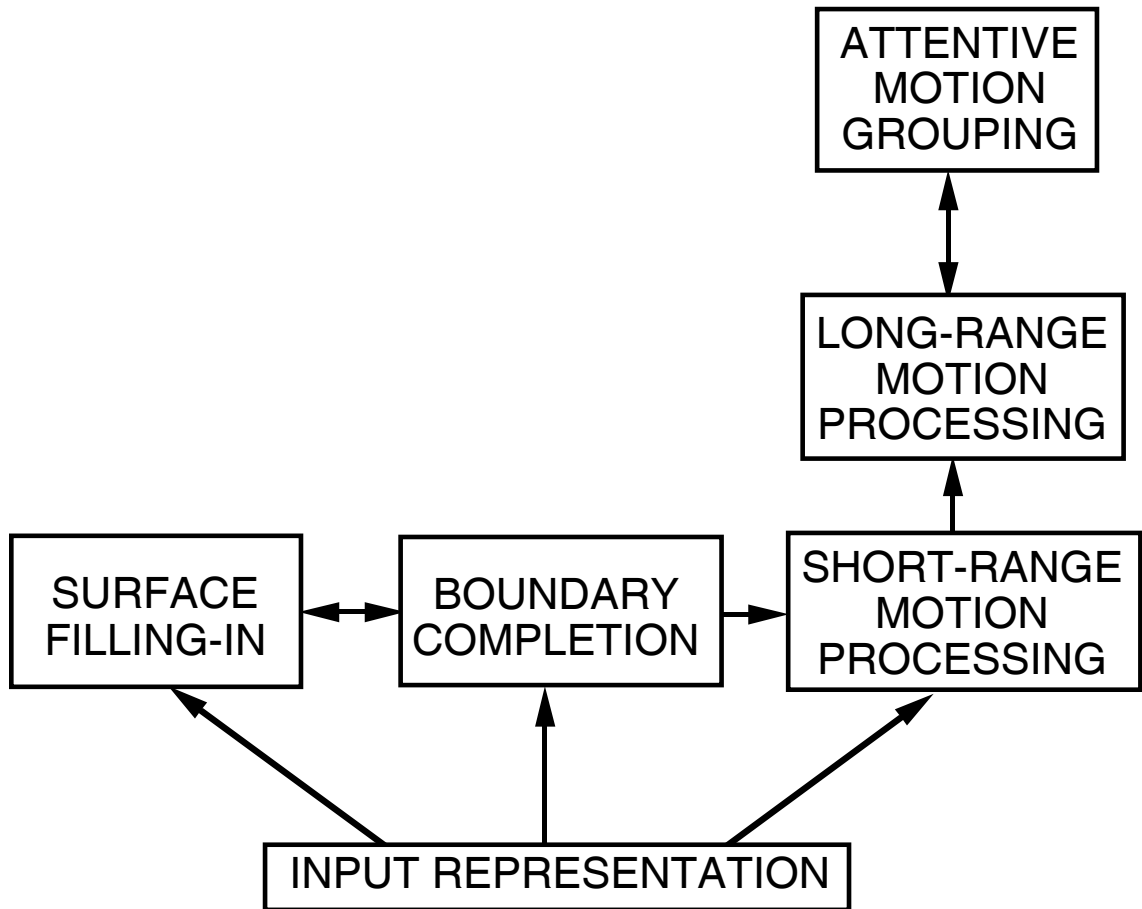
Figure 1: **Schematic of parallel processing streams and interactions that lead to for-motion percepts. Surface filling-in is proposed to occur in the blob V1 → thin stripe V2 → V4 cortical processing stream, boundary completion in the interblob V1 → in-terstripe V2 → V4 stream, and motion processing in the V1 → MT → MST stream.**

Maunsell and van Essen, 1983; Newsome, Gizzi, and Movshon, 1983) that can be used to solve the global aperture problem (Chey, Mingolla, and Grossberg, 1994, 1997; Grossberg and Mingolla, 1993). Pooling many orientations into a single direction of motion causes a loss of stereo acuity within the motion processing stream (Logothetis *et al*, 1990; Schiller, Logothetis, and Charles, 1990). The form-motion interaction across streams enables the form stream to input its emergent form-and-color-and-depth, or FACADE, information into the motion stream. As a result, the motion stream can better detect and track in depth the moving objects whose 3-D boundaries and surfaces pop-out within the form stream.

This form-motion interstream interaction is crucial in our explanations of the line motion illusion and its generalizations. We suggest that percepts which arise from it be called *formotion* percepts since they involve the active *formation* of *form-motion* percepts.

At what cortical processing stages does the form-motion interstream interaction occur? Gross-berg (1991) suggested that it outputs from the V1 → V2 form processing stream *after* the stage at which 3-D boundaries are formed, and inputs to the V1 → MT motion processing stream *before* the stage of long-range motion filtering. In particular, the input to the motion stream needs to occur before the stage where information from opposite contrast polarities and multiple orientations are pooled into motion directions. The interstream interaction was thus predicted to occur via a V2

$\rightarrow$ MT connection. This interaction is shown schematically as the interaction between boundary processing and short-range motion processing in Figure 1. In addition, it is assumed that the motion stream responds to transient changes in the form stream. Otherwise, a stationary form could generate persistent signals to the motion system. Transient cell responses are also used to directly activate the motion system, as described in more detail below.

Francis and Grossberg (1996a) have computationally modeled this form-motion interaction and used it to simulate data that link the persistence of boundary segmentations in the form stream to the quality of apparent motion in the motion stream, including Korte's Laws (Korte, 1915). Herein we extend this approach to show how many formotion percepts emerge from preattentive waves of boundary growth and color filling-in within the form stream, waves of long-range apparent motion within the motion stream, and interactions between these changing form and motion signals via the form-motion linkage. The results were first reported in Baloch and Grossberg (1996).

Our main results may be reduced to an analysis of the conditions under which a motion wave occurs; namely, a wave of neural activity across the model processing level that computes long-range apparent motion. As shown below, such a motion wave can be generated *directly* within the V1 $\rightarrow$ MT processing stream by using the long-range motion processing mechanisms of that stream (see Figure 1), or *indirectly* via the form-motion interaction in response to transients of boundary growth or decay, and of color filling-in, within the form processing stream. Our analysis discusses how each of these processes respond to formotion inputs.

In order to distinguish a motion wave that may be due to an indirect form-motion interaction from a motion wave that is directly generated by the long-range apparent motion process, we call the latter a G-wave, for reasons that are made clear below. We also explain how various combinations of activity onsets and offsets, or relative onset rates or offset rates, can lead to such a G-wave (Figure 2). The proposal that onset and offset combinations can lead to a G-wave was first used to simulate data about long-range apparent motion (Grossberg and Rudd, 1989, 1992). G-waves occur in the motion stream at the long-range motion filter, whose functional role is to combine motion estimates from multiple orientations, contrast polarities, and both eyes into a pooled estimate of motion direction. Here we show that when evolving boundary and surface signals input to the motion stream, in addition to the offsets and onsets that are directly converted into G-waves by early motion mechanisms, then almost all formation data known to us can be explained. We hereby explain formotion data as manifestations of the mechanisms whereby 3-D forms are generated and tracked as they move in a prescribed direction.

The second theme concerns how visual attention may be attracted towards combination of object onsets, offsets, or motion during bottom-up visual information processing. As schematized in Figure 1, the model proposes that the long-range apparent motion mechanism that generates G-waves can also engage spatial attention. G-waves help spatial attention to track even intermittently viewed targets by smoothly interpolating their intermittently viewed positions (Grossberg, 1991, 1997b). G-waves can smoothly interpolate intermittent target views even if they represent targets moving with variable speed (Grossberg and Rudd, 1989, 1992). G-wave properties are consistent with data showing that spatial attention can travel across variable distances in equal time (Kwak, Dagenbach, and Egeth, 1991; Remington and Pierce, 1984), is controlled by the magnocellular processing stream (Steinman *at al*, 1996), and has a center-surround organization (Steinman *et al*, 1995).

The model also proposes how, once attention is engaged, top-down attentional priming can influence the direction of perceived motion (Groner, Hofer, and Groner, 1986; Sekuler and Ball, 1977; Stelmach, Herdman, and McNeil, 1994). The model suggests that this top-down process helps to solve the aperture problem by capturing ambiguous motion signals and defining an attended
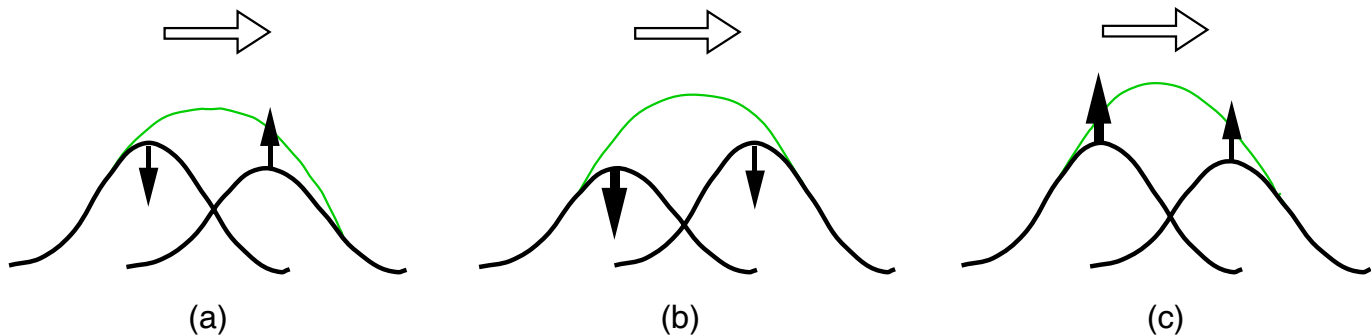
Figure 2: Some conditions leading to a continuous motion wave in response to two or more spatially and temporally disjoint inputs. Suppose that these inputs activate long-range Gaussian kernels whose total activity summates before the peak activity is selected by a center-surround network. Under appropriate spatiotemporal conditions, the peak activity moves continuously from one input position to the other if (a) the first Gaussian decays as the second Gaussian grows, (b) the first Gaussian decays faster than the second one, or (c) the first Gaussian grows faster than the second one.

object's global direction and speed of motion (Chey *et al.*, 1997). In this conception, output cells from the long-range motion filter go through another directionally-selective filter whose cells compete to choose a winning direction. The winning cells send top-down signals back to the long-range filter cells; see Figure 1. These top-down signals select long-range filter cells that encode the same direction and inhibit cells that do not. Attention is hereby focussed on the cells which conform to the winning direction. Thus motion capture, which seems to be an automatic and preattentive process, is proposed to be carried out by the same circuit that permits top-down attention to selectively focus on a desired direction.

Various data support this conception. Cavanagh (1992) has described an attention-based motion process, in addition to a "low-level" or automatic motion process, and has shown that it provides accurate velocity judgments. By suggesting how this attentive process and motion capture are linked, the model explains how the attentive process yields accurate velocity judgments. Within the context of formotion experiments, the model clarifies how displays that activate the motion system can also focus spatial attention. Von Grünau, Dubé, and Kwas (1996) and von Grünau, Racette, and Kwas (1996) have carried out a number of experiments with which to disentangle and measure the preattentive motion and attentional priming effects. They showed that attentional priming develops slower, consistent with the proposal that it is activated through a feedback process.

The directional choice in the feedback circuit is proposed to occur in the ventral part of cortical area MST, which has large directionally tuned receptive fields that are specialized for detecting moving objects (Tanaka *et al.*, 1993). In this interpretation, MST_ν can attentionally modulate MT cells, which are proposed to include the long-range filter cells. Consistent with this proposal, Treue and Maunsell (1996) have shown that attention can modulate motion processing in cortical areas MT and MST in behaving macaque monkeys. O'Craven *et al.* (1996) have shown using fMRI that attention can modulate the MT/MST complex in humans. This interpretation leads to the prediction that MST_ν cells make a directional choice that is used to overcome aperture ambiguities in MT cell responses. Top-down signals from MST_ν cells are proposed to select MT cells that encode an object's direction of motion and to suppress those that do not (Chey *et al.*, 1997).

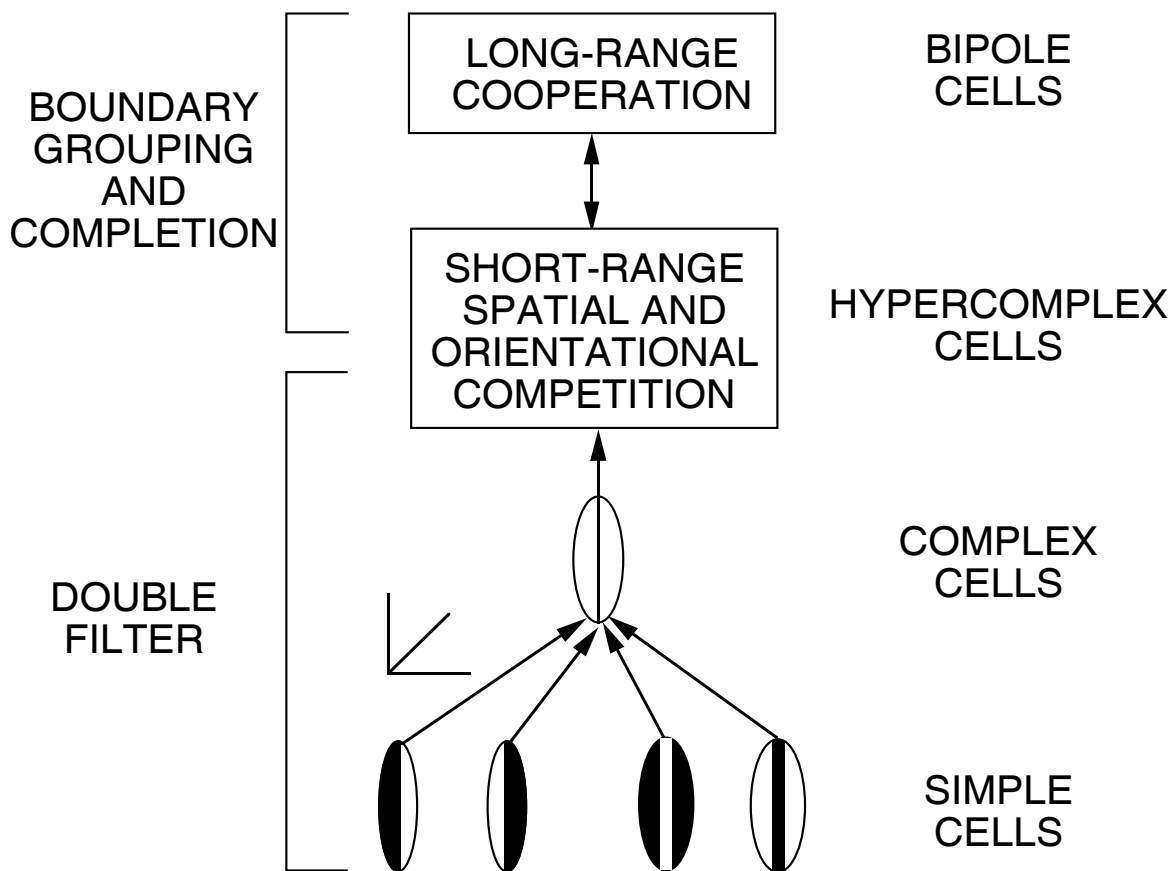Our exposition begins with a brief summary of models of boundary completion and color filling-

Figure 3: Schematic of a monocular single-scale version of the BCS model.

in to highlight properties that are important for explaining formotion percepts. Then relevant modeling properties of short-range and long-range motion processing are summarized. A theorem is stated that characterizes when a G-wave can occur in both long-range apparent motion and formotion percepts. Then typical formotion data are analyzed and explained using the model, and some key examples are simulated. Equations, parameters, and proofs of theorems are provided in the Appendices.

3   Waves of Boundary Completion

The Boundary Contour System (BCS) model was introduced to explain how the brain generates 3-D boundary segmentations in response to edges, textures, shading, and stereo information (e.g., Grossberg, 1994; Grossberg and Mingolla, 1985a, 1985b, 1987). A schematic of a single-scale monocular version of the BCS model is given in Figure 3.

Figure 3 shows that the model consists of two parts: a double filter and a boundary grouping and completion network. In the double filter, complex cells pool together half-wave rectified outputs from simple cells that are sensitive to opposite contrast polarities. Complex cells hereby compute an oriented full-wave rectification of a scene. The second filter processes complex cell outputs via short-range center-surround interactions across space and orientation. These interactions generate the receptive fields of hypercomplex (or endstopped complex) cells. Variants of this filter have become standard in models of texture segregation (e.g., Chubb and Sperling, 1989; Grossberg and Mingolla, 1985b; Malik and Perona, 1990; Sutter, Beck, and Graham, 1989).

The present analysis focuses upon how boundaries are grouped and completed, notably upon transients that can create waves of boundary formation from priming to test stimuli. In the bound-
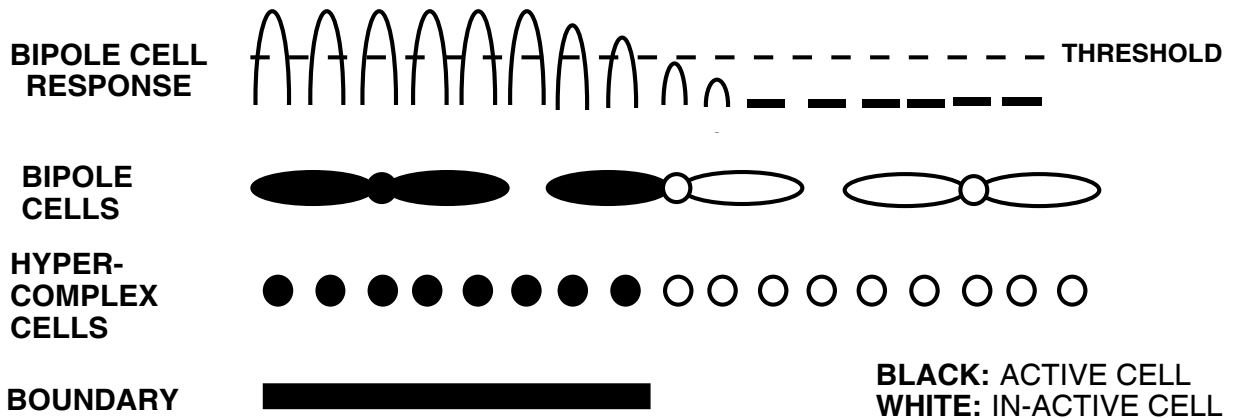
Figure 4: Long-range cooperation during boundary completion: The bipole cells on the left are active suprathreshold because they receive input from one or more horizontal lobes and the cell body. The bipole cells in the middle receive input from the left horizontal lobe only and are thus active subthreshold. The bipole cells on the right are not active. When the edge is suddenly extended to the right, the bipole cells near the middle reach their thresholds earlier than the bipole cells on the right.

ary completion network, bipole cells cooperatively group together inputs from hypercomplex cells whose positions and orientations are similar to those of the bipole cell receptive field. The bipole cell receptive field has two oriented horizontal lobes in addition to the cell body. A bipole cell can fire if it receives enough oriented input to both lobes, or to at least one lobe and the cell body. (Variants in which input to the cell body alone can fire the cell are also possible.)

Activated bipole cells compete across position and orientation before generating positive feedback signals to like-oriented hypercomplex cells at the same position. These feedback signals help to create and enhance spatially and orientationally consistent boundary groupings, while inhibiting inconsistent ones. Hypercomplex boundary signals with the most cooperative support from bipole grouping thereupon further excite the corresponding bipole cells. These bottom-up and top-down cooperative-competitive interactions rapidly converge to a final boundary segmentation. These concepts have been used to explain and predict behavioral and neural data about boundary segmentation in a number of reports (e.g., Francis and Grossberg, 1996a, 1996b; Francis, Grossberg, and Mingolla, 1995; Gove, Grossberg, and Mingolla, 1995; Grossberg, 1994; Grossberg and Mingolla, 1985a, 1985b, 1987). Grossberg *et al* (1997) have analyzed how such mechanisms may be embedded in cortical layers, columns, and maps. Grossberg *et al* (1995) have shown that such circuits are competent to process complex imagery.

To understand how boundary waves are formed, two key properties of the BCS are needed: similar orientations facilitate each other via *long-range cooperation* and dissimilar orientations inhibit each other via *short-range competition*. These two factors together accelerate the formation of smooth contours and slow down the formation of abrupt changes in boundary orientation.

Long-range cooperation
Long-range cooperation is mediated by bipole cells. Because a bipole cell cannot fire if only one horizontal lobe of its receptive field is activated, bipole cells in the vicinity of a priming edge may not fire suprathreshold even if they are more active than cells that are farther away. If the priming edge is extended by a test stimulus, the bipole cells closer to it reach their threshold earlier than those that are farther away. In Figure 4, the bipole cells on the left are active suprathreshold because

7

they receive input from the priming edge within one or more horizontal lobes and the cell body. The bipole cells in the middle are only active subthreshold because they receive priming input only from the left horizontal lobe. Bipole cell activities become gradually smaller as a function of their distance from the priming edge. The bipole cells on the right are not active at all because they receive no input. When the edge is suddenly extended to the right, the bipole cells near the middle reach their thresholds earlier than the bipole cells on the right. A wave of boundary completion ensues from the priming stimulus towards the test edge. This example illustrates how a boundary wave can contribute to the line motion illusion even if there is no top-down attentional priming.

Short-range competition
Dissimilar orientations inhibit each other in a spatial neighborhood via short-range competition between dissimilarly-oriented bipole cells and hypercomplex cells. Figure 5 illustrates the effect of such short-range competition on the temporal growth of an edge close to a dissimilarly oriented edge. The vertical bipole cell in Figure 5b is active in response to the vertical edge and inhibits the horizontal bipole cell in its neighborhood. The horizontal bipole cell in Figure 5a does not receive an inhibitory input as no vertical edge is present in its neighborhood. When a horizontal edge is now presented to both horizontal bipole cells in Figure 5 simultaneously, the bipole cell in Figure 5a is activated faster than the bipole cell in Figure 5b. Thus if, as in Figure 5c, a test bar turns on between a bar with which it is parallel and a bar with which it is perpendicular, the boundary will grow more quickly from left to right, without the intervention of top-down attention or parsing and matching rules.

4   Surface Formation and Color Filling-In

Boundary and surface representations are both needed to generate a 3-D representation of a scene. The Feature Contour System (FCS) model was introduced in order to explain how the brain discounts the illuminant and uses the discounted signals to fill-in surface representations of brightness, color, depth, and form. Behavioral and neural data that have been explained by the FCS are contained in numbers of reports (e.g., Arrington, 1994; Cohen and Grossberg, 1984; Francis and Grossberg, 1996b; Grossberg and Todorović, 1988; Paradiso and Nakayama 1991; Takeichi, Watanabe, and Shimojo, 1992; Watanabe and Cavanagh, 1992; Watanabe and Sato, 1989; Watanabe and Takeichi, 1990). In all these examples, interactions between BCS and FCS mechanisms determine the final percept. Here we show how addition of a test stimulus that touches one or more priming stimuli can reorganize both BCS boundaries and FCS filtering and filling-in signals in such a way as to generate a wave of filling-in that correlates well with various formotion percepts. This filling-in wave can occur in parallel with a boundary wave in the same direction. To see how this works, a brief review of FCS concepts is needed.

The first stage of the FCS model is a monocular preprocessing stage that consists of on-center off-surround (ON cell) and off-center on-surround (OFF cell) receptive fields. These cells discount the illuminant and compute Weber-law modulated contrast ratios of the image. In the simplest monocular version of the FCS, these activities are half-wave rectified and topographically input to a Filling-in Domain, or FIDO, which is a regular array of intimately connected cells such that contiguous cells can easily pass activity to each other.

Each FIDO also receives boundary signals from the BCS. These boundary signals act as inhibitory gating signals that stop the spread of activation across boundaries. The net effect of these interactions is that the activation which is initiated by FCS signals is diffused and averaged within the boundaries generated by BCS signals. Figure 6 shows a FIDO along with its FCS activating and BCS gating signals.
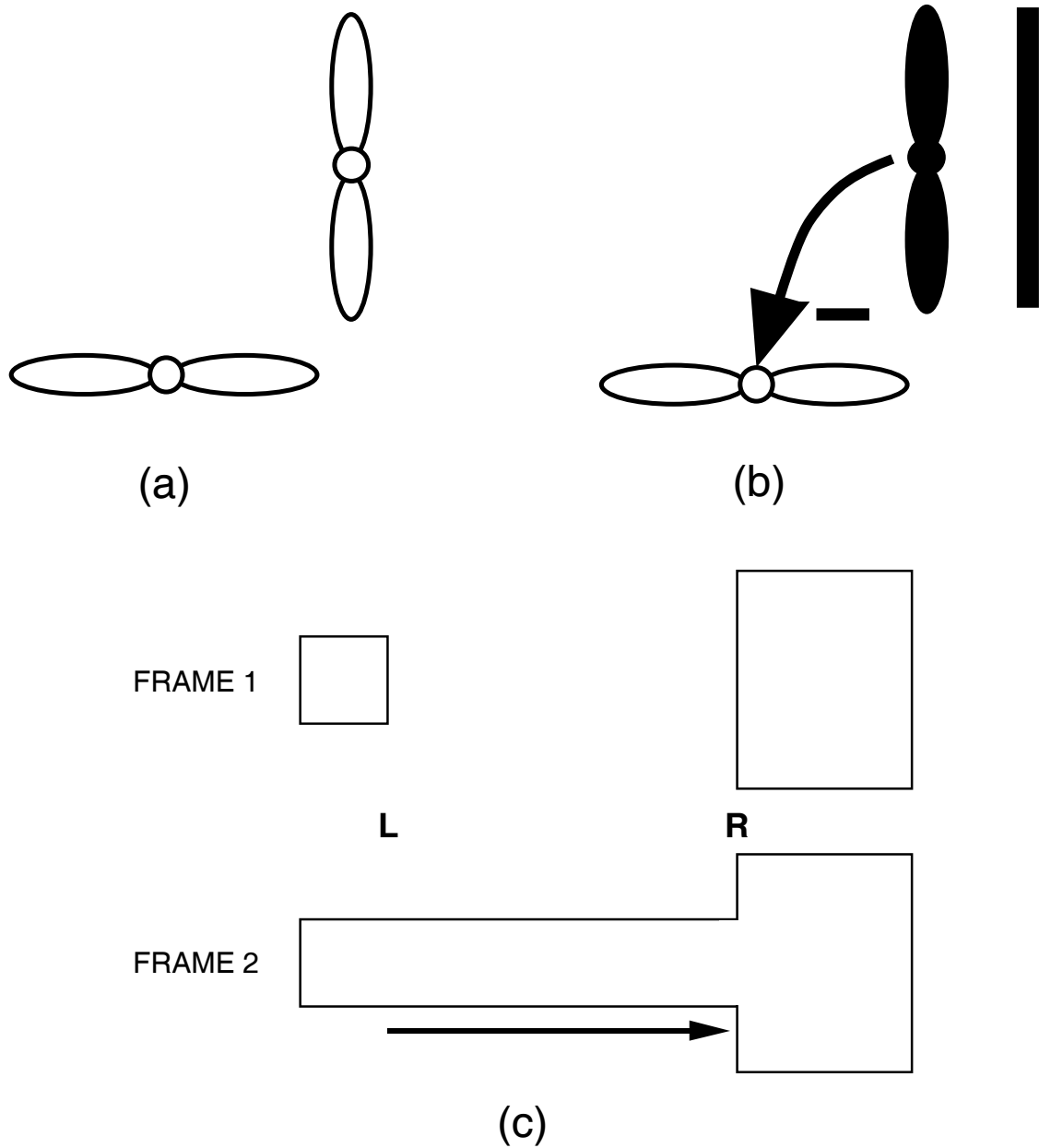
Figure 5: Short-range competition during boundary completion: The active vertical bipole cell in (b) inhibits the horizontal cell in its neighborhood while the horizontal cell in (a) does not receive any inhibitory input as no vertical edge is present. (c) When a horizontal edge is now presented to both horizontal bipole cells at L and R simultaneously, the bipole cell at L is activated faster than the bipole cell at R.

BCS INPUTS

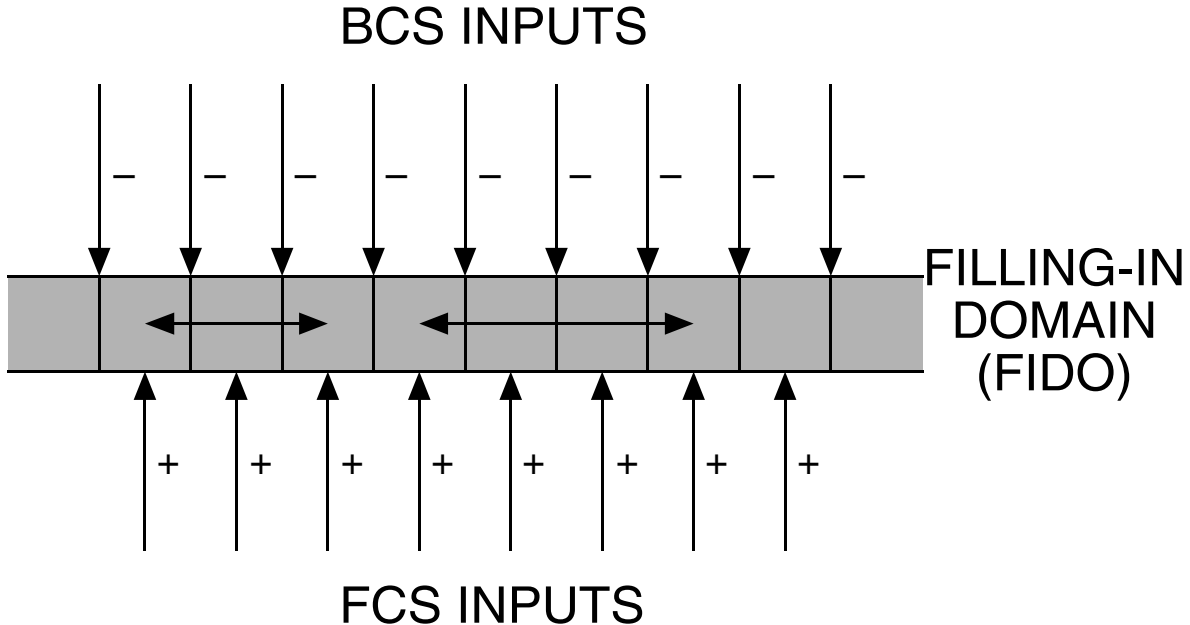FILLING-IN DOMAIN (FIDO)

FCS INPUTS

Figure 6: Filling-In DOmain (FIDO): FCS inputs initiate filling-in of the area between the active boundaries via a diffusion process. The FIDO diffusion is excited by FCS inputs and gated by inhibitory BCS boundary signals.

These BCS-FCS interactions help to explain properties of color-specific priming experiments (Faubert and von Grünau, 1992, 1995; Tse *et al*, 1995, 1996). When a color boundary is removed by presenting the same color next to it, the color can rapidly fill-in from the existing color to the remaining boundaries. Figure 7a illustrates this property. The area to the left is filled with red color due to prior presentation of a priming stimulus. Color-inducing signals exist all along the boundary of the prime, as indicated by the black circles. When the area between **L** and **R** receives a red test bar, the boundary at **L** between the prime and the test bar collapses quickly while the new boundary at **R** and the new color-inducing signals along the test bar boundaries (indicated by white circles) are growing. Since the color-inducing signals at **L** have already filled-in the prime when the inhibitory boundary-gating signal is removed at **L**, color can start to immediately diffuse from the left while the new test bar region gradually responds to its new color inputs. This is perceived as a wave of color filling-in from **L** to **R**.

Why does color seem to flow into the test bar from a prime with the same color, rather than one of an opponent color? One factor is the ON cell opponent color receptive field, which slows down the growth of an opponent color in the inhibitory surround of an existing color. Figure 7b shows a red priming bar to the left and a green priming bar to the right. The figure also shows two on-green off-red opponent receptive fields close to the red and green regions. Black circles indicate primed color-inducing cells and white circles indicate test bar color-inducing cells. When the region between **L** and **R** receives a green test bar, the on-green off-red receptive field at left (**L**) receives inhibitory input from the red-surround while the on-green receptive field at right (**R**) does not have inhibitory red in its surround. Thus the ON cell receptive field at right becomes active faster than the receptive field at left. Since cells at this preprocessing stage input to both the FCS and the BCS, the boundary at **R** collapses more quickly and the red color-inducing signals build up more quickly, while the red/green boundary at **L** and the corresponding green color-inducing signals at **L** change more slowly. As a result, green color fills-in from **R** to **L**.

The model suggests that waves of boundary and surface formation can influence perception in

(a)

(b)

LEGEND

RED
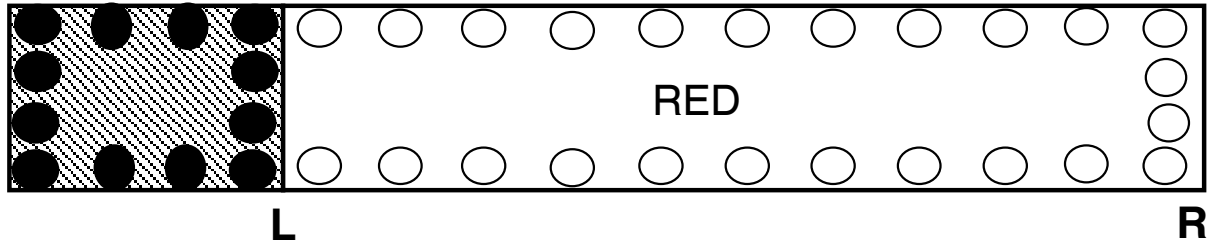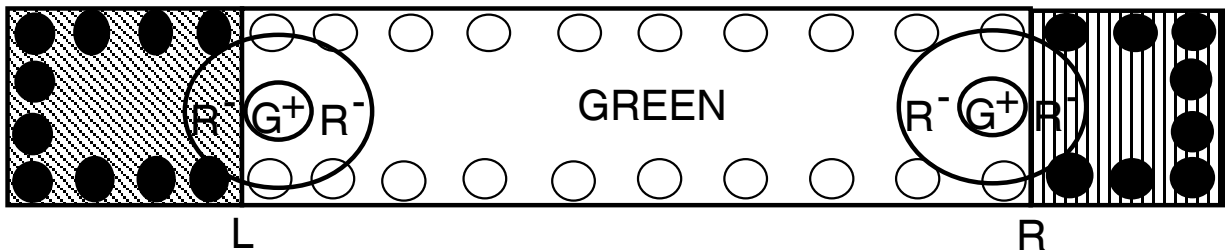
GREEN

● COLOR INDUCING SIGNALS (FRAMES 1 & 2)

○ COLOR INDUCING SIGNALS (FRAME 2)

Figure 7: Opponent inputs to color filling-in: (a) When the area between L and R is filled with red color, the boundary at L collapses and the color fills-in from L to R. (b) When an opponent color (green) is presented next to an existing color (red), the off-surround slows the growth of opponent color. Therefore green color-inducing signals buildup slowly at L while the boundary at R collapses quickly. The green color flows from R to L.

at least two ways. They can generate a percept of moving form by propagating upward through the interblob and blob visual cortical streams into areas V4 and IT of the What cortical processing stream. Or they can generate a percept of motion, or more properly of formotion, via an interstream interaction from (say) area V2 to MT, and then upwards towards parietal cortex in the Where and How processing stream (Goodale and Milner, 1992; Mishkin, Ungerleider, and Macko, 1983; Ungerleider and Mishkin, 1982). The form-motion model developed by Francis and Grossberg (1996a) and Grossberg (1991) proposes that the motion mechanisms which are activated by this interstream interaction react to transients of the BCS boundary formation process. The 3-D vision model of Grossberg (1994, 1997a) proposes that FCS surface filling-in processes can modulate BCS boundary processes and, thus, motion mechanisms via a surface-boundary-motion interstream interaction. We simulate this transient interstream interaction below.

Before doing so, some additional points need to be made. Transients in boundary formation at, say, the *discrete* locations **L** and **R** in Figure 5c or Figure 7 can cause a *continuous* G-wave to occur within the motion processing stream. This sort of G-wave is distinct from the boundary and filling-in waves that are formed within the BCS and FCS. It is due to mechanisms of long-range apparent motion processing which are now reviewed.

5   Long-Range Apparent Motion and Formotion Waves

Grossberg and Rudd (1989, 1992) introduced a neural model to explain data about short-range and long-range apparent motion, among other motion phenomena. In the model, mechanisms that are sensitive to short-range motion input to long-range motion processing mechanisms, as in the right column of Figure 1. The long-range mechanisms interact, in turn, with a long-range attentive grouping process. Together these processes have been used to simulate parametric data about how the brain overcomes aperture ambiguities to generate a coherent representation of a moving object's direction and speed (Chey *et al*, 1994, 1997).

The basic idea of how discrete events in time generate a continuous long-range motion wave is very simple. Suppose that the neural activity due to one event decays while the activity due to a later, spatially displaced, event grows (Figure 2a). Let these activities be processed by a spatially long-range Gaussian filter before they are added up. Then the peak activity of the Gaussian sum moves continuously from the position of the first event to that of the second event if their spatiotemporal overlap falls within certain bounds; for example, if the two events are separated by a distance less than half the size of the Gaussian. Such a motion wave is called a G-wave because it is a general property of Gaussianly filtered signals that gradually grow and decay through time.

In many experiments on long-range apparent motion, the *offset* of a first flash is followed by the *onset* of a second flash to generate a G-wave. In some formotion experiments, the same is true. For example, when a red test bar turns on next to a red priming bar, the boundary where they touch shuts off as the opposite boundary turns on (Figure 7a). In contrast, when a test bar turns on between two priming bars in a split priming experiment, both of the boundaries at the test-prime interfaces shut off. In Figure 5c, the vertical boundary at **L** where the test and prime are collinear shuts off faster than the vertical boundary at **R** where they are not collinear. The later boundary persists longer due to cooperative support by vertical bipole cells that receive inputs from the vertical edges of the prime (Francis *et al*, 1994). Appendix D contains the proof that a G-wave can also occur from a fast decaying signal to slow decaying signal. Thus the test bar appears to grow out of the collinear priming bar towards the noncollinear priming bar. This G-wave is created within the motion stream by transient boundary signals from the form stream that are delivered via the form-motion interaction (Figure 1). Such a G-wave can add its motion to the motions derived

from waves of boundary completion and surface filling-in. None of these effects involve higher-order "parsing and matching" rules, as these are commonly understood.

When the bar in the split priming experiment of Figure 5c is removed, a motion in the opposite direction is typically perceived. A G-wave may also be created under these circumstances, since the vertical boundary at the right in Figure 5c grows more quickly than the boundary at the left. This happens because the boundary at the right receives additional collinear activation of its bipole cells from the vertical edges of the priming bar that are not eliminated by the test bar. Appendix D proves that a G-wave can occur from a fast growing signal to a slow growing signal. In all, we can now state in intuitive terms the

Formotion Wave Theorem

A G-wave may be generated from (a) a decaying signal to a growing signal, (b) a fast decaying signal to a slow decaying signal, and (c) a fast growing signal to a slow growing signal under appropriate spatiotemporal conditions.

Figure 2 summarizes these three cases. Grossberg (1991, 1997b) has suggested that such G-waves may help the brain to continuously track moving targets even if they are only intermittently seen and move with variable speeds. The G-wave accomplishes this by generating a continuously moving focus of spatial attention that may be used to command orienting movements towards the target.

6 Analysis of Formotion Experiments

With this background, the basic formotion experiments can now be more systematically analyzed. Figure 8a summarizes the line motion experiment of Hikosaka, Miyauchi, and Shimojo (1993a). In Frame 1, a box or spot is presented followed by a bar or line in Frame 2 that is contiguous to the box. Although the bar is presented all at once, it appears to grow out of the box as indicated by the arrow. If the bar is now removed, a motion in the opposite direction towards the box is perceived. If this experiment is repeated, the bar again appears to grow out of box and then shrink back to it. Hikosaka *et al* (1993a) argued that the spot in Frame 1 attracts spatial attention towards it. Their explanation does not, however, account for motion in the reverse direction when the bar is removed, since one would expect the gradient of spatial attention around the box to be obliterated by the test bar and by the shift of attention that it causes away from the box. In addition, the reverse motion occurs even if the bar is left on until all traces of the original attention gradient would have dissipated.

In an extension of this experiment, von Grünau and Faubert (1994) and Faubert and von Grünau (1995) studied the effect of Stimulus Onset Asynchrony (SOA) on split motion. In Figure 8b, two boxes of the same color (or luminance) are presented at two separate locations and one of the boxes is delayed. This is followed by a bar joining the two boxes. The bar appears to emerge from both boxes. If the SOA is zero, this split motion collides in the middle. The collision point moves closer to the first box as SOA is increased. In Figure 8c, the boxes in Frame 1 are of two different colors (or luminances) and the bar in Frame 2 matches in color (or luminance) one of the boxes. The motion is always perceived away from the box that matches the color (or luminance) of the bar as indicated by the arrow in Figure 8c, and a motion in the opposite direction is seen when the bar is removed. Faubert and von Grünau, (1995) did not report any effect of varying SOA on the perceived direction of motion and attribute this to attribute priming effects that override any attentional gradient. Tse and Cavanagh (1995) also reported a similar experiment as a case against the gradient of attention argument.

FRAME-1

FRAME-2

FRAME-1

FRAME-2

L   R

(a)

FRAME-1

FRAME-2

FRAME-3

L   R

(b)

FRAME-1

FRAME-2

L   R

(c)

FRAME-1

FRAME-2

L   R

(d)

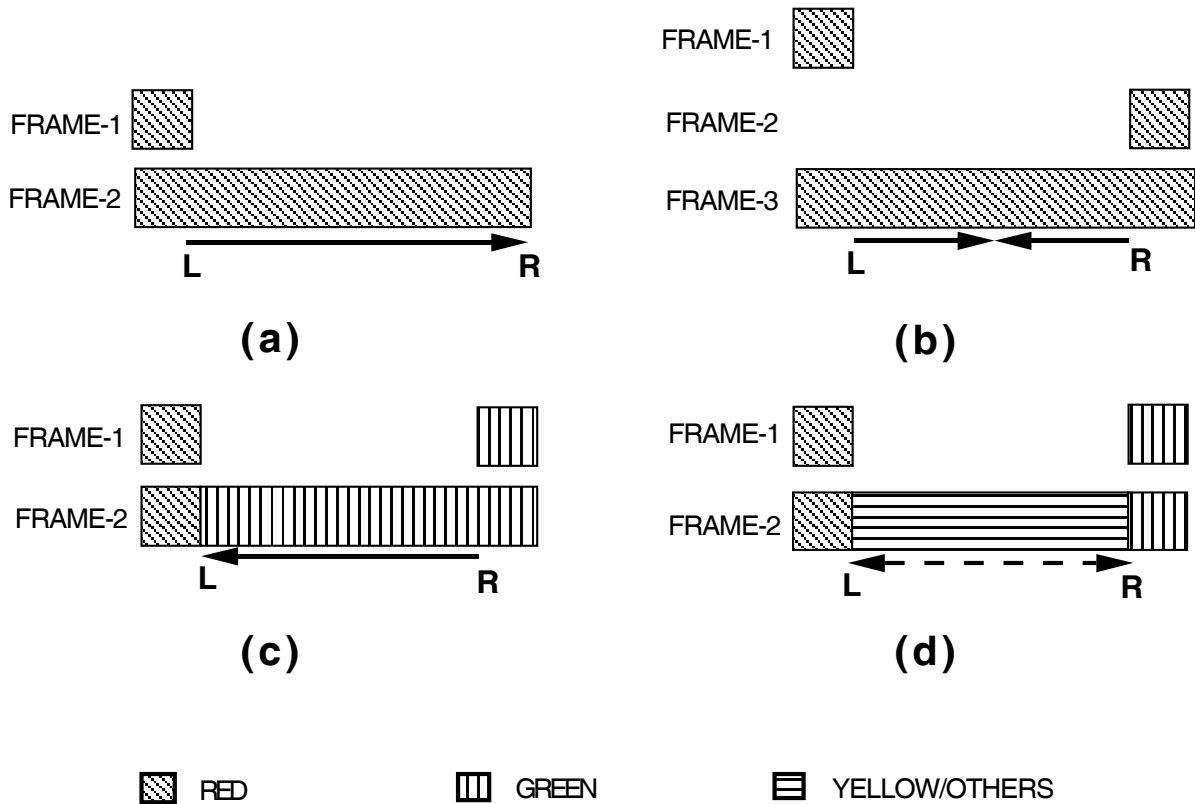RED   GREEN   YELLOW/OTHERS

Figure 8: Formotion experiments: The stimuli in each frame are presented discretely but a continuous motion is observed as indicated by the arrow. See text for details.

Let us now examine these cases in view of the neural models of boundary formation, color filling-in and form-motion interaction that were outlined earlier. The first factor is the formation of the horizontal boundary of the bar in Frame 2. In Figure 8a, the bipole cells that are closer to the horizontal edges of the box (i.e., near **L**), start out with signals that are larger than for the bipole cells that are away from the box. When the bar is presented, the bipole cells near **L** reach their threshold earlier than the cells that are away from **L**. Therefore the horizontal boundaries of the bar grow from **L** to **R** and a wave of boundary completion ensues. When the bar is removed, the bipole cells closer to the box continue to receive bottom-up signals from the horizontal edges of the box and therefore decay slowly. Hence, the portions of edges closer to the box persist longer than those away from the box. A wave of boundary erosion ensues from **R** to **L**.

The second factor is color filling-in. When the bar is presented next to a box in Frame 2, the boundary at **L** collapses quickly and color fills in from the left in its FIDO. The diffusion is bounded by new horizontal bar boundaries that grow from **L** to **R** and the new vertical boundary at **R**. The third factor is the form-motion interaction and formation of a motion G-wave. The offset of the edge at **L** and the onset of the edge at **R** generates a G-wave from **L** to **R**. Similarly, when the bar is removed, the edge at **R** decays and the edge at **L** grows, so a G-wave is generated from **R** to **L**. Lastly, these bottom-up motion signals attract spatial attention.

Now consider split motion with same-color boxes and varying SOA (see Figure 8b). When the SOA is zero (i.e., Frame 2 coincides with Frame 1), the horizontal boundary growth and color filling-in favor both directions equally while G-wave formation does not favor either direction. For

example, the horizontal boundaries of the bar in Frame 3 receive long-range cooperative bipole signals from the horizontal edges of both boxes and grow simultaneously from both sides to meet in the middle of the bar. Similarly, vertical edges at **L** and **R** are removed simultaneously and color fills-in from both ends to meet in the middle. Finally, both vertical edges decay at the same rate and no G-wave is generated. The non-zero SOA cases are more involved, since we need to understand what happens to the bottom-up spatiotemporal signals when the signals corresponding to the box that appears later are still growing and the bar is presented. For example, the boundary signals corresponding to the box that appears later may still be growing when the bar is presented in Frame 3 while the boundary signals corresponding to the box that appeared earlier may have reached their peak values, depending on the SOA. Therefore, the bipole cells of the horizontal edge closer to the earlier box are at an advantage and grow faster than the ones closer to the box that appears later. As a result, the boundary wave from the left progresses further than the boundary wave from the right.

But this the opposite of what is observed! Does this mean that the model is wrong? A more probing analysis shows that this is not so. The first thing to note is that the boundary signals do not themselves activate the motion system. *Transients* of the boundary signals activate the motion system. Transient cells respond to boundary signals in order to prevent a stationary but sustained boundary from relentlessly generating motion signals. In all of the previous examples with boundary waves, either a single wave existed, so the transients tracked that wave, or a pair of equal waves existed, as is the zero SOA case, so the transients did not favor either one. In the non-zero SOA case, by contrast, transients of the boundary signals favor the second box over the first box, because it generates larger transient signals when the test bar occurs. Let us now consider the decay of vertical boundaries of the boxes that they share with the bar. Since the boundary signals for the second box may be smaller than the boundary signals for the first box, depending on SOA, the boundary at **R** collapses faster than the boundary at **L**. A G-wave therefore progresses from the second box to the first box. Finally, due to this order of boundary decay, color filling-in progresses further from the second box than the first box. These various effects are simulated below. They particularly support the model's claim that transients of the boundary signals activate the motion system. This property was used in Francis and Grossberg (1996a) to simulate Korte's laws (Korte, 1915). The model hereby suggests an unsuspected mechanistic link between Korte's laws and the Faubert-von Grünau split motion data.

For the attribute priming split motion experiment with different colored boxes and zero SOA (Figure 8c), horizontal boundary growth does not favor either direction as the same amount of support is available from both sides. However, the build-up of green color near the red box (at **L**) is slower because the opponent red color in the off-surround inhibits the green in the on-center. When the vertical boundary at **R** collapses, the green color rapidly flows from the right and a color filling-in wave ensues. When the bar is removed, the red in the off-surround inhibits the green so the green-sensitive cells decay faster at the left, and the color erodes from **L** to **R**. For the G-wave factor, when the green bar is presented at Frame 2, the boundary at **R** decays and simultaneously the red/green boundary at **L** grows, or at least decays more slowly than the boundary at **R**. In either case, a G-wave is generated from the fast decaying signals at **R** to the growing signals or slow decaying signals at **L**. When the bar is removed, the red/green boundary at **L** decays while the boundary at **R** either grows due to transient responses to the green input, or at least decays more slowly than that at **L**, especially if the red and green are isoluminant. In either case, G-wave is generated from the fast decaying signals at **L** to the growing signals or slow decaying signals at **R**. Lastly, these bottom-up motion signals can attract spatial attention.

Faubert and von Grünau (1995) showed that a non-zero SOA had much less of an effect in an attribute priming experiment in which the boxes had different colors. There was always a

strong tendency for motion percept to emerge from the box that matched the color of the test bar. The model suggests that this happens because the color of the test bar and its matching box are processed by a different filling-in domain than the box with an opponent color. Color can thus flow *only* from the matching box towards the test bar in this case. This property helps to explain why the percept seems to grow from the box which matches the color of the test bar even if the SOA is non-zero.

The previous analysis suggests, however, that the boundary transient wave and G-wave favor the reverse direction in the non-zero SOA case. Does this mean that the color wave is in opposition to the boundary wave and G-wave? Several factors work against this conclusion in the attribute priming case. For one, when the test bar turns on, it causes a rapid decay of the boundary that it shares with the box of the same color. The boundary between the test bar and the box of opponent color decays more slowly, or may even grow under some circumstances. Thus a G-wave forms in the same direction as the color wave. When the color wave inputs to the motion stream via a form-motion interaction, it can join the G-wave to strengthen their combined effect.

Additional factors also work against the boundary transients favoring the second box. One such factor is that the onset of the test bar does not obliterate the vertical boundary which it makes with the box of opponent color. As a result, orientational competition from this boundary can slow down the growth of the horizontal boundary from the second box, along with its transients. In addition, in the full 3-D version of FACADE theory (Grossberg, 1994, 1997a), it is shown that filled-in surface representations send feedback to the boundaries that support them, thereby confirming and strengthening these boundaries while inhibiting boundaries corresponding to larger distances from the observer. This operation realizes a *boundary-surface consistency* property. Thus as a color wave forms, it sends positive feedback to the boundary system which can force the boundary wave to grow in the same direction as the color wave.

Faubert and von Grünau (1995) also investigated the non-zero SOA case under dichoptic presentation in which the first box or second box was presented to the same eye as the test bar, while the other box was presented to the other eye. They showed that the eye of origin influenced the percept strongly, specially at short SOAs. This effect is clarified once again by the 3-D version of FACADE theory, which suggests why the first stage of color filling-in takes place in *monocular* filling-in domains. Thus the first filling-in event will be biased by the box that appears in the same eye as the test bar. This filling-in bias will tend to alter the percept much as in the attribute priming experiment with non-zero SOA. In addition, however, the dichoptic presentation will also engage slower binocular interactions that are discussed in Grossberg (1994, 1997a).

An experimental variation developed by us of the split motion experiment with attribute priming is summarized in Figure 8d. This experiment studies the effects of top-down attentional priming on formotion percepts by neutralizing bottom-up feature factors. In Frame 1, a green box is presented on the right and a red box on the left. In Frame 2, a bar is switched on and off periodically between these boxes. The bar starts with green color, changes to cyan, then yellow, then orange, then red, and finally to yellow again. Each color is presented for ten trial cycles each. Motion is perceived from the green box to the red ($\mathbf{R}$ to $\mathbf{L}$) for green, cyan and yellow bars. It changes direction from red to green ($\mathbf{L}$ to $\mathbf{R}$) for orange, red and yellow bars. The important observation is the reversal of perceived direction of motion for yellow, which has equal red and green content. Since none of the bottom-up factors favors any direction of motion, top-down attentional priming may be inferred to determine the perceived direction of motion. This direction is primed by the perceived direction immediately prior to the presentation of yellow color. Such priming may be accomplished when bottom-up signals that code this direction activate top-down motion grouping signals as in Figure 1. These top-down signals persist until the yellow bar is presented, thereby priming the system to generate the observed hysteresis effect. These signals have elsewhere been
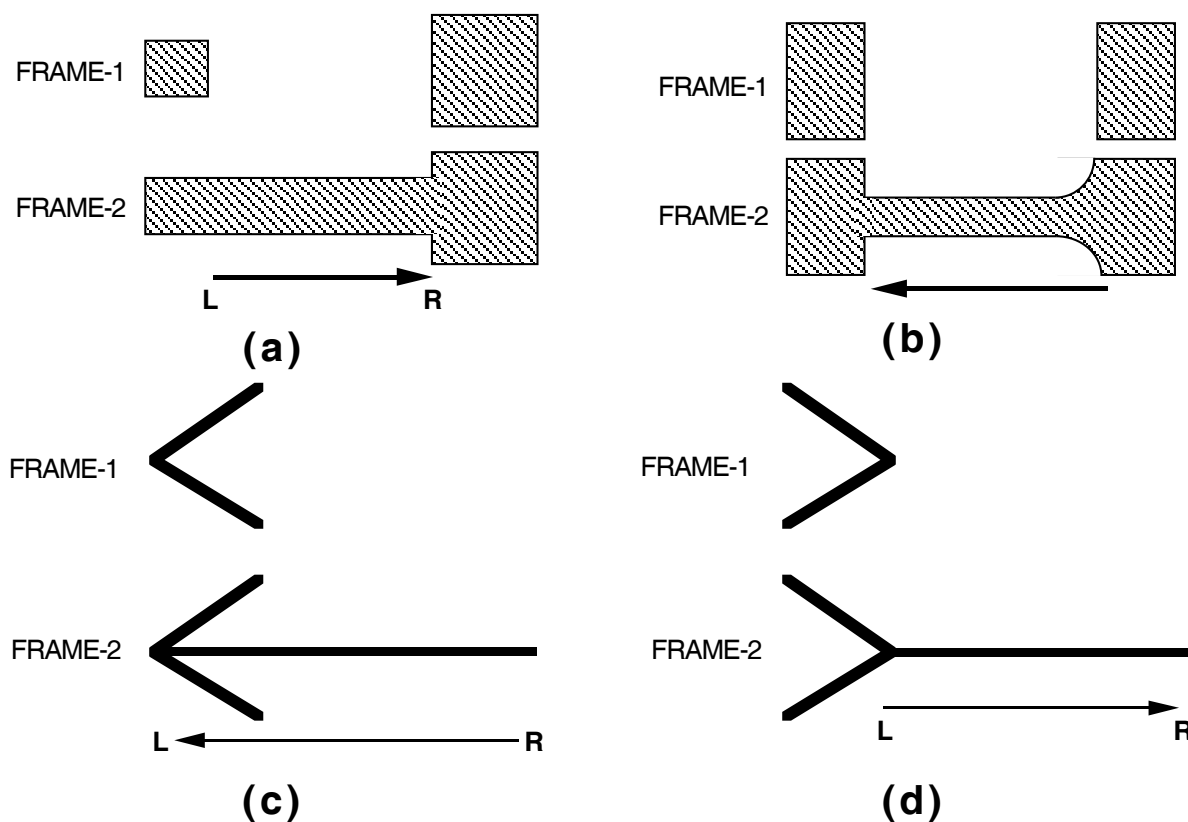
Figure 9: **Formotion when more complex figures change shape:** These experiments highlight some important features of boundary formation (short-range competition) and form-motion interaction (G-wave between signals that increase or decrease simultaneously). See text for details.

used to explain the motion capture that overcomes aperture ambiguities (Chey *et al*, 1997). They can also focus attention upon a primed direction of motion. Thus, in the absence of other stronger bottom-up factors, top-down attentional priming may have a rate-limiting effect upon the direction of perceived motion.

We now consider some of the formotion experiments in which more complex figures change shape (Tse *et al*, 1996). These experiments highlight some important features of boundary formation (short-range competition) and form-motion interaction (G-wave between signals that increase or decrease simultaneously) that were not responsible for motion perception in the experiments described in Figure 8. Figure 9 sketches four such experiments.

In Figure 9a a small box and a large box of the same color are presented at Frame 1. At Frame 2, the area between the two boxes is filled with a bar matching the height of the smaller box. A motion from the small box to the large box is reported, as indicated by the arrow. When the bar is removed (i.e., go from Frame 2 to Frame 1), motion in the opposite direction is reported. Consider the horizontal boundary formation along the bar. In addition to the long-range cooperation from horizontal edges of the small box (near **L**), the horizontal boundary close to the large box (near **R**) receives inhibitory signals from the vertical boundaries of the large box. This results in faster boundary growth at **L**. Therefore the horizontal boundary forms from the small box to the large box, and a wave of boundary completion ensues from **L** to **R**. When the bar is removed, long-range cooperation from the small box and short-range competition from the large box makes the

horizontal boundaries erode from **R** to **L**.

Similarly, in Figure 9b, the smooth contours at **R** receive long-range cooperative signals, since bipole cells pool signals from a range of orientations (Gove, Grossberg, and Mingolla, 1995; Grossberg and Mingolla, 1985b). As before, the vertical boundaries at **L** inhibit the horizontal boundaries. A wave of boundary completion ensues from **R** to **L**. The experiments in Figures 9c and 9d were described by Professor Shimojo during a personal communication (Miyauchi, Hikosaka, and Shimojo, unpublished) in support of our boundary formation model. The arrowhead in Figure 9c competes with the horizontal line, since orientational competition is spread over a range of orientations, peaking at perpendicular ones (Gove, Grossberg, and Mingolla, 1995; Grossberg and Mingolla, 1987). This competition slows down boundary growth near the arrowhead. Thus the line is seen to grow towards the arrowhead. On the contrary, the arrowhead in Figure 9d supports the growth of the horizontal boundary in its neighborhood due to cooperative orientational pooling by horizontal bipoles of the "relatable" orientations of the arrowheads (Gove et al., 1991; Grossberg and Mingolla, 1985b; Kellman and Shipley, 1991). This expedites its growth so that the line appears to grow away from the arrowhead.

When the test bar is turned on during Frame 2 of Figure 9a, the middle portion of the vertical boundary at **R**, though decaying, is supported by the remaining active portions via long-range cooperation (they try to form an illusory contour). Therefore, the vertical boundary at **L** decays faster than the corresponding portion of the vertical boundary at **R**, even though both are same in length. This has two effects. First, the vertical boundary at **L** is removed earlier than the vertical boundary at **R** and color starts to fill-in from **L** to **R**. A color filling-in wave ensues from **L** to **R**. Second, a motion G-wave is generated from the fast decaying boundary to the slow decaying boundary (from **L** to **R**).

When the test bar is removed, the vertical boundary at **R** forms more quickly (due to long-range vertical bipole cooperation) than does the vertical boundary at **L**. A G-wave is hereby generated from the fast growing edge at **R** to the more slowly growing edge at **L**. The same arguments apply to the experiments summarized in Figure 9b-d.

The experiments that are summarized in Figure 10 illustrate that formotion obeys 3-D pop-out rules for modal completion, amodal completion and illusory contour formation (Baloch and Grossberg, 1996; Tse et al, 1996). In Figure 10a the bar appears to move from the left and complete amodally behind the Kanizsa triangle. Grossberg (1994) has modeled how the illusory contours and surface of the Kanizsa triangle pop-out on BCS and FCS representations that represent a closer depth than those which represent the bar. Once this obstruction of the bar boundaries has been eliminated, the bar boundaries can be collinearly completed "behind" the triangle. This happens on a BCS representation that represents a farther depth than the Kanizsa triangle. Then all of our prior remarks about boundary waves go through on this farther BCS representation. In addition, the offset and onset events of the bar may be mapped via a depth-selective form-motion interaction from the farther BCS boundary representation to the motion processing stream (Francis and Grossberg, 1996a; Grossberg, 1991), as indicated in Figure 1. Then a G-wave can be generated by the bar in its depth-selective motion network using mechanisms of the formotion wave theorem.

Figure 10b provides another example of modal and amodal completion and illustrates how formotion percepts can be derived from figures that are formed due to illusory contours. The models of boundary formation, color filling-in and form-motion interaction suggest that the same rules govern both real and illusory contours (Francis, Grossberg, and Mingolla, 1994; Grossberg and Mingolla, 1985a, 1985b). An illusory rectangle morphs into an illusory square (Figure 10c) much as a box morphs into a bar (Figure 8a). Collapse in Frame 2 of the illusory contour that formed the right edge of the bar in Frame 1 occurs when a new illusory contour forms on the right
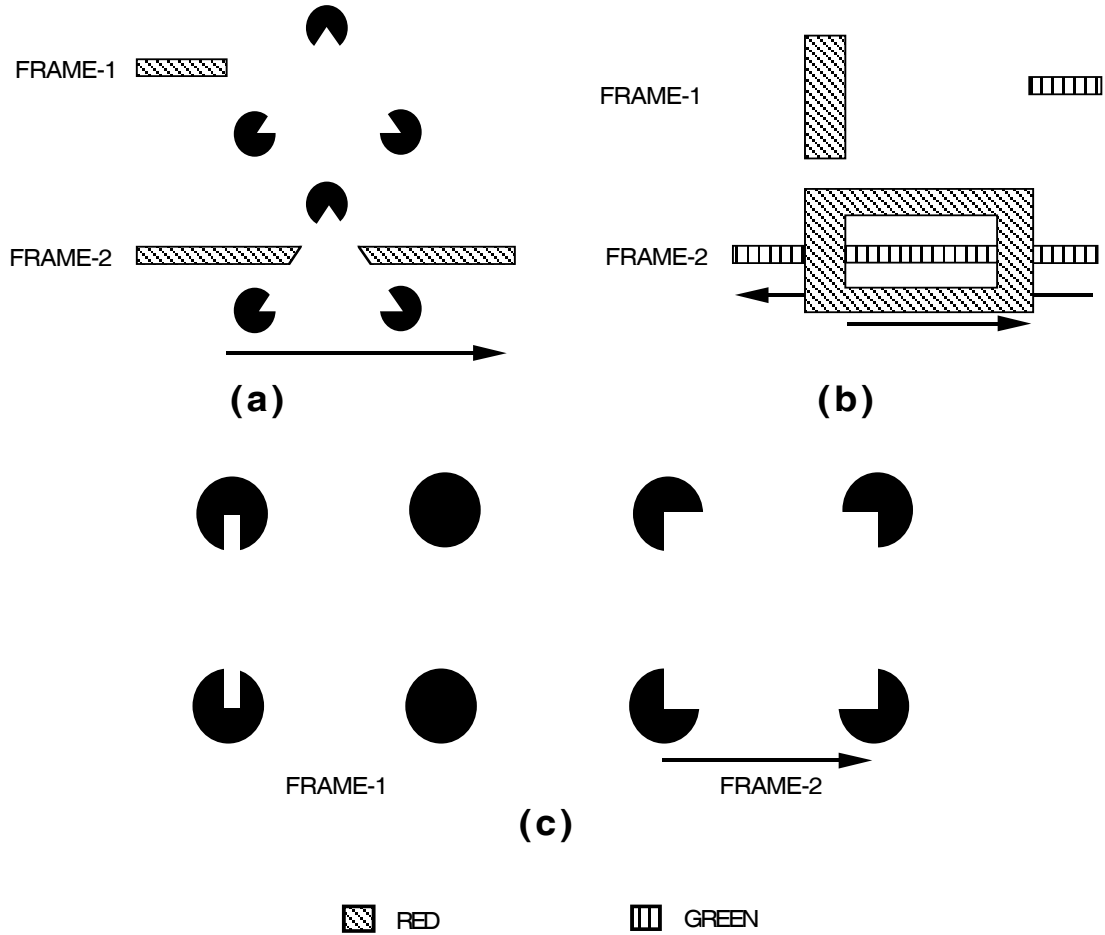
Figure 10: Morphing figures obey 3-D pop-out rules for modal completion, amodal completion, and illusory contours. See text for details.

edge of the square. A G-wave from left to right is hereby generated. In addition, the horizontal boundaries at the top and bottom of the bar enable two horizontal boundary waves to form from left to right at the bottom and top side of the illusory square.

7   Simulation of boundary waves

The remainder of the article describes simulations of these phenomena. A simplified version of the Boundary Contour System that models the boundary grouping and completion network is first simulated. Its equations and parameters are given in Appendix A. The experiment given in Figure 9a illustrates both long-range cooperation and short-range competition. The results of the simulations for this experiment are discussed here. Figure 11 shows the simulation layout and node assignment. Nodes 1-5 are assigned to the top horizontal edge of the box on the left, which is active all the time (i.e., in both Frames 1 and 2). The top horizontal edge of the bar is assigned nodes 6-25 and is active during Frame 2 only. The right vertical edge of the small box at **L** is assigned nodes 26-30 and is active during Frame 1 only. The left vertical edge of the box at **R** is assigned nodes 31-45. The middle portion of this edge (nodes 36-40) is active during Frame 1 only. Bipole cells of size 9 and a cross-orientational inhibitory region of size 5 are used. We simulate temporal dynamics of wave-like horizontal boundary growth (nodes 6-25) and the temporal decay of vertical edges at **L** (nodes 26-30) and **R** (nodes 36-40). The bottom horizontal edge has the same temporal dynamics as the top, so only the top is simulated. The different rates of vertical edge decay are
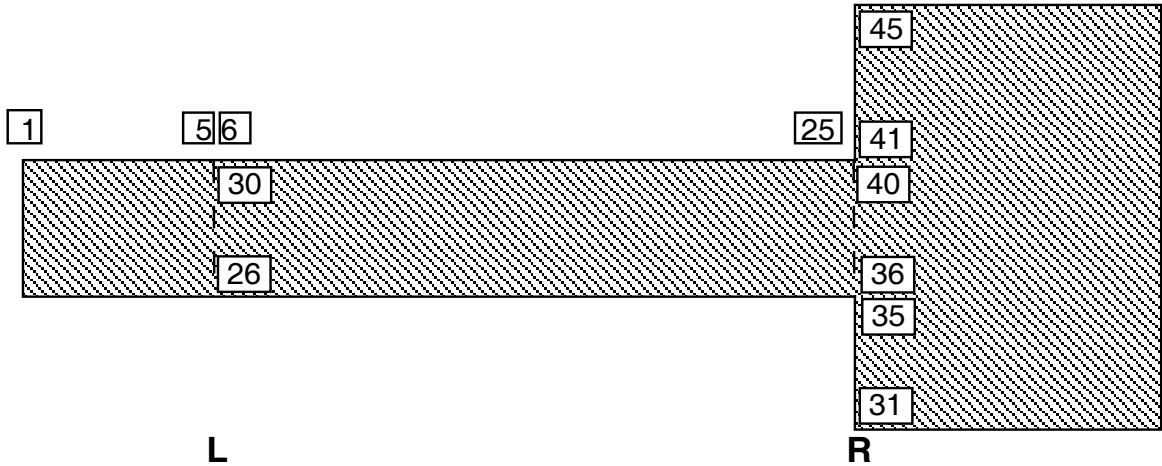
Figure 11: Boundary wave simulation layout and node assignment: The top horizontal boundary of the box at the left is assigned 5 nodes (1-5) and it is on during both Frames 1 and 2. The top horizontal boundary of the bar is assigned 20 nodes (6-25) and it is on during Frame 2 only. The vertical boundary at L is assigned 5 nodes (26-30) and it is on during Frame 1 only. The vertical boundary at R is assigned 15 nodes (31-45) and it is on during both Frames 1 and 2 except for the middle segment (node 36-40) which is on during Frame 1 only.

used later for motion G-wave simulations using the form-motion interaction model.

Figure 12 shows the output from the bipole cells (Figure 12a) and the hypercomplex cells (Figure 12b) of the first 5 nodes of horizontal edge of the bar (nodes 6-10). The activity is plotted on the vertical axis and the time on the horizontal axis. The time shown is from units 0.5 to 1.0 (i.e., the duration of Frame 2). Note that nodes to the left (closer to **L**) are activated earlier.

The time taken by bipole cells and hypercomplex cells corresponding to the horizontal bar to become active suprathreshold is plotted against each node in Figure 12c and 12d, respectively. A wave of boundary completion is seen from node 6 to node 25 (**L** to **R**). The effect of inhibition from the vertical boundary at **R** is also noticeable. The last five cells of the horizontal bar (cells 21-25) that receive cross-orientational inhibitory input from vertical boundaries are delayed even longer to become active suprathreshold, as indicated by an increase in the slope towards the right of the curve. Figure 12e and 12f shows how long it takes the vertical boundary at bipole and hypercomplex cells, respectively, to decay to a threshold value at **L** (nodes 26-30) and **R** (nodes 36-40). The boundary at **L** decays faster.

8 Simulation of illusory contour formation

To further test the model's ability to accomplish boundary completion, a parametric study of illusory contour formation was undertaken using the same equations and parameters as for the example discussed above. The gap over which a boundary can be completed is a function of model parameters, including the size of bipole cells. Simulations were carried out by varying the size of bipole cells and the size of the gap. The system is able to complete the boundary if the gap is approximately 27% or less of the size of a bipole cell. The result for bipoles of size 15 is given in Figure 13a and 13b, where the times it takes bipole cells and hypercomplex cells to become active
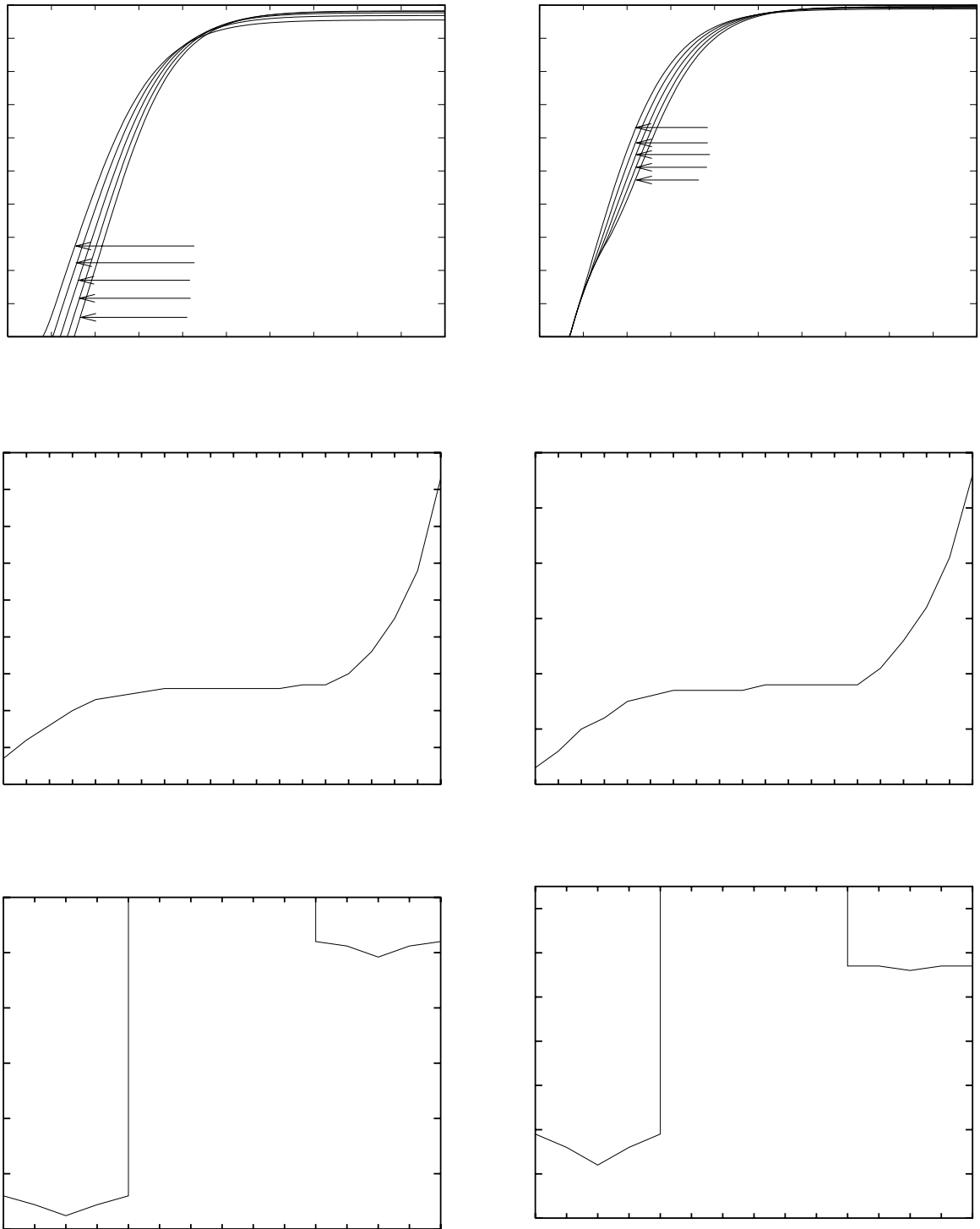
Figure 12: Simulation results of boundary formation: (i) Output from bipole cells (a) and hypercomplex cells (b) of the first 5 nodes of horizontal boundary of the bar. The boundary starts to grow from the left. (ii) The time taken by bipole cells (c) and hypercomplex cells (d) of the horizontal bar to become active suprathreshold. A wave of boundary completion is seen from node 6 (L) to node 25 (R). (iii) Time taken by bipole cells (e) and hypercomplex cells (f) of the vertical boundary at L (nodes 26-30) and the portion of vertical boundary at R (nodes 36-40) to decay subthreshold. The boundary at L decays faster than the boundary at R. See text for details.

suprathreshold are plotted against cell numbers 1-15. The gap was at nodes 6-9, as shown at the figure top. The broken portion of the line indicates the gap. Boundaries with more cooperative support at the right line form first, boundaries over the gap form last, and boundaries near the middle of the gap form before boundaries near the gap ends.

9   Simulation of color filling-in

The diffusion of signal in a filling-in domain is simulated to confirm that a wave of color filling-in is generated from an existing color region through a quickly collapsing boundary to the remaining boundaries. Simulation layout for a one-dimensional diffusion case (horizontal) and node assignment are given in Figure 14a. The equations and parameters are described in Appendix B.

The box at the left (nodes 1-5) is active during Frames 1 and 2. The bar (nodes 6-25) is active during Frame 2 only. The vertical boundary at **L** decays and the vertical boundary at **R** grows during Frame 2. The boundary signals are generated by the BCS model. The parameter values for BCS are the same as used in the boundary formation examples (given in Appendix A). The boundary signals gate (inhibit) the diffusion signals at both ends of the box (nodes 1 and 5) in Frame 1. When the bar is presented in Frame 2, the boundary gating signal at **L** decays and allows the diffusion to fill-in from the left while the boundary signal at **R** grows and blocks further filling-in.

Figure 14b plots the time when nodes become active suprathreshold on the vertical axis versus the cell numbers corresponding to the bar (nodes 6-25) on the horizontal axis. A wave of color filling-in from node 6 to 25 (**L** to **R**) is seen. This simulation uses the property that BCS boundaries are filling-in *generators* as well as filling-in *barriers*. To the present, only the barrier or gating property has been used. Grossberg (1987, 1994) showed that both properties are needed to explain data about 3-D color perception. In their capacity as filling-in generators, boundaries allow filling-in to occur only if they are collinearly interpolated between the opponent color responses to a stimulus. This property is realized by letting filling-in domains that represent opponent colors input to double-opponent cells. In the present example, the generator property implies that color signals become effective as their contiguous boundary signals do. Color filling-in hereby rides the boundary wave. Without this BCS-FCS interaction, the maximum time of filling-in occurs closer to **R** than **L**. With the BCS-FCS interaction, the maximum time occurs at **R**; see Figure 14 below.

10   Simulation of form-motion interaction

The key features of form-motion interaction as given in Figure 1 are simulated here. Since a number of results where a G-wave occurs from decaying signals to growing signals have been described elsewhere (Grossberg and Rudd, 1989, 1992), we describe here one example for the case when a G-wave is formed from fast decaying boundary signals to slow decaying boundary signals (Figure 9a). The simulation layout and node assignment are given in Figure 15a. The equations and parameter values are given in Appendix C.

The temporal responses of vertical edges at **L** and **R** from the BCS simulation described in Section 7 for boundary formation and decay are used here as input to the transient filter that mediates the form-motion interaction. For simplicity, edges at **L** and **R** are represented by single nodes (nodes 6 and 25 respectively). Both edges grow in Frame 1 and generate a transient ON response in the transient filter circuit. When the bar is presented during Frame 2, a transient OFF response is generated. The ON and OFF transient responses of boundaries at **L** and **R** are given in Figure 15b. The motion filters at these edge locations also respond to the onset of signals (ON
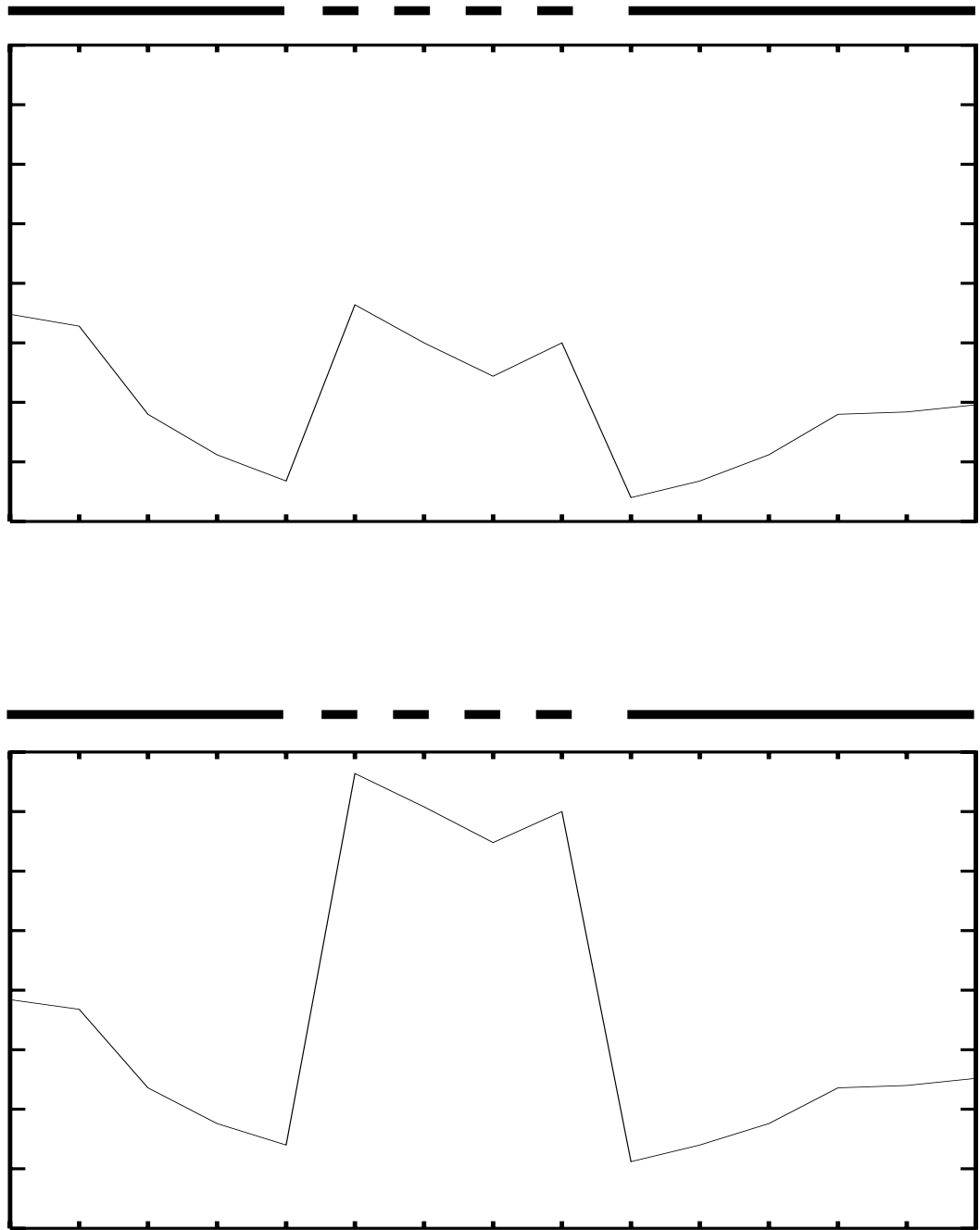
Figure 13: Simulation results of illusory contour formation: The time taken by the bipole cells (a) and the hypercomplex cells (b) to complete the boundary across the gap in a line (shown at the top of the plots). See text for details.
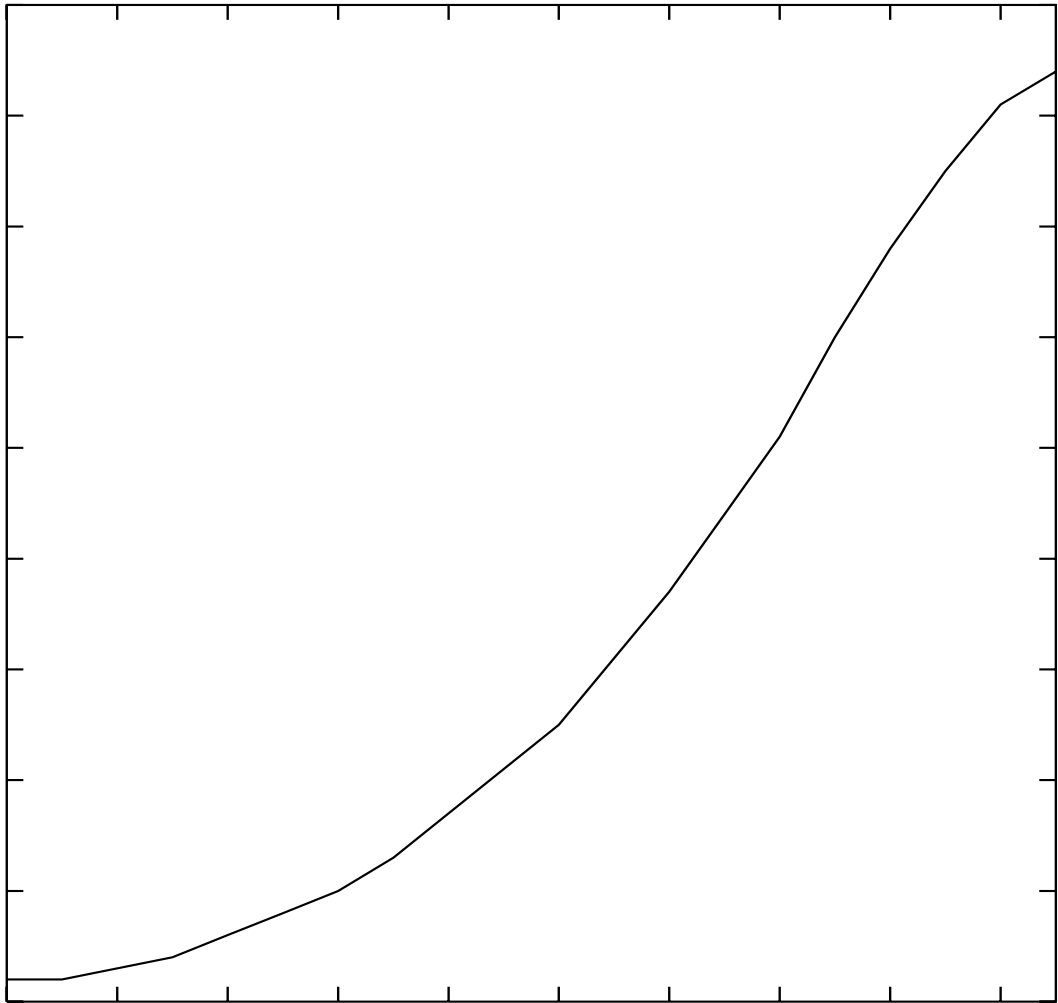
Figure 14: (a) Color filling-in wave simulation layout and node assignment. The box at the left is 5 nodes wide and the bar is 20 nodes wide. During Frame 1, the FIDO receives boundary signals at nodes 1 and 5. During Frame 2, the boundary signal at node 5 is removed and a new boundary signal at node 25 is activated. (b) Simulation result of color filling-in wave. The time taken by nodes corresponding to the bar (nodes 6-25) to become active suprathreshold is shown. The color fills in from node 6 (L) to node 25 (R).

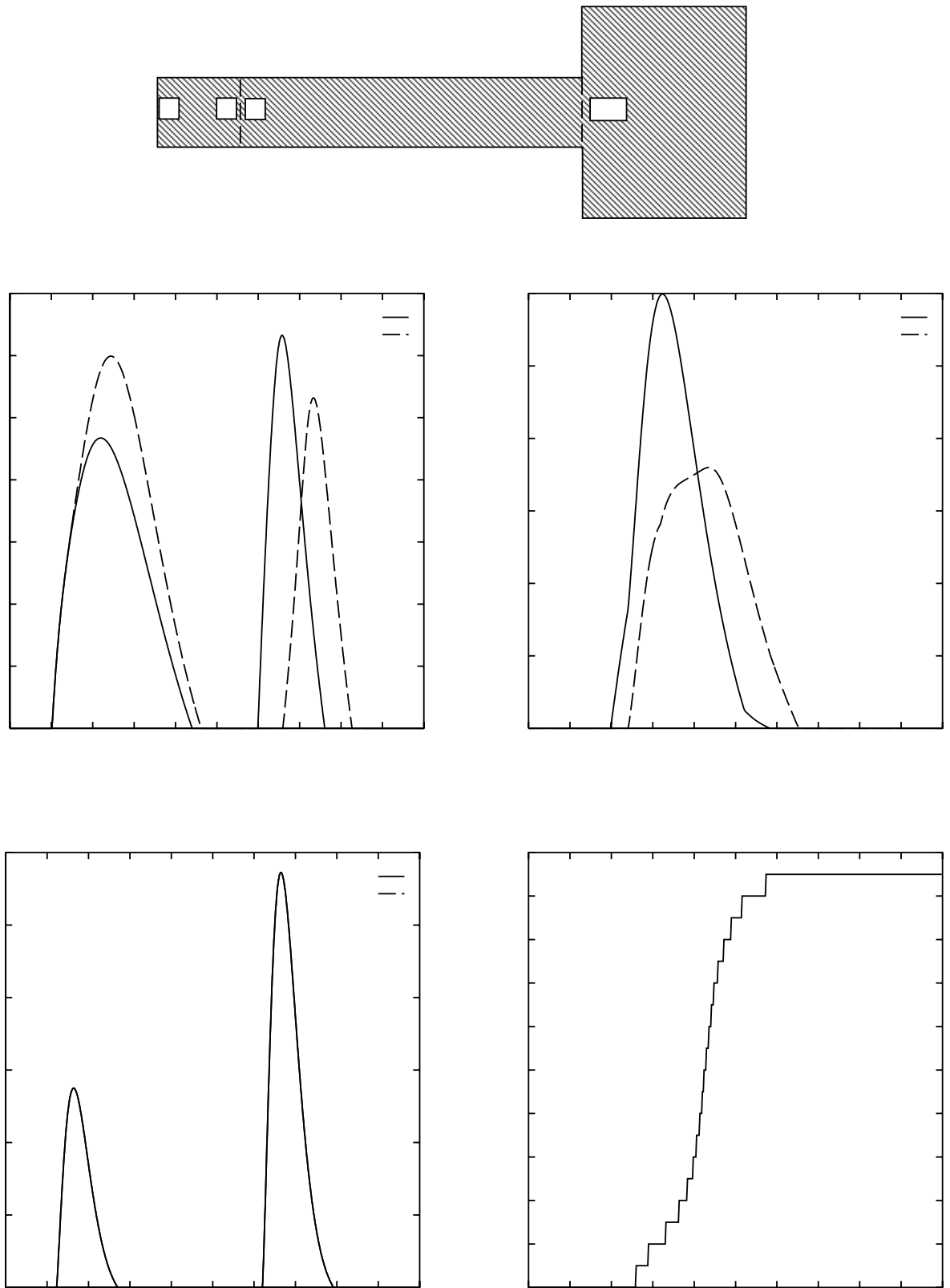Figure 15: Form-motion interaction simulation: (a) Simulation layout and node assignment. For simplicity, the boundaries at L and R are represented by single nodes. The box at the left is 5 nodes wide and the gap between the two vertical boundaries is 20 nodes. (b) Transient filter responses at L and R. (c) Low-level motion filter responses at L and R. (d) Total transient signals. (e) Motion G-wave through time. See text for details.

cells) in Frame 1 and transient ON responses are generated. When the bar is presented in Frame 2, its OFF cells are activated and transient OFF responses are generated. The responses from the ON and OFF channels of motion filter at **L** and **R** are given in Figure 15c.

Since the boundary at **L** (node 6) decays faster than the boundary at **R** (node 25), the transient OFF response from the transient filter at **R** occurs later than that at **L**. However, the motion filter at **L** and **R** receive simultaneous color inputs directly from the input representations (see Figure 1), so the transient ON and OFF responses at **L** and **R** overlap. The ON responses from the transient filter for edges at **L** and **R** also overlap during Frame 1 because both boxes at the left and right turn on simultaneously. However, starting at Frame 3, when the bar is removed, the decaying portion of the vertical boundary at **R** will be supported by the remaining vertical boundaries of the large box via long-range bipole cell cooperation and will grow faster than the vertical boundary at **L**, which does not have such support. Therefore, the ON response of the transient filter at **L** will occur later than that at **R**.

The transient responses are then relayed through the form-motion interaction. These signals during Frame 2 are shown in Figure 15d. The total response at **L** is earlier than the total response at **R**. These signals are passed through a long-range spatial filter in the motion stream (see Figure 1) and spatially compete at the motion wave layer. A G-wave results, as shown in Figure 15e.

11   Simulation of split motion

The split motion experiment illustrated in Figure 8b is simulated to demonstrate how boundary, surface, and motion waves formed from both sides collide, and how the collision point moves closer to the box that appears earlier as the SOA is increased (Faubert and von Grünau, 1995). Figure 16a shows the simulation layout and node assignment for the split motion experiment. Nodes 1-5 are assigned to the top horizontal edge of the box on the left, which is active all the time. Nodes 26-30 are assigned to the top horizontal edge of the box on the right which is turned on after a variable SOA and remain active thereafter. The bar is 20 cells wide and is assigned nodes 6-25, which becomes active after a fixed duration of 0.2 time units after the box on the right has been presented. We simulate the temporal dynamics of horizontal boundary growth (boundary wave), transients of boundary formation (boundary transient wave), color filling-in wave, and G-wave for varying SOA and observe the location of the collision point. SOA were 0.0, 0.06, 0.1, 0.2, and 0.4 time units. SOAs were selected such that their ratio to the fixed time at which the bar is presented after the second box is switched ON is the same as in the data by Faubert and von Grünau (1995).

Figure 16b shows the location of the collision point for boundary wave simulation as SOA is increased from 0.0 to 0.4. The collision point is seen moving from the center of the bar for SOA = 0.0 towards the second box (on the right) as the SOA is increased. Figures 16c-e show in contrast that, the *transients* of boundary formation, the direction of color filling-in, and the G-wave all favor motion from the second box to the first. Thus all of the factors which influence the motion system demonstrate the experimentally observed pattern of results. This fact lends additional weight to the hypothesis that transients of onset and offset events play a key role in determining motion percepts. The extent to which the collision point is shifted closer to the first box as the SOA is increased is determined by the values of the parameters. For example, the collision point for the G-wave is shifted closer to the first box than in Figure 16 if the parameters A, B and C are reduced; see Equations C1, C2 and C3; Appendix C.
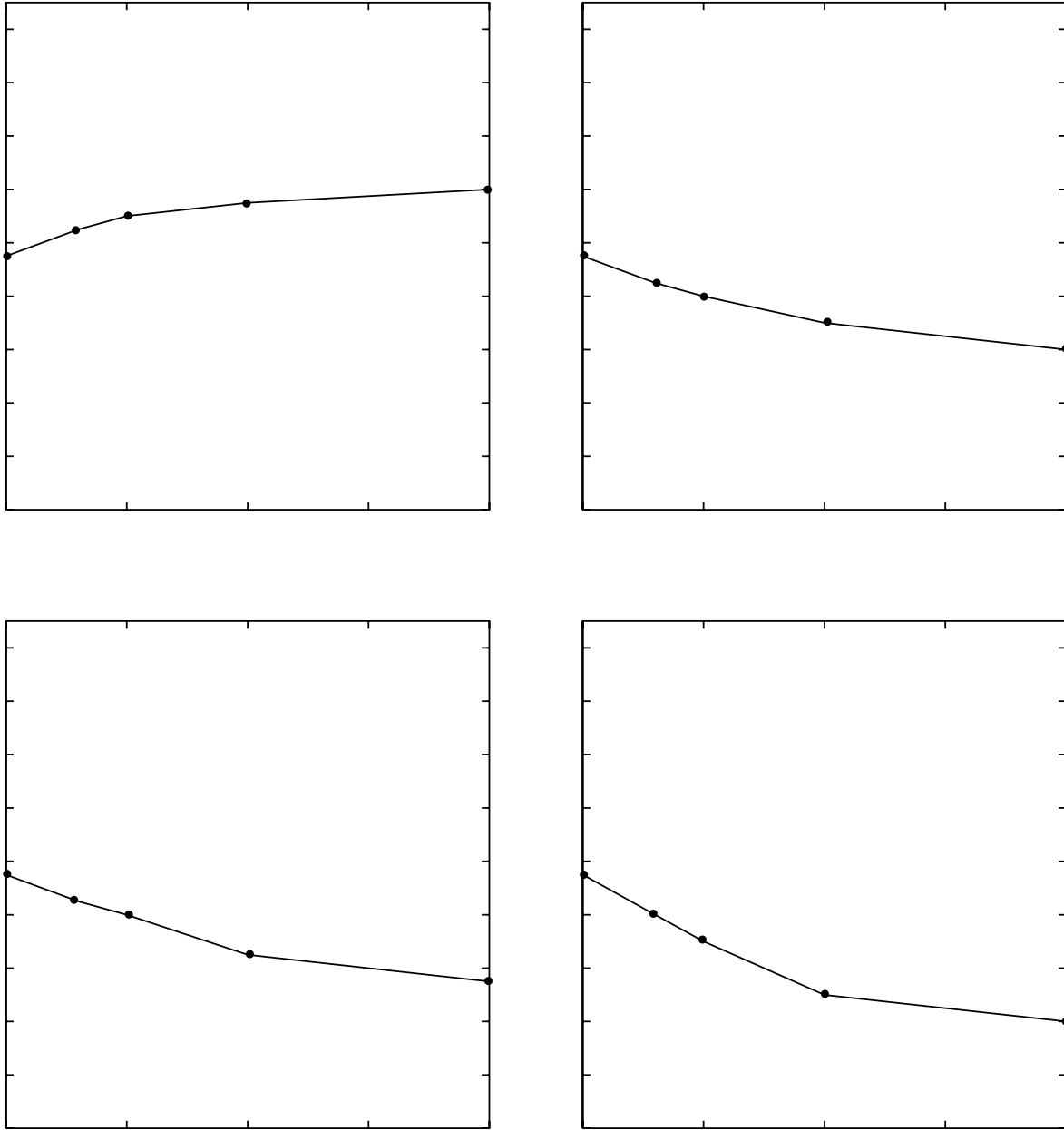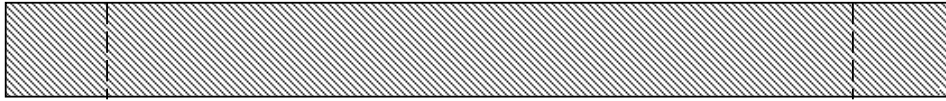
Figure 16: (a) Split motion simulation layout and node assignment. The top horizontal boundary of the box at the left is assigned nodes 1-5, and it is active all the time. The top horizontal boundary of the box on the right is assigned nodes 26-30. Its presentation is delayed by variable time (SOA) and it is active thereon. The horizontal portion of the bar is assigned nodes 6-25. Its presentation is delayed by a fixed time of 0.2 units after the box on the right is presented. (b) Boundary completion wave. (c) Boundary transients wave. (d) Color filling-in wave. (e) G-wave. See text for details.

12   Summary:

Experiments wherein spatially overlapping stimuli presented discretely in time generate continuous form-motion percepts have recently generated a great deal of interest. These percepts have been referred to as the illusory line motion illusion by Hikosaka, Miyauchi, and Shimojo (1993a), motion induction by von Grünau and Faubert (1994), and transformational apparent motion by Tse, Cavanagh, and Nakayama (1996). We call them a formotion percepts because they involve the active formation of percepts via a form-motion interstream interaction. The ongoing debate is whether formotion phenomena are due to a gradient of visual attention or as a consequence of bottom-up feature processing. Our results suggest that both opinions are partly correct, but that neither position, taken singly or together, is sufficient. Instead, in this paper we have demonstrated using simulations of some key experiments that formotion perception is a consequence of four major factors for visual information processing of spatiotemporal visual signals.

1. **Boundary completion:** Existing boundaries expedite growth of like-oriented boundaries in their neighborhood. This property has been modeled by a Boundary Contour System (BCS) wherein long-range cooperation between like-oriented boundary signals and short-range competition between dissimilarly oriented boundaries allow smooth contours to grow faster than abrupt boundaries when both are presented simultaneously.

2. **Color filling-in:** Color diffuses rapidly from existing color regions to new color boundaries. Also, opponent colors compete, which slows down the growth of new opponent colors in the neighborhood of an existing color surface. These properties are represented in a Feature Contour System (FCS). A Filling-In-DOmain (FIDO) receives color-inducing opponent signals which diffuse between existing boundaries computed by the BCS.

3. **Form-motion interaction:** We prove a Formotion Wave Theorem that states "a motion wave is generated from a decaying edge to growing edge, from a fast decaying edge to slow decaying edge, and from a fast growing edge to a slow growing edge under appropriate spatiotemporal conditions". These combinations are consequences of a form-motion interaction wherein signals from low-level motion filters interact with transients of boundary signals through a long-range spatial filter. How the system combines effects of the boundary and motion systems contributes to all the examples discussed herein.

4. **Spatial attention:** Bottom-up motion signals attract spatial attention by activating the top-down process whereby motion capture occurs. In the absence of other stronger bottom-up signals, earlier motion signals or directed attention may prime the direction of perceived motion in later trials. One example of this top-down priming is given in Figure 8d. The case of a yellow colored bar between a red box and green box balances all of the bottom-up feature factors, as both ends provide equal signals to horizontal boundary formation and both vertical boundaries decay simultaneously. In such a case, top-down attentional priming can play a rate limiting role. If the motion in previous trials was from a green box, the motion is seen from that direction for the yellow bar. However, if the motion in previous trials was from a red box, motion seems to emerge from that direction when the color of the bar is switched to yellow. Thus, motion in either direction is seen based on top-down attentional priming of motion direction.

The models of boundary formation, color filling-in and form-motion interaction follow the same rules for morphed illusory contours as for real boundaries and surfaces. Formotion figures also obey 3-D pop-out rules for modal completion, amodal completion and illusory contour formation.

FACADE theory suggests how such 3-D boundaries are completed amodally, or modally, and why Kanizsa figures pop-out (Grossberg, 1994, 1997a). The 3-D examples in Figure 10 can be explained by FACADE theory using these 3-D mechanisms in combination with the form-motion interstream interaction.

One fact that complicates understanding of formotion percepts is that not all of the above factors contribute equally, or at all, in every experiment. For example, in the line drawings of Figures 9c and 9d, boundary waves seem to produce the main motion sensation. To determine the extent of contribution by each of these factors for each experiment requires more parametric experimental protocols.

One way to approach such a study is to develop experiments wherein the effects of some factors negate each other and by controlling how much each factor contributes. One such experiment that we are currently studying starts out with a red box in Frame 1 at the right. In Frame 2, a bar is presented that forms a bent boundary with the box at right. It is the same experiment as given in Figure 9a without the small box on the left, so that there is no long-range cooperative boundary signal from the left. Now horizontal boundary formation favors motion to the right because the vertical edge of the box inhibits horizontal boundary growth in its vicinity. In contrast, color filling-in favors motion to the left. In addition, a leftward G-wave is generated from the right vertical decaying edge to the left vertical growing edge. The motion percept in this experiment thus depends upon the extent of orientational inhibition. There are also some directed top-down attentional priming effects. Experiments such as these, that trade off transient factors within the boundary, surface, and motion streams, may prove valuable as a diagnostic tool for teasing apart the several parallel mechanisms that contribute to formotion percepts.

References

Albright, T.D., Desimone, R., and Gross, C.G. (1984). Columnar organization of directionally sensitive cells in visual area MT of the macaque. *Journal of Neurophysiology*, **51**, 16-31.

Allman, J., Miezin, F., and McGuinness, E. (1985). Direction and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT). *Perception*, **14**, 105-126.

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

Baloch, A. and Grossberg, S. (1996). Neural dynamics of morphing motion. *Investigative Ophthalmology and Visual Science*, **37**, 3419.

Baloch, A., Grossberg, S., Mingolla, E., and Nogueira C.A.M. (1996). A neural model of first-order and second-order motion perception and magnocellular dynamics. Technical Report CAS/CNS-TR-96-030. Boston, MA: Boston University.

Cavanagh, P. (1992). Attention based motion perception. *Science*, **257**, 1563-1565.

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

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

Chey, J., Grossberg, S., and Mingolla, E. (1997). Neural dynamics of motion grouping: From aperture ambiguity to object speed and direction. *Journal of the Optical Society of America - A*, in press.

Chubb, C. and Sperling, G. (1989). Two motion perception mechanisms revealed through distance-driven reversal of apparent motion. *Proceedings of the National Academy of Sciences - USA*, **86**, 2985-2989.

Faubert, J. and von Grünau, M. (1992). The extent of split attention and attribute priming in motion induction. *Perception*, **21**, 105b.

Faubert, J. and von Grünau, M. (1995). The influence of two spatially distinct primers and attribute priming on motion induction. *Vision Research*, **35**(22), 3119-3130.

Francis, G. and Grossberg, S. (1996a). Cortical dynamics of form and motion integration: Persistence, apparent motion and illusory contours. *Vision Research*, **36**(1), 149-173.

Francis, G. and Grossberg, S. (1996b). Cortical dynamics of boundary segmentation and reset: persistence, afterimages, and residual traces. *Perception*. In Press.

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

Goodale, M.A. and Milner, D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, **15**, 20-25.

Gove, A., Grossberg, S. and Mingolla, E. (1995). Brightness perception, illusory contours, and corticogeniculate feedback. *Visual Neuroscience*, **12**, 1027-1052.

Groner, R., Hofer, D., and Groner, M. (1986). The role of anticipation in the encoding of motion signals-sensitization or bias. In F. Klix and Hagendorf (Eds.), **Human memory and cognitive capabilities**. Amsterdam: Elsevier.

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

Grossberg, S. (1977). Apparent motion. Unpublished manuscript.

Grossberg, S. (1980). How does a brain build cognitive code? *Psychological Review*, **87**, 1-51.

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

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

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

Grossberg, S. (1997a). Cortical dynamics of 3-D figure-ground perception of 2-D pictures. Technical Report CAS/CNS-TR-95-013. Boston, MA: Boston University, *Psychological Review*, in press.

Grossberg, S. (1997b). How is a moving target continuously tracked behind occluding cover? In Watanabe, T. (Ed.), **High level motion processing**, Cambridge, MA: MIT Press, in press.

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

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

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

Grossberg, S., Mingolla, E., and Ross, W. (1994). A neural theory of attentive visual search: Interactions of boundary, surface, spatial, and object representations. *Psychological Review*, **101**(3), 470-489.

Grossberg, S., Mingolla, E., and Ross, W. (1997). Visual brain and visual perception: How does the visual cortex do perceptual grouping. *Trends in Neurosciences*, **20**(3), 106-111.

Grossberg, S., Mingolla, E., and Williamson, J. (1995). Synthetic aperture radar processing by a multiple scale neural system for boundary and surface representation. *Neural Networks*, **8**, 1005-1028.

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

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

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

Hikosaka, O., Miyauchi, S., and Shimojo, S. (1993a). Focal visual attention produces illusory temporal order and motion sensation. *Vision Research*, **33**(9), 1219-1240.

Hikosaka, O., Miyauchi, S., and Shimojo, S. (1993b). Voluntary and stimulus-induced attention detected as motion sensation. *Perception*, **22**, 517-526.

Kanizsa, G. (1951). Sulla polarizzazione del movimento gamma. *Archivio di Psccologia, Neurologia e Psichiatria*, **3**, 224-267.

Kanizsa, G. (1979). **Organization in vision: Essays in Gestalt Perception**. New York: Praeger Press.

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

Kenkel, F. (1913). Untersuchungen über den Zusammenhang zwischen Erscheinungsgröße und Erscheinungsbewegung bei einigen sogenannten optischen Täuschungen. *Zeitschrift für Psychologie*, **67**, 358-449.

Korte, A. (1915). Kinematoskopische Untersuchungen. *Zeitschrift für Psychologie*, 194-296.

Kwak, H.-W., Dagenbach, D., and Egeth, H. (1991). Further evidence for a time-independent shift of the focus of attention. *Perception and Psychophysics*, **49**, 473-480.

Livingstone, M.S. and Hubel, D.H. (1984). Anatomy and physiology of a color system in the primate visual cortex. *Journal of Neuroscience*, **4**, 309-356.

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

Malik, J. and Perona, P. (1990). Preattentive texture discrimination with early vision mechanisms. *Journal of the Optical Society of America A*, **7**, 923-932.

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

Mishkin, M., Ungerleider, L.G., and Macko, K.A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences*, **6**, 414-417.

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

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

O'Craven, K.M., Rosen, B.R., Kwong, K.K., and Savoy, R.L. (1996). Detecting the effects of voluntary attention on a visual motion processing region in human cortex. Submitted for publication.

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

Remington, R. and Pierce, L. (1984). Moving attention: Evidence for time-invariant shifts of visual selective attention. *Perception and Psychophysics*, **35**, 393-399.

Schiller, P.H., Logothetis, N.K., and Charles, E.R. (1990). Role of color-opponent and broad-band channels in vision. *Visual Neuroscience*, **5**, 321-346.

Shimojo, S., Miyauchi, S., and Hikosaka, O. (1992). Visual motion sensation yielded by non-visually driven attention. *Investigative Ophthalmology and Visual Science*, **32**, 1354.

Sekuler, R. and Ball, K. (1977). Mental set alters visibility of moving targets. *Science*, **198**, 60-62.

Steinman, B.A., Steinman, S.B., and Lehmkuhle (1995). Visual attention mechanisms show a center-surround organization. *Vision Research*, **35**, 1859-1869.

Steinman, B.A., Steinman, S.B., and Lehmkuhle (1996). Transient visual attention is dominated by the magnocellular stream. *Vision Research*. In Press.

Stelmach, L.B. and Herdman, C.M. (1991). Directed attention and perception of temporal order. *Journal of Experimental Psychology: Human, Perception and Performance*, **17**, 539-550.

Stelmach, L.B., Herdman, C.M., and McNeil, R. (1994). Attentional modulation of visual processes in motion perception. *Journal of Experimental Psychology: Human, Perception and Performance*, **20**, 108-121.

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

Takeichi, H., Watanabe, T., and Shimojo, S. (1992). Illusory occluding contours and surface formation by depth propagation. *Perception*, **21**, 177-184.

Tanaka, K., Sugita, Y., Moriya, M., and Saito, H.A. (1993). Analysis of object motion in the ventral part of the medial superior temporal area of the macaque visual cortex. *Journal of Neuroscience*, **69**, 128-142.

Treue, S. and Maunsell, J.H.R. (1996). Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature*, **382**, 539-541.

Tse, P. and Cavanagh, P. (1995). Line motion occurs after surface parsing. *Investigative Ophthalmology and Visual Science*, **36**, 1919.

Tse, P., Cavanagh, P., and Nakayama, K. (1996). The role of parsing in high level motion processing. In Watanabe, T. (Ed.), **High level motion processing**, Cambridge, MA: MIT Press, in press.

Ungerleider, L.G. and Mishkin, M. (1982). Two cortical visual systems: Separation of appearances and location of objects. In D.L. Ingle, M.A. Goodale, and R.J.W. Mansfield (Eds.), **Analysis of visual behavior**. Cambridge, MA: MIT Press.

von Grünau, M., Dubé, S., and Kwas, M. (1996). Two contributions to motion induction: A preattentive effect and facilitation due to attentional capture. *Vision Research*, **36**, 2447-2457.

von Grünau, M. and Faubert, J. (1994). Intraattribute and interattribute motion induction. *Perception*, **23**, 913-928.

von Grünau, M., Racette, L., and Kwas, M. (1996). Measuring the attentional speed-up in the motion induction effect. *Vision Research*, **36**, 2433-2446.

Watanabe, T. and Cavanagh, P. (1992). Depth capture and transparency of regions bounded by illusory and chromatic contours. *Vision Research*, **32**, 527-532.

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

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

Appendices: Equations, parameters and theorems

A Boundary Wave Dynamics:

The equations and parameters follow of a simplified version of the BCS boundary grouping and completion network.

A.1 Hypercomplex Cells:

The hypercomplex cell activity x_i at position i obeys the equation

$$\frac{d}{dt}x_i = -Ax_i + (B - x_i)[I + G\ f(z_i)], \tag{A1}$$

where I is the bottom-up input and $f(z_i) = [z_i - \Gamma_1]^+$ is the feedback signal from the associated bipole cell.

A.2 Bipole Cells:

The bipole cell activity y_i at position i obeys the equation

$$\frac{d}{dt}y_i = -Dy_i + E\ g(L_i) + F\ g(C_i) + E\ g(R_i) - H\sum_{j \in \sigma_i} W_{ji}\ y_j, \tag{A2}$$

the weights W_{ji} are signal strengths from dissimilar oriented bipole cells, y_j, in a Gaussian neighborhood σ_i:

$$W_{ji} = \frac{1}{\sigma_i\sqrt{2\pi}}e^{-\frac{s^2}{2\sigma_i^2}}, \tag{A3}$$

and s is the spatial distance between nodes i and j. Terms $Eg(L_i)$, $Fg(C_i)$, and $Eg(R_i)$ define the long-range cooperative process and term $-H\sum_{j \in \sigma_i} W_{ji}y_j$ defines the short-range competitive process. The kernels L_i and R_i define the left and right lobes of the bipole receptive field, and kernel C_i defines the effect of a centered input on the cell body, where

$$L_i = \sum_{k=1}^{N}[x_{i-k} - \Gamma_2]^+, \tag{A4}$$

$$C_i = \eta\ [x_i - \Gamma_2]^+, \tag{A5}$$

$$R_i = \sum_{k=1}^{N}[x_{i+k} - \Gamma_2]^+, \tag{A6}$$

and parameter *2N+1* represents the size of a bipole cell receptive field, η equals N, and the function g(x) satisfies

$$g(x) = \frac{\alpha\ x}{\beta + x}. \tag{A7}$$

The simulations parameters are: A = B = 1.0, D = 10.0, E = F = 50.0, G = 1.0, H = 0.1, N=4, $\sigma_i = 5$, $\Gamma_1 = 0.75$, $\Gamma_2 = 0.25$, $\alpha = 1.0$ and $\beta = 2N + 1 = 9$ (the size of bipole cell).

B Color Filling-In Wave Dynamics:

The activity z_i of a cell in position i of a one-dimensional filling-in domain obeys the equation:

$$\frac{d}{dt}z_i = -Az_i + D\left(z_{i-1} - z_i\right) + D\left(z_{i+1} - z_i\right) + F_i. \tag{B1}$$

The input F_i is the signal from color-inducing cells gated by boundary signals. For simulations, A = 1.0, D = 50.0 and F_i was 1.0 when the boundary signal at position i was suprathreshold. Equation (B1) approximates the properties that the color cell activities quickly reach equilibrium, are (approximately) equal because the contrast of the bar is uniform across space, and their effect on filling-in is gated on whenever the boundary strength exceeds threshold.

C Motion G-Wave Dynamics:

Equations and parameters of a simplified form-motion interaction model outlined in Figure 1 are described here. The transient response of boundary signals is detected by a transient filter. An opponent processing circuit called a gated dipole is used as a transient filter (Grossberg, 1972, 1980). Offset of an input to the ON (or OFF) channel of such a circuit can generate an antagonistic rebound response in the OFF (or ON) channel. In the present application, the ON channel of the dipole circuit represents onset of a boundary signal and the OFF channel represents offset of the boundary signal. A dipole circuit is also used as a motion filter to detect the motion signals due to local intensity changes. A local increase in intensity is detected by an on-center off-surround network of ON cells and a local decrease in intensity is detected by an off-center on-surround network of OFF cells. In the present application, the ON channel of a dipole circuit receives input from ON cell and the OFF channel receives input from OFF cells. In this way, a transient ON response is generated by either the onset of an ON cell or the offset of an OFF cell (local increase in brightness). Similarly, an OFF response is generated by either the onset of an OFF cell or the offset of an ON cell (local increase in darkness). These transient responses are combined with the transient responses due to onset or offset of boundaries at those locations. These composite signals from the form and motion streams are passed through a long-range spatial Gaussian filter in the motion stream and spatially compete at the motion wave layer.

C.1 Gated Dipole Transient Filter:

A gated dipole circuit equations are as follows:
ON-Channel Input Stage

$$\frac{d}{dt}u_{1i} = -Au_{1i} + x_i + \gamma, \tag{C1}$$

where x_i is activity from the hypercomplex cells as described in (A1) and γ is a tonic arousal level.
Off-Channel Input Stage

$$\frac{d}{dt}u_{2i} = -Au_{2i} + \gamma, \tag{C2}$$

where γ is the same level of arousal as in (C1).
ON-Transmitter Production - Inactivation

$$\frac{d}{dt}v_{1i} = B(1 - v_{1i}) - C[u_{1i}]^+ v_{1i}, \tag{C3}$$

where $[w]^+ = max(w, 0)$ denotes half-wave rectification.

Off-Transmitter Production - Inactivation

$$\frac{d}{dt}v_{2i} = B(1 - v_{2i}) - C[u_{2i}]^+ v_{2i}. \tag{C4}$$

Transmitter-Gated ON-Activation

$$\frac{d}{dt}u_{3i} = -Au_{3i} + D[u_{1i}]^+ v_{1i}. \tag{C5}$$

Transmitter-Gated OFF-Activation

$$\frac{d}{dt}u_{4i} = -Au_{4i} + D[u_{2i}]^+ v_{2i}. \tag{C6}$$

Normalized Opponent ON-Activation

$$\frac{d}{dt}u_{5i} = -Au_{5i} + (E - u_{5i})u_{3i} - (F + u_{5i})u_{4i}. \tag{C7}$$

Normalized Opponent OFF-Activation

$$\frac{d}{dt}u_{6i} = -Au_{6i} + (E - u_{6i})u_{4i} - (F + u_{6i})u_{3i}. \tag{C8}$$

ON-Channel Output Stage

$$\frac{d}{dt}u_{7i} = -Au_{7i} + G[u_{5i}]^+. \tag{C9}$$

OFF-Channel Output Stage

$$\frac{d}{dt}u_{8i} = -Au_{8i} + G[u_{6i}]^+. \tag{C10}$$

ON Output:

$$u_i^{ON} = [u_{7i} - \Lambda]^+. \tag{C11}$$

OFF Output:

$$u_i^{OFF} = [u_{8i} - \Lambda]^+. \tag{C12}$$

The simulation parameters are: A = 10.0, B = 0.1, C = 0.75, D = 200.0, E = 100.0, F = 100.0, G = 200.0, γ = 10.0, Λ = 3.0.

C.2 Motion Filter:

A dipole circuit is used to represent the transient response to moving stimuli in the motion stream (Baloch *et al.*, 1996; Nogueira *et al.*, 1993). The ON-channel of the dipole responds to the net increase in a changing input while the OFF-channel responds to a net decrease. The dipole can hereby receive phasic input at either ON or OFF channels. The ON channel gets phasic input from DB (Dark-Bright) or s^+ inputs while the off-channel gets phasic input from BD (Bright-Dark) or s^- inputs. The dynamics of the dipole are the same as for the transient filter described earlier except for the following differences:

ON-Channel Input Stage

$$\frac{d}{dt}u_{1i} = -Au_{1i} + s_i^+ + \gamma. \tag{C13}$$

36

OFF-Channel Input Stage

$$\frac{d}{dt}u_{2i} = -Au_{2i} + s_i^- + \gamma. \tag{C14}$$

The simulation parameters are: A = 5.0, B = 0.25, C = 10.0, D = 100.0, E = 100.0, F = 100.0, G = 200.0, γ = 10.0, Λ = 0.5, s_i^+ = 1.0 for bright inputs, and s_i^- = 1.0 for dark inputs. These parameters enable direct motion inputs to generate faster transients than inputs from the form stream.

C.2.1 Motion Wave Layer:

Input:

$$I_i(t) = T_1(t)e^{-\frac{(i-1)^2}{2K^2}} + T_M(t)e^{-\frac{(M-i)^2}{2K^2}}, \tag{C15}$$

Where

$$T_i = \left[u_i^{ON}\right]^{\mathcal{TF}} + \left[u_i^{ON}\right]^{\mathcal{MF}} + \left[u_i^{OFF}\right]^{\mathcal{TF}} + \left[u_i^{OFF}\right]^{\mathcal{MF}}. \tag{C16}$$

Superscripts $\mathcal{TF}$ and $\mathcal{MF}$ represent Transient and Motion filters, respectively. Thus $T_1(t)$ is the sum of signals from transient and motion cells at position 1 and $T_M(t)$ is the sum of signals from Transient and Motion cells at position 2. Simulation parameters are: M = 20 and K = 11.

Output:

$$\frac{d}{dt}w_i = -Aw_i + (B - w_i)\,\eta\,I_i. \tag{C17}$$

Simulations parameters are: A = B = 1.0 and η = 2.0. We do not simulate a short-range spatial filter here because it does not significantly influence any of the results.

D Formotion Wave Theorem:

A motion G-wave may be generated from (a) a decaying signal to a growing signal, (b) a fast decaying signal to a slow decaying signal, and (c) a fast growing signal to a slow growing signal, if and only if the spatial distance between signals (L) is less than twice the size of Gaussian filter (K) i.e., $L < 2K$.

D.1 Proof: Decaying to growing (Grossberg, 1977)

Denote the output of the long-range Gaussian filter at the motion wave layer by T(w,t), where w varies over a continuum of cells. The activity x_0 at position 0 decays while the activity x_L at position L grows:

$$T(w,t) = x_0(t)\,e^{\frac{-w^2}{2K^2}} + x_L(t)\,e^{\frac{-(w-L)^2}{2K^2}}. \tag{D1}$$

Let $x_0(t)$ and $x_L(t)$ be defined by

$$\frac{dx_0}{dt} = -Ax_0 + J_0, \tag{D2}$$

and

$$\frac{dx_L}{dt} = -Ax_L + J_L, \tag{D3}$$

where $x_0(0) = x_L(0) = 0$. Then

$$x_0(t) = \int_0^t e^{-A(t-v)} \, J_0(v) \, dv, \tag{D4}$$

and

$$x_L(t) = \int_0^t e^{-A(t-v)} \, J_L(v) \, dv. \tag{D5}$$

Let input

$$J_0(t) = \begin{cases} J & \text{if } 0 \le t \le T \\ 0 & \text{if } t > T \end{cases}, \tag{D6}$$

and

$$J_L(t) = \begin{cases} J & \text{if } T + I \le t \le 2T + I \\ 0 & \text{if } t > 2T + I \end{cases}. \tag{D7}$$

where I is the ISI between the stimuli. Then for $T + I \le t \le 2T + I$,

$$x_0(t) = \frac{J}{A}(1 - e^{-AT})e^{-A(t-T)}, \tag{D8}$$

and

$$x_L(t) = \frac{J}{A}(1 - e^{-A(t-T-I)}). \tag{D9}$$

Substituting (D8) and (D9) in (D1) yields:

$$T(w,t) = \frac{J}{A}(1 - e^{-AT})e^{-A(t-T)} \, e^{\frac{-w^2}{2K^2}} + \frac{J}{A}(1 - e^{-A(t-T-I)}) \, e^{\frac{-(w-L)^2}{2K^2}}. \tag{D10}$$

The maximum values of $T(w,t)$ occur only at locations $w = w(t)$ such that

$$\frac{\partial T(w,t)}{\partial w} = 0. \tag{D11}$$

Such locations obey the equation

$$\frac{e^{A(t-T)} - e^{AI}}{1 - e^{-AT}} = \frac{w}{L - w} \, e^{\frac{L(L - 2w)}{2K^2}}. \tag{D12}$$

The function,

$$f(t) = \frac{e^{A(t-T)} - e^{AI}}{1 - e^{-AT}}, \tag{D13}$$

is an increasing function of t. We wish to determine when the positions $w = w(t)$ at which $T(w,t)$ is maximal increase as a function of t. In order for this to happen, the right hand side of Equation (D12), namely function

$$g(w) = \frac{w}{L - w} \, e^{\frac{L(L - 2w)}{2K^2}}, \tag{D14}$$

must also be an increasing function of w, for all $0 \le w \le L$, since then we can solve for

$$w = g^{-1}(f(t)) \tag{D15}$$

as an increasing function of w for all $0 \leq w \leq L$. Function $g(w)$ is monotone increasing if $g'(w) > 0$, which holds if and only if function

$$h(w) = (L - w)[1 - \frac{Lw}{K^2}] + w \tag{D16}$$

satisfies

$$h(w) > 0. \tag{D17}$$

In order for Equation (D15) to hold for all $0 \leq w \leq L$, the minimum of $h(w)$ for $0 \leq w \leq L$ must be positive. The minimum of $h(w)$ occurs at $w = L/2$, and equals

$$h\left(\frac{L}{2}\right) = \frac{L}{2}\left(2 - \frac{L^2}{2K^2}\right). \tag{D18}$$

The number $h(L/2)$ is positive iff $0 < L < 2K$.

D.2   Proof: Fast decaying to slow decaying

Again we start with the total input $T(w,t)$ to the motion wave layer. The activity at position 0 decays faster than the activity at position L. The function $T(w,t)$ is given in Equation (D1). The activity $x_0(t)$ at position 0 decays at a rate A for $T + I \leq t \leq 2T + I$ and is given in (D8). Similarly, the activity $x_L(t)$ at position L decays at a rate B for $T + I \leq t \leq 2T + I$ and is given by:

$$x_L(t) = \frac{J}{B}(1 - e^{-BT})e^{-B(t-T)}. \tag{D19}$$

Substituting (D8) and (D19) in (D1) yields:

$$T(w,t) = \frac{J}{A}(1 - e^{-AT})e^{-A(t-T)} \; e^{\frac{-w^2}{2K^2}} + \frac{J}{B}(1 - e^{-BT})e^{-B(t-T)} \; e^{\frac{-(w-L)^2}{2K^2}}. \tag{D20}$$

The maximum values of $T(w,t)$ occur only at locations $w = w(t)$ that obey (D11). Such locations obey the equation

$$\frac{A}{B} \frac{(1 - e^{-BT}) \; e^{-B(t-T)}}{(1 - e^{-AT}) \; e^{-A(t-T)}} = \frac{w}{L - w} \; e^{\frac{L(L - 2w)}{2K^2}}. \tag{D21}$$

The function

$$f(t) = \frac{A}{B} \frac{(1 - e^{-BT}) \; e^{-B(t-T)}}{(1 - e^{-AT}) \; e^{-A(t-T)}}, \tag{D22}$$

is an increasing function of t if $A > B$. The right hand side of (D21), which is the same as function g(w) in (D12), is an increasing function of w, for all $0 \leq w \leq L$ and T(w,t) is maximal iff $0 < L < 2K$. Therefore the maximum of total activity at the motion wave layer moves continuously from position 0 to L.

D.3 Proof: Fast growing to slow growing

Here in the total input $T(w, t)$, the activity at position 0 grows faster than the activity at position L. The function $T(w, t)$ is given in (D1). The activity $x_L(t)$ at position L grows at a rate A for $T + I \leq t \leq 2T + I$ and is given in (D9). Similarly, the activity $x_0(t)$ at position 0 grows at a rate B for $T + I \leq t \leq 2T + I$ and is given by:

$$x_0(t) = \frac{J}{B}(1 - e^{-Bt}). \tag{D23}$$

Substituting (D9) and (D23) in (D1) yields:

$$T(w, t) = \frac{J}{B}(1 - e^{-Bt}) \, e^{\frac{-w^2}{2K^2}} + \frac{J}{A}(1 - e^{-At}) \, e^{\frac{-(w-L)^2}{2K^2}}. \tag{D24}$$

The maximum values of $T(w, t)$ occur only at locations $w = w(t)$ that obey (D11). Such locations obey the equation

$$\frac{B}{A} \frac{1 - e^{-At}}{1 - e^{-Bt}} = \frac{w}{L - w} \, e^{\frac{L(L - 2w)}{2K^2}}. \tag{D25}$$

The function

$$f(t) = \frac{A}{B} \frac{1 - e^{-Bt}}{1 - e^{-At}}, \tag{D26}$$

is an increasing function of t if $B > A$. The right hand side of (D21), which is the same as function g(w) in (D12), is an increasing function of w, for all $0 \leq w \leq L$ and T(w,t) is maximal iff $0 < L < 2K$. Therefore the maximum of total activity at the motion wave layer moves continuously from position 0 to L.