HOW IS A MOVING TARGET CONTINUOUSLY TRACKED BEHIND OCCLUDING COVER?

Stephen Grossberg

January 1996

Technical Report CAS/CNS-96-001

Boston University Center for Adaptive Systems and
Department of Cognitive and Neural Systems
111 Cummington Street
Boston, MA 02215

HOW IS A MOVING TARGET CONTINUOUSLY TRACKED BEHIND OCCLUDING COVER?

Stephen Grossberg†

Department of Cognitive and Neural Systems
and
Center for Adaptive Systems
Boston University
677 Beacon Street
Boston, Massachusetts 02215 §

December, 1995

To appear in Watanabe, T. (Ed.)
High level motion processing
Cambridge, MA: MIT Press

Technical Report CAS/CNS-TR-96-001
Boston, MA: Boston University

Keywords: synchronization, binding problem, perceptual framing, visual cortex, neural network, boundary contour system, temporal order judgment, spatial pooling

---

†Supported in part by the Office of Naval Research (N00014-95-1-0657 and N00014-95-1-0409).

§Acknowledgments: The author wishes to thank Diana Meyers for her valuable assistance in the preparation of the manuscript.

1 Introduction: Form-Motion Interactions in the Forest Primeval

Imagine a predator or prey darting intermittently behind protective cover in a forest or jungle. As the animal moves, patterns of light and shade play upon its body through the overhanging foliage. These moving regions mingle with the movements caused on the animal's body as its limbs and muscles deform the textured patterns on its coat. The luminance and color contours between these moving regions may move in a variety of directions that do not necessarily point in the direction of the animal's physical movement. Rather, a scintillating mosaic of moving contours may be generated that could easily prevent its detection.

Detecting an animal under these rather typical conditions in the wild is obviously of great ecological importance to an observer, whether predator or prey. Doing so, however, requires that the visual system of the observer be able to solve several difficult perceptual problems, such as: How does the observer separate the forest cover from the moving animal itself, so that cover and animal can be independently recognized? In particular, how does the cover pop-out in front of the animal so that the contours of the animal are not confused with those of the cover? How is the scintillating mosaic of moving contours reorganized into a coherent object percept with a unitary direction-of-motion? How are the intermittent appearances of the animal integrated into a continuous motion percept that is amodally completed behind the cover, thereby enabling the animal to more easily be tracked? How does this amodally completed motion percept adapt itself to the animal's variable speed of locomotion? How does attention track the animal's moving trajectory?

The problem of cover pop-out is a special case of the more general problem of 3-D figure-ground separation, which implicates mechanisms of depthful boundary segmentation and surface representation within the visual form perception system. Computing a coherent direction-of-motion may also make use of form boundaries that group the animal's textures, as well as motion boundaries that use the form boundaries to propagate unambiguous motion direction signals from the animal's outer contours to the ambiguous scintillating mosaic within. This latter process realizes a type of "apparent motion", since it transforms the actual motion inputs received from the animal into a different motion percept. The interaction between texture and motion boundaries is, in addition, a type of form-motion interaction. Finally, the process whereby intermittent views, or flashes, of the animal are interpolated as it moves at variable speeds is a type of long-range apparent motion. This process may use form-motion interactions to group the target "flashes" that are intermittently seen by the observer before these groupings are interpolated by the motion system. This process also invokes figure-ground separation to amodally complete the apparent motion signals behind the occluding cover.

2 Neural Models of Visual Form and Motion

Neural network models of how these several perceptual processes work have been progressively developed during recent years (e.g. Chey *et al.*, 1995a, 1995b; Francis and Grossberg, 1996a, 1996b; Francis *et al.*, 1994; Grossberg, 1983, 1987a, 1991, 1994). Other successful models have focused upon processes of visual filtering (Chubb and Sperling, 1989; Wilson and Kim, 1994; Malik and Perona, 1990; Reichardt, 1961; Sutter *et al.*, 1989; van Santen and Sperling, 1984, 1985). The present models include filtering as well as grouping processes,

and have shown how the requirements of grouping suggest additional constraints that help to derive the filters that feed the grouping processes.

The present chapter reviews several of these model processes, including how long-range apparent motion interpolates intermittent flashes; how occluding objects and their boundaries are separated from, and pop-out in front of, occluded objects; how form-motion interactions can use contours generated within the form system to generate motion percepts within the motion system; how the motion signals of an intermittently viewed, occluded object may be amodally completed behind the occluding cover; and how spatial attention may track the amodally completed trajectory.

3 An Introduction to Apparent Motion

Exner (1875) provided the first empirical evidence that the visual perception of motion was a distinct perceptual quality, rather than being merely a series of spatially displaced static percepts over time. He did this by placing two sources of electrical sparks close together in space. When the sparks were flashed with an appropriate temporal interval between them, observers reported a compelling percept of continuous motion of a single flash from one location to another, even though neither flash actually moved. At shorter temporal intervals, flashes look simultaneous and stationary. At longer intervals, they look like successive stationary flashes, with no intervening motion percept. When the spatiotemporal parameters of the display are suboptimal, a "figureless" or "objectless" motion called *phi motion* is perceived, wherein a sense of motion without a clearly defined form is perceived. A smooth and continuous motion of a perceptually well-defined form is called *beta motion*, and typically occurs at a larger interstimulus interval, or ISI, between the offset of one flash and the onset of the next flash.

This classical demonstration of apparent motion was followed by a series of remarkable discoveries, particularly by gestalt psychologists, concerning the properties of motion perception. It was noticed that a decrease in ISI causes the speed of the interpolating motion to increase (Kolers, 1972). A motion percept can also smoothly interpolate flashes separated by different distances, speeding up if necessary to cross a longer distance at a fixed ISI. If a more intense flash follows a less intense flash, the perceived motion can travel backwards from the second flash to the first flash. This percept is called *delta motion* (Kolers, 1972; Korte, 1915). *Gamma motion* is the apparent expansion at the onset of a single flash, or its contraction at its offset (Bartley, 1941; Kolers, 1972). A similar expansion-then-contraction may be perceived when a region is suddenly darkened relative to its background, and then restored to the background luminance.

If a white spot on a gray background is followed by a nearby black spot on a gray background, then motion between the spots can occur while the percept changes from white to black at an intermediate position. Likewise, a red spot followed by a green spot on a white background leads to a continuous motion percept combined with a binary switch from red to green along the motion pathway (Kolers and von Grünau, 1975; Squires, 1931; van der Waals and Roelofs, 1930, 1931; Wertheimer, 1912/1961). These results show that the motion mechanism can combine visual stimuli corresponding to different colors, or even opposite directions-of-contrast. Complex tradeoffs between flash luminance, duration, distance, and ISI in the generation of motion percepts were also discovered. For example, the minimum

ISI for perceiving motion increases with increasing spatial separation of the inducing flashes. This property is sometimes called Korte's Third Law (Boring, 1950; Kolers, 1972; Korte, 1915). A similar threshold decrease with distance occurs in the minimum stimulus onset asynchrony, or SOA, which is the difference between the flash onset times. Interestingly, whereas the minimum ISI decreases with flash duration, the minimum SOA increases with flash duration.

These discoveries raised perplexing issues concerning the nature of the long-range brain interaction that generates a continuous motion percept between two stationary flashes. Why is this long-range interaction not perceived when only a single light is flashed? In particular, why are not outward waves of motion signals induced by a single flash? How does a motion signal get generated from the location of the first flash after the first flash terminates, and only after the second flash turns on? How does the motion signal adapt itself to the variable distances and ISIs of the second flash, by speeding up or slowing down accordingly? In particular, how can the motion signal adapt to the ISI between two flashes even though such adaptation can only begin after the first flash is over? I like to call this the ESP Problem.

The figural organization of motion stimuli can also influence motion percepts. The Ternus displays provide a classical example (Ternus, 1926/1950). In Frame 1 of a Ternus display, three white elements are placed in a horizontal row on a black background (or conversely). After an ISI, in Frame 2 all three elements are shifted to the right so that the two rightward elements in Frame 1 are in the same locations as the two leftward elements in Frame 2. Depending on the ISI, the observer perceives either of four percepts. At very short ISIs, all four elements appear simultaneous. At long ISIs, observers do not perceive motion at all. At ISIs slightly longer than those yielding simultaneity, the leftmost element in Frame 1 appears to jump to the rightmost element in Frame 2. This percept is called *element motion.* At somewhat longer ISIs, all three flashes seem to move together between Frame 1 and Frame 2. This is called *group motion.*

The percept of group motion might suggest that Ternus percepts are due to a cognitive process that groups the flashes into attended objects, and that motion perception occurs only after object perception. Such an explanation is not, however, made easily consistent with the percept of element motion. It has been argued, for example, that at short ISIs, the visual persistence of the brain's response to the two rightmost flashes of Frame 1 continues until the two leftmost flashes of Frame 2 occur (Braddick, 1980; Braddick and Adlard, 1978; Breitmeyer and Ritter, 1986; Pantle and Petersik, 1980). As a result, nothing changes at these two flash locations when Frame 2 occurs, so they do not seem to move. This type of explanation suggests that at least part of the apparent motion percept is determined at early processing stages. It does not, however, explain how we see element motion. In particular, why does not the element motion percept collide with the two stationary flash percepts? What kind of perceptual space can carry element motion across, or over, the stationary flashes?

Reverse-contrast Ternus motion also suggests that motion properties may be determined at early processing stages. In this paradigm, three white spots on a gray background in Frame 1 are followed by three black spots on a gray background in Frame 2 (see Figure 1). At the ISIs where element motion previously occurred, group motion now occurs (Pantle and Picciano, 1976). How does a change of contrast between Frame 1 and Frame 2 obliterate element motion? Does it do so by altering the effects of visual persistence on Frame 2?

Figure 1

A unified answer to all of these questions has recently been developed in a neural model of motion perception that clarifies the functional significance of many apparent motion percepts (Grossberg, 1991; Grossberg and Mingolla, 1993; Grossberg and Rudd, 1989, 1992). Perhaps the simplest such model is schematized in Figure 2. It is called a Motion Oriented Contrast-Sensitive Filter, or MOC Filter, and will be described in greater detail in Section 16. The entire model of motion segmentation consists of multiple copies of the MOC Filter, each corresponding to a different range of receptive field sizes, and each of which inputs to a grouping, or binding, network that is called the Motion Oriented Cooperative Competitive Loop, or MOCC Loop. Taken together, these MOC Filters and MOCC Loops are called the Motion Boundary Contour System, or Motion BCS.

The Motion BCS is designed to possess the minimal number of processing stages that are capable of tracking an object's direction-of-motion independent of whether the object's several parts are darker than or lighter than the background upon which they are moving. Grossberg and Rudd (1992) showed that each of the MOC Filter's processing stages is needed to explain data about beta motion, gamma motion, delta motion, Ternus motion, reverse motion, and related types of motion. The model's dynamics thereby illustrate how seemingly paradoxical apparent motion data may be explained as emergent properties of ecologically useful design constraints on the tracking of real moving objects.

Figure 2

4 Variable-Speed Apparent Motion and Continuous Tracking

I will first focus on one key process of the MOC Filter; namely, how large variations in ISI at fixed distance can be continuously interpolated, and how large variations in distance at fixed ISI are interpolated within a near-constant amount of time (Kolers, 1972). Given the ecological interpretation of apparent motion as a target tracking signal in a cluttered environment, this property means that the apparent motion signal can continuously interpolate the trajectory of a target that moves with variable speed behind variably spaced and sized occluders. The mechanism that achieves this is posited to exist between Levels 4 and 5 in the MOC Filter of Figure 2. It is a surprisingly simple mechanism and utilizes components that are generally familiar to psychologists: temporally decaying responses to stimuli input to a Gaussian filter whose responses are sharpened by lateral inhibition. Remarkably, in response to temporally successive inputs to the Gaussian filter at different locations, a traveling wave can be generated from the first input location to the second input location, and the peak of this wave, which is contrast-enhanced by lateral inhibition, generates a focal activation that speeds up or slows down with increases or decreases of distance or ISI as in the data.

5 G-Waves for Long-Range Apparent Motion

How are long-range apparent motion signals generated in such a model? Figure 3 schematizes how a flash at Level 1 (Figure 3a) leads to a focal activation at Level 5 (Figure 3c) after it activates the long-range Gaussian filter that joins Level 4 to Level 5 (Figure 3b). The broad

Gaussian activation of Level 5 is sharpened into a focal activation by lateral inhibition, or competition, among the Level 5 cells.

Figure 3

Figure 4 shows how this input activation looks in time. The input to Level 1 (Figure 4a) generates a slowly decaying temporal trace (Figure 4b) that has been called "visual inertia" by Anstis and Ramachandran (1987). When this trace is fed through the Gaussian filter, it generates a spatially distributed input to Level 5 that waxes and wanes through time, without spreading across space (Figure 4c). The maximum value of this input does not move. Hence a single flash does not cause a movement across space.

Figure 4

Suppose, however, that two locations both input through overlapping regions of the Gaussian filter, and that the activation in response to a flash at the first location is decaying while activation is growing in response to a flash at the second location (Figure 5). Under these circumstances, the *total* input to Level 5 from both flashes is the sum of a temporally waning Gaussian plus a temporally waxing Gaussian, as in Figure 6a. Under appropriate conditions, this sum represents a wave whose maximum travels continuously in time from the location of the first flash to the location of the second flash. After competition selects the location of this maximum, it represents a continuously moving peak of activation between the flashes (Figure 6b).

Figure 5

In summary, the time- and space-averaged responses to individual flashes do not change their positions of maximal activation through time. In this sense, nothing moves. When a series of properly timed and spaced flashes is presented, however, the sum of the temporally and spatially averaged responses that they generate can produce a continuously moving peak of activity between the positions of the stroboscopic flashes. This is an emergent property of network dynamics, rather than a property of any cell acting alone.

Figure 6

6 Motion Speed-Up and Multiscale Coherence

This Gaussian wave, called a G-wave, was discovered and mathematically analysed in Grossberg (1977). These results waited twelve years for publication in Grossberg and Rudd (1989) because it took that long to understand how a long-range Gaussian filter could be fit into a larger theory of motion perception, such as the Motion BCS, that also includes a role for such other processes as transient cells and short-range spatial interactions. A G-wave occurs whenever waxing and waning activation traces interact via a spatial Gaussian kernel under appropriate spatiotemporal conditions. One can therefore expect them to arise in a number of brain systems. This section analyses when a G-wave occurs and how it can speed up or slow down to fit changing ISI and flash distance conditions. It also analyses how such waves remain superimposed in space when they are processed by multiple Gaussian filters of variable size. The reader who is not interested in the mathematical mechanism of G-waves can skip directly to Section 7.

The basic mathematical framework for proving G-wave properties is straightforward. Let $\frac{dx}{dt}$ denote the rate of change through time of a cell's activity, or potential, x. Let activity x_i obey a leaky integrator equation in response to its input J_i at position i. In particular, let flashes occur at positions $i = 0$ and $i = L$. Then

$$\frac{dx_0}{dt} = -Ax_0 + J_0 \tag{1}$$

defines the activity x_0 and input J_0 at position 0, and

$$\frac{dx_L}{dt} = -Ax_L + J_L \tag{2}$$

does the same at position L, where $x_0(0) = x_L(0) = 0$. Equations (1) and (2) may be integrated to find:

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

and

$$x_L(t) = \int_0^t e^{-A(t-v)} J_L(v) dv. \tag{4}$$

Let the inputs J_0 and J_L switch on to the constant value J for a duration T at times 0 and $T + I$ respectively, as in:

$$J_0(t) = \begin{cases} J & \text{if } 0 \le t \le T \\ 0 & \text{if } T < t \end{cases}$$

and

$$J_L(t) = \begin{cases} J & \text{if } T + I \le t \le 2T + I \\ 0 & \text{if } 2T + I < t \end{cases} \tag{6}$$

where I is the ISI between the flashes. Then for times t between the onset and offset of J_L $(T + I \le t \le 2T + I)$,

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

and

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

By (7) and (8), $x_0(t)$ is decreasing while $x_L(t)$ is increasing.

Let $x_0(t)$ and $x_L(t)$ pool their outputs at Level 5 via the long-range Gaussian filter

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

as in Figure 6. For simplicity, replace index i by a continuum of cells at positions w in Level 5. Then the total input to position w of Level 5 is

$$T(w, t) = x_0(t) \exp[\frac{-w^2}{2K^2}] + x_L(t) \exp[\frac{-(w - L)^2}{2K^2}].$$

By (7) and (8),

$$T(w,t) = \frac{J}{A}\Big[(1-e^{-AT})e^{-A(t-T)}\exp[\frac{w^2}{2K^2}] + (1-e^{-A(t-T-I)})\exp[\frac{(w-L)^2}{2K^2}]\Big] \tag{11}$$

The first result shows under what combinations of parameters the maximum value of $T(w,t)$ moves continuously from position $w = 0$ towards position $w = L$ through time. It also characterizes the maximum flash separation that can generate a G-wave in response to a Gaussian with size parameter K in (9).

Theorem 1 (Apparent Motion)
The maximum of $T(w,t)$ moves continuously from position $w = 0$ to position $w = L$ if and only if

$$L < 2K. \tag{12}$$

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

$$\frac{\partial T(w,t)}{\partial w} = 0. \tag{13}$$

By (11), such locations obey the equation

$$\frac{e^{A(t-T)} - e^{AI}}{\quad} = \frac{w}{\quad}\exp\Big[\frac{L(L-2w)}{2K^2}\Big]. \tag{14}$$

The function

$$f(t) = \frac{e^{A(t-T)} - e^{AI}}{1 - e^{-AT}} \tag{15}$$

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 (14), namely function

$$g(w) = \frac{w}{L-w}\exp\Big[\frac{L(L-2w)}{2K^2}\Big], \tag{16}$$

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

$$w = g^{-1}\big(f(t)\big) \tag{17}$$

as an increasing function of w for all $0 \leq w \leq L$.

Function $g(w)$ is monotone increasing if $g'(w) > 0$. Differentiating (16) shows that this condition holds if and only if the function

$$h(w) \equiv (L-w)[1 - \frac{Lw}{K^2}] + w \tag{18}$$

satisfies

$$h(w) > 0. \tag{19}$$

In order for (19) 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 $w = \frac{L}{2}$, and equals

$$h(\frac{L}{2}) = L(1 - \frac{L^2}{4K^2}).$$

The number $h(\frac{L}{2})$ is positive if (12) holds.

The next result proves that the apparent motion signal reaches the position $w = \frac{L}{2}$ midway between positions $w = 0$ and $w = L$ at a time $t_{\frac{1}{2}}$ that is independent of L and K. Independence of L illustrates how the wave speeds up to travel over larger interflash distances. Independence of K means that wave speed does not depend upon the width K of the Gaussian filter, just so long as the filter is wide enough to support a wave. This means that the speed-up property enables multiple spatial scales in the motion perception system to generate G-waves that are all spatially coincident (Figure 7). Because of this property, a coherent motion percept may be synthesized from data from all the spatial scales of long-range motion filters that are activated by the stimulus.

Theorem 2 (Equal Half-Time Property)
The time at which the motion signal reaches position $w = \frac{L}{2}$ is

$$t_{\frac{1}{2}} = T + \frac{1}{A} ln\left[e^{AI} + (1 - e^{-AT})\right]. \tag{21}$$

Proof: By (17), we need to compute $t = f^{-1}(g(w))$ when $w = \frac{L}{2}$, namely

$$t_{\frac{1}{2}} = f^{-1}\left(g(\frac{L}{2})\right).$$

By (16),

$$g(\frac{L}{2}) = 1.$$

Equation (21) follows immediately from (23), (14), and (16)

Figure 7

7 Spatial Attention Shifts and Target Tracking by the Where Cortical Stream

In Grossberg (1991), I suggested that long-range apparent motion, including the property of motion speed-up, could be used to help track rapidly moving objects that may be perceived only intermittently. It was suggested that a G-wave could give rise to certain spatial shifts in attention, such as those reported by Ericksen and Murphy (1987), Kwak, Dagenbach, and Egeth (1991), LaBerge and Brown (1989), and Remington and Pierce (1984). Thus if a targeted predator or prey rapidly moved across a scene, perhaps darting behind protective cover, then a G-wave could interpolate these temporally discrete views with a continuous motion signal that adapts its speed to the varying speed of the target. Such a continuous motion signal could activate the spatial attention system, where it could be used to predict

the location and speed of the target, and to command motor responses accordingly. The results of Kwak, Dagenbach, and Egeth (1991) and of Remington and Pierce (1984) are of particular interest, since they reported that spatial attention, just like long-range apparent motion, could speed up to cover variable distances in equal time.

The hypothesis that a G-wave may give rise to a spatial attention shift is consistent with the fact that the motion perception cortical processing stream, labeled magnocellular in Figure 8, is part of a larger Where processing stream that includes cortical region MT as well as parietal cortex. The MOC Filter is hypothesized to model aspects of the V1→ MT processing stream and thus feeds into parietal cortex. The Where processing stream computes the locations of targets with respect to an observer and helps to direct attention and action towards them (Goodale and Milner, 1992; Mishkin, Ungerlieder, and Macko, 1983; Ungerlieder and Mishkin, 1982). Data consistent with the prediction that a G-wave interpolative process occurs in the Where stream have recently been recorded from the posterior parietal cortex of monkeys (Albright, 1995; Assad and Maunsell, 1995). In contrast, the form perception processing stream, labeled parvocellular in Figure 8, is part of a larger What processing stream that includes region V4 as well as inferotemporal cortex. The What processing stream generates representations of object form and color and uses them to recognize the objects based upon prior learning (Desimone, 1991, 1992; Desimone, Schein, Moran, and Ungerleider, 1985; Desimone and Ungerleider, 1989; Gochin, 1990; Gochin, Miller, Gross, and Gerstein, 1991; Harries and Perrett, 1991; Miller, Li, and Desimone, 1991; Mishkin, 1982; Mishkin and Appenzeller, 1987; Perrett, Mistlin, and Chitty, 1987; Riches, Wilson, and Brown, 1991; Schwartz, Desimone, Albright, and Gross, 1983; Spitzer, Desimone, and Moran, 1988).

Figure 8

In evaluating the hypothesis that long-range apparent motion inputs to spatial attention mechanisms, it needs to be kept in mind that a spatially continuous motion signal is generated only under certain spatiotemporal conditions, that the speed of the motion signal is nonuniform in time (see Grossberg and Rudd, 1992), and that spatially discrete jumps in activation may occur in cases where continuous motion is not observed; for example, if $L > 2K$ in (12). These properties may help to clarify some of the variability in data about how quickly attention shifts under different conditions and whether it does so continuously or discretely.

8 Amodal Completion of Motion Signals Behind Occluding Forms

The above discussion clarifies how a long-range apparent motion G-wave can interpolate continuous motion signals between intermittently seen targets. It remains to explain how occluding objects can pop-out in front of a moving target and how a G-wave can be amodally completed behind the occluders in response to intermittent views of the target.

Our model of how this happens includes the following features. In addition to a motion BCS model of how V1 → MT interactions within the Where cortical processing stream generate boundary segmentations of motion signals, the model also analyses aspects of visual processing in the What cortical processing stream. It includes a static BCS model of how the V1 → V2 → V4 Interblob stream generates 3-D boundary segmentations of

object form (Francis, Grossberg, and Mingolla, 1994; Grossberg, 1987a, 1994; Grossberg and McLoughlin, 1995; Grossberg and Mingolla, 1985a, 1985b, 1987) and a parallel Feature Contour System (FCS) of how the V1 → V2 → V4 Blob stream generates 3-D surface representations of object brightness and color (Arrington, 1994; Cohen and Grossberg, 1984; Grossberg, 1994; Grossberg and Mingolla, 1985a; Grossberg and Todorović, 1988; Grossberg and Wyse, 1991; Pessoa, Mingolla, and Neumann, 1995).

All boundaries that are computed within the static BCS and motion BCS are perceptually invisible. This is true because these boundary systems pool signals from detectors that are sensitive to opposite contrast polarities. This pooling process completes boundaries around objects whose background texture and shading break the total object boundary into fragments with opposite contrast polarity, as in Figure 10. Such pooling occurs, for example, at the complex cell stage of the Static BCS.

Figure 9

A thorough analysis of how occluding objects pop-out in front of occluded objects would take us too far afield. For such an analysis, see Grossberg (1994). Here a few key points are made in order to clarify the main factors that seem to be involved. The static BCS includes multiple BCS copies; see Figure 10. Each BCS copy can generate a 3-D boundary segmentation of scenic data that lies within a prescribed depth range from the observer. Other things being equal, the BCS boundaries of occluding objects generate inhibitory signals that detach boundaries of occluded objects from them. In other words, at the T-junctions where occluded objects intersect occluding objects, the stems of the T's are detached from their tops. As a result of this operation, the boundaries that surround occluding objects may be completed but those around occluded objects are broken.

These BCS copies interact reciprocally with FCS copies. The BCS-to-FCS signals (see pathways 6 in Figure 9) capture those monocular FCS color and brightness signals (from pathways 5 in Figure 9) that are consistent with the copy's BCS boundary. These FCS signals diffusively fill-in a surface representation within the corresponding FCS copy. This surface is attributed to the same depth range as the BCS copy itself. Because boundaries around occluded objects are broken, only the occluding objects can be successfully filled-in by this operation. Further processing alters this situation as follows.

Reciprocal FCS-to-BCS signals (pathways 7 in Figure 9) are derived from the contours of successfully filled-in regions, notably the surfaces of occluding objects. These feedback signals excite, and thereby strengthen, boundaries at the BCS copy that activates their own surface. They inhibit boundaries at their locations within BCS copies that represent larger depths. Before this inhibitory feedback acts, the boundaries of occluding objects could be represented at these larger depths too, due to the action of a size-disparity correlation that gives rise to the depth-selective BCS cells. After the inhibitory feedback acts, these redundant boundaries of occluding objects are inhibited and, along with them, their inhibition of the occluded boundaries is removed.

Two consequences result from FCS-to-BCS feedback: First, the boundaries that are shared by occluding and occluded objects become selectively attached, or bound, to the surface representation of the occluding object. "Boundary ownership" by occluding objects is hereby achieved. Second, as soon as occluder boundaries are inhibited within the BCS copies that represent more distant depths, then the occluded boundaries can be completed

behind the occluders, whether as static illusory contours or moving G-waves. The illusory contours that amodally complete the occluded parts of static forms are hypothesized to be computed within the Interblob V1 → V2 → V4 stream. The G-waves that interpolate intermittent views of a moving target are hypothesized to use a combination of V1 → MT motion interactions and V1 →V2 → MT form-motion interactions.

Figure 10

9 Form-Motion Interactions

The remainder of the chapter reviews recent modeling results concerning how form-motion interactions occur; in particular, how computations of visual form within the Static BCS generate inputs to the Motion BCS where they influence percepts of motion, such as the motion of a moving target behind intermittent occluding cover. More generally, these results concern the large issue of how the visual system integrates visual form and motion information to generate a coherent percept of moving forms. As noted above, many percepts of form are emergent properties of images and scenes, much as illusory contours help to group textured scenes into detectable objects (Beck, Prazdny, and Rosenfeld, 1983). It is also know known that many motion percepts depend on detection of form. Chubb and Sperling (1991) described motion percepts that are undetectable by some types of motion analysis and noted how detection of moving form might explain these percepts. Cavanagh and Mather (1989) argued that some properties of apparent motion require detection of moving forms. Wilson, Ferrera, and Yo (1993) described how integration of luminance-based and form-based motion could help to explain percepts of moving plaids.

Neurophysiological evidence of form and motion integration includes neurophysiological studies suggesting that properties of motion (Albright, Desimone, and Gross, 1984; Maunsell and van Essen, 1983; Mikami, Newsome, and Wurtz, 1986a) and apparent motion (Mikami, Newsome, and Wurtz, 1986b; Newsome, Mikami, and Wurtz, 1986) are represented in the processing stream of visual cortex that includes area MT. Von der Heydt, Peterhans, and Baumgartner (1984) and Peterhans and von der Heydt (1989) have reported evidence that the interblob cortical stream generates emergent properties of visual form, notably in area V2. Given these parallel cortical processing streams for form and motion processing, at what processing stages do form and motion processing interact? One possible link is between cortical areas V2 and MT (DeYoe and van Essen, 1988).

Grossberg (1991) outlined a model that suggested how a V1 → V2 → MT link between form and motion processing mechanisms could play two related functional roles. First, such a connection would allow the motion detecting system to respond to diverse perceptual groupings, such as illusory contours and segmentations of textures. Second, it was explained how the model form processing mechanisms could make finer disparity matches than could its motion processing mechanisms, in a manner compatible with neurophysiological evidence about depth sensitivity from cells in V2 (von der Heydt, Hänny, and Dürsteler, 1981) and MT (Logothetis, Schiller, Charles and Hurlbert, 1990; Schiller, Logothetis and Charles, 1990a, 1990b). Thus, the form-to-motion pathway could help the motion system create motion signals at the computed depths of the perceptual groupings, thereby integrating form and motion data into consistent percepts of moving forms in depth, including movements of

grouped visual targets behind intermittent occluding cover. This analysis suggested at what processing stages outputs from the form system should input to the motion system. In particular, the form inputs arrive at the motion system prior to the Gaussian filter that generates G-waves.

Francis and Grossberg (1996a) developed the Grossberg (1991) proposal by modeling how to link the static BCS models of emergent boundary segmentation and the Motion BCS model of motion perception to explain and simulate data about form-motion interactions. The model has thus far been used to analyse form and motion percepts that mechanistically links together several types of data that heretofore have been treated separately. In particular, it links together data about the persistence of static images with data about the quality of apparent motion. The key idea is to relate the time taken to generate and reset a persistent boundary segmentation in the form cortical stream (through V2) with threshold properties of apparent motion in the motion cortical stream (through MT). The model hereby uses the *dynamic* characteristics of the form processing system to explain data about motion perception.

As a first step in analyzing how an emergent form in V2 could influence a motion percept in MT, Francis and Grossberg (1996a) simulated visual displays that generate 2-D percepts of apparent motion of moving forms. These results can be viewed as a probe of how a moving form within a particular depth plane influences the corresponding motion percept. The analysis simulated three challenging sets of data, using a fixed set of model parameters, to test the model's competence:

(I) Illusory contours move in apparent motion and do not obey the inverse relationship between upper interstimulus interval (ISI) thresholds and stimulus duration that is characteristic of luminance-based contours (Mather, 1988; Ramachandran, 1985; von Grünau, 1979).

(II) Apparent motion can occur between one stimulus defined by illusory contours and a second stimulus defined by luminance contrast (Cavanagh, Arguin and von Grünau, 1989; von Grünau, 1979).

(III) Korte's Laws: For luminance-based stimuli, both upper and lower ISI thresholds are inversely related to flash duration. The range of ISIs capable of producing apparent motion narrows as the spatial separation between the flashes increases (Kolers, 1972; Korte, 1915; Neuhaus, 1930).

These data and how the model explains each of them are reviewed below.

10 Apparent motion of illusory contours

Several authors have shown that illusory contours can move in apparent motion (Mather, 1988; Ramachandran, 1985; von Grünau, 1979). From an ecological point of view, such a property would help the visual system to continuously track the motion of a textured target that is grouped by illusory contours as it intermittently emerges from occluding cover. A simple version of this competence was explored in an experiment by Ramachandran (1985), wherein subjects saw an illusory Kanizsa square on the left side and a jumbled set of lines on the right side. During the second time period, the pac man circles that induced the Kanizsa square filled up and the illusory Kanizsa square disappeared. At the same time, lines on the right within a region defined by an illusory square disappeared. Subjects reported seeing

motion of the illusory square from the left to the right. Features in the two images cannot be matched, but the illusory Kanizsa squares which these features induce can be matched. Thus the emergent illusory form activated the motion system. Von Grünau (1979) reported similar results.

Figure 11

Mather (1988) investigated the temporal properties of illusory contour apparent motion. Figure 11a (from Mather) shows contour plots of reports of seeing apparent motion between two illusory Kanizsa squares as a function of the inducing stimuli duration for two subjects. Note the inverted-U shape of the top curves that divide regions of 55% perceived motion. This curve represents the upper ISI threshold for perceiving apparent motion of the illusory contour. Figure 11c replots the curves on non-logarithmic axes. The shape of these curves is unlike the data derived from apparent motion of luminance-based contours. For example, as described below, the upper ISI threshold values for illusory contour motion exceed those of luminance-based contour motion. The inverted-U shape of threshold ISIs as a function of stimulus duration also differs from the inverse relation of threshold ISI to stimulus duration that is found for luminance-based contours.

11 Persistence of illusory contours

The model explains these properties of illusory contour apparent motion as a consequence of form-motion interactions. The illusory contours are generated in the form perception system (Static BCS) and input to the motion perception system (Motion BCS). Key properties of Mather's motion perception data are explained using properties of the stationary illusory contours that are computed in the Static BCS. In particular, data on illusory contour persistence of Meyer and Ming (1988) are described by curves that are remarkably similar to the upper ISI thresholds found by Mather (1988). Francis, Grossberg, and Mingolla (1994) simulated the persistence of illusory and luminance-based contours using the Static BCS model. As in the data of Meyer and Ming (1988), the Static BCS representation of an illusory contour lasts longer than that of a real contour and exhibits an inverted-U relationship between persistence and stimulus duration. These properties were traced in Francis *et al.* (1994) to an analysis of why the illusory contour boundary takes longer to form than a luminance-based boundary (increasing portion of the inverted-U curve) and has fewer reset signals to shut it off (decreasing portion of the inverted-U curve). This analysis is reviewed in Section 15 for completeness.

As noted above, the model suggests how the Static BCS interacts with the Motion BCS to model the cortical V1 $\rightarrow$ V2 $\rightarrow$ MT interaction. In particular, the persistence of illusory contour inputs from the form model (Static BCS) to the motion model (Motion BCS) determines the upper ISI threshold of apparent motion of the contour. In this way, the dynamic characteristics of form processing are used below to explain the data in Figure 11a. Figure 11b summarizes ISI thresholds for computer simulations of illusory contour motion in the model that qualitatively match the properties of the curves found by Mather (1988). In particular, Figure 11c plots the ISI thresholds from the subjects in Mather's study and the model. The results are similar in magnitude and qualitative shape. This simulation used

the same parameters for the Static BCS as that were used in Francis *et al.* (1994) to simulate data from Meyer and Ming (1988), who directly measured the inverted-U relationship between persistence and stimulus duration.

12 Interattribute Motion

Von Grünau (1979) observed that subjects can sometimes see apparent motion between an illusory contour and a contour defined by luminance edges. Cavanagh, Arguin, and von Grünau (1989) generalized this result by showing that subjects reported seeing motion between stimuli defined by any combination of attributes, including luminance, color, texture, relative motion, or stereopsis. They also noted that motion between stimuli of different attributes is weaker than motion between stimuli of the same attribute. Taken together, these studies suggest that interattribute motion is thus a result of form and motion integration.

The Static BCS is capable of responding to multiple types of form-supporting cues, including luminance, color, texture, shading, and stereo cues (Cruthirds, Grossberg, and Mingolla, 1993; Graham, Beck, and Sutter, 1992; Grossberg, 1994; Grossberg and Mingolla, 1985a, 1985b, 1987; Grossberg and McLoughlin, 1995; Grossberg, Mingolla, and Williamson, 1995; Sutter, Beck, and Graham, 1989). The Motion BCS is capable of responding to a wide range of apparent motion, first-order motion, and second-order motion cues (Chey, Grossberg, and Mingolla, 1995a, 1995b; Grossberg and Rudd, 1989, 1992; Nogueira, Grossberg, and Mingolla, 1993). Thus many properties of interattribute motion could, in principle, be explained by interactions between the Static BCS and the Motion BCS. From an ecological perspective, such interattribute apparent motion could facilitate target tracking under viewing conditions that vary greatly as the target moves from one viewing context to another.

13 Korte's Laws

Figure 12a shows the upper and lower ISI threshold values for apparent motion of luminance-based stimuli (Kolers, 1972; after Neuhaus, 1930). This figure shows that as stimulus duration increases from 10 to 45 to 90 milliseconds, each upper and lower ISI threshold curve decreases at every spatial separation. Moreover, as the distance between the two stimuli increases, the range of ISIs that produce apparent motion narrows, with the upper ISI decreasing and the lower ISI increasing for every stimulus duration. These properties are often collectively referred to as Korte's laws (Korte, 1915).

Figure 11

Grossberg and Rudd (1992) explained the characteristics of the lower ISI thresholds and the role of spatial separation using the MOC Filter of Figure 2. In the Grossberg and Rudd simulations, the MOC Filter was sensitive only to changes in luminance. In the form-motion model of Francis and Grossberg (1996a), the boundary segmentation outputs of the Static BCS input to the MOC Filter. The MOC Filter is thus rendered sensitive to changes in form as well as to changes in luminance. The fast responses of transient cells to luminance flashes allow the MOC Filter to respond quickly to moving stimuli. Form processing by the BCS may be slower than motion processing, so that persisting form inputs from the Static

BCS to the MOC Filter often outlast the effects of purely luminance based inputs. These form-motion interactions are shown below to be sufficient to explain the properties of the upper ISI thresholds in Korte's Laws.

In particular, Francis *et al.* (1994) showed that increasing the duration of a stationary form input decreases the persistence of the boundary representation in the Static BCS, much as experiments on visual persistence (e.g., Bowen, Pola and Matin, 1974) report an inverse relationship between persistence and stimulus duration. As in the case of illusory contours above, the persistence of form signals determines the upper ISI threshold of apparent motion. The remaining properties of Korte's Laws — namely, the lower ISI thresholds and influences of spatial separation — are explained below in terms of MOC Filter properties. Figure 12b summarizes the computer simulations of how the form-motion model simulates Korte's Laws. The model reproduces all the qualitative properties of the classical Neuhaus (1930) data, as well as the most important quantitative property: The surprisingly large 350 msec gap between the smallest lower ISI threshold and the largest upper ISI threshold.

14 The Static Boundary Contour System

In the remaining sections, the Static BCS and MOC filter are more completely defined and used to summarize simulations of the above data. Grossberg (1984) and Cohen and Grossberg (1984) introduced the Static BCS model. Grossberg (1987a, 1987b, 1994) and Grossberg and Mingolla (1985a, 1985b, 1987) developed the model to simulate how the visual system detects, completes, and regularizes boundary segmentations in response to regions of different luminance, color, texture, shading, or stereo signals. The Static BCS computations for single-scale monocular processing will be reviewed herein. They consist of a series of filtering and grouping stages; see Figure 13. The first stage, schematized as an unoriented annulus in Figure 13, models relevant features of the shunting on-center off-surround interactions at the retinal and LGN levels. These cells compensate for variable illumination and enhance regions of local contrast in the image. Both on-center off-surround ON cells and off-center on-surround OFF cells exist *in vivo*, but are not needed herein. Their complementary responses to images are modeled elsewhere (Gove, Grossberg, and Mingolla, 1995; Grossberg, Mingolla, and Williamson, 1995; Grossberg and Wyse, 1991; Pessoa, Mingolla, and Neumann, 1994).

These model LGN cells input to pairs of like-oriented simple cells that are sensitive to opposite contrast polarity, or direction-of-contrast. Simple cell activities are half-wave rectified to generate output signals. Pairs of simple cells sensitive to opposite contrast polarity pool their half-wave rectified output signals at like-oriented complex cells. By pooling outputs from opposite polarity simple cells, complex cell outputs are rendered insensitive to direction-of-contrast, as are all subsequent cell types in the model. That is how "all boundaries are invisible" obtains in the model. Complex cells activate hypercomplex cells through an on-center off-surround network (first competitive stage) whose off-surround carries out an endstopping operation. In this way, complex cells excite hypercomplex cells of the same orientation and position, while inhibiting hypercomplex cells of the same orientation at nearby positions. One role of this spatial competition is to spatially sharpen the neural responses to oriented luminance edges. Another role is to initiate the process, called *end cutting*, whereby boundaries are formed that abut a line end at orientations perpendicular or oblique to the

orientation of the line itself (Grossberg, 1987a; Grossberg and Mingolla, 1985b).

The pooling of oppositely polarized, half-wave rectified signals at complex cells has the same net effect as an oriented full-wave rectified filter. The sequence of filtering (by simple cells) followed by full-wave rectification and subsequent lower-frequency oriented filtering (by hypercomplex cells) has been used in a number of models of texture segregation (Bergen and Landy, 1991; Malik and Perona, 1990; Sutter, Beck and Graham, 1989) since its introduction in the BCS by Grossberg and Mingolla (1985b).

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

Output from the higher-order hypercomplex cells feed into cooperative bipole cells that initiate long-range boundary grouping and completion. Bipole cells fire if both of their receptive fields are sufficiently activated by inputs from hypercomplex cells whose orientations are approximately colinear with the axes of the bipole receptive fields. Bipole cells hereby realize a type of long-range cooperation among the outputs of active hypercomplex cells. For example, a horizontal bipole cell, as in Figure 13, is excited by activation of horizontal hypercomplex cells that input to its horizontally oriented receptive fields. A horizontal bipole cell is also inhibited by activation of vertical hypercomplex cells. This inhibition prevents colinear horizontal grouping from penetrating a region filled with vertical or oblique orientations, a property called spatial impenetrability (Grossberg, 1987a; Grossberg and Mingolla, 1987).

Bipole cells were predicted to exist in Cohen and Grossberg (1984) and Grossberg (1984) shortly before cortical cells in area V2 with similar properties were reported by von der Heydt, Peterhans, and Baumgartner (1984). Grossberg and Mingolla (1985a, 1985b) have used bipole cell properties to simulate and explain data about illusory contour formation, neon color spreading, and texture segregation. Grossberg (1994) used them to explain data about depth perception and figure-ground separation. These same properties play a key role in the explanations herein of apparent motion of illusory contours and interattribute apparent motion.

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

it from the MOCC Loop of the Motion BCS (see Section 3). Cruthirds *et al.* (1993) have achieved better fits to texture grouping data using both the BCS filter and grouping than the filter alone.

15 Temporal dynamics of Boundary Persistence and Reset

The positive feedback within the hypercomplex-bipole feedback loop creates hysteresis that could cause uncontrolled boundary persistence and image smearing in response to image motion (Burr, 1980). Francis *et al.* (1994) showed that luminance edge or illusory contour inducer inputs to the BCS can trigger feedback interactions that, if left unchecked, could last for hundreds of simulated milliseconds. This happens because hypercomplex and bipole cell activities at a particular position and orientation decay away only when bipole cell output centered at the same position and orientation weakens. Since bipole cell activation depends on inputs to both receptive fields, bipole activation near the ends of contours weakens first after inputs shut off. As these bipole cells lose activation, so do all other cells of the same orientation and position. This decay causes more bipole cell activities to decay, which continues the process. The net effect of these spatial and temporal interactions is that boundary activities erode from contour ends to the contour middle. This erosion is observable in the simulations in Figure 19 below.

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

The Static BCS corrects this problem by using its resonant feedback to maintain segmentations of unmoving scenic objects, while it actively resets segmentations corresponding to rapidly changing scenic objects. Image smearing is hereby controlled in a form-sensitive way. The BCS uses the same circuit to create resonant boundaries and to reset them. Two types of mechanism accomplish this. The first mechanism uses the orientational competition that converts model hypercomplex cells into higher-order hypercomplex cells. Consider how this competition works between pairs of mutually perpendicular cells. Pairs of mutually perpendicular complex, hypercomplex, and higher-order hypercomplex cells, designated in gray within Figure 13, define a type of opponent processing circuit that Grossberg (1972) has called a gated dipole. The gates in the dipole are habituative transmitters that multiply signals in the pathways from complex to hypercomplex cells (square synapses in Figure 13). Such a gated dipole can rapidly inhibit a bipole cell when its inputs shut off or are removed due to image motion.

Figure 13

For example, suppose that a horizontal edge turns on horizontally oriented complex, hypercomplex, and bipole cells, thereby generating a horizontal boundary segmentation. Offset of the horizontal edge can cause an antagonistic rebound of activity in the corresponding gated dipoles, leading to activation of vertically oriented hypercomplex cells, which inhibit

horizontal bipole cells by the same mechanisms that ensures spatial impenetrability.

The rebound is generated as follows. When the horizontal input is on, horizontal transmitter gates habituate. The net result is an overshoot of input to horizontal bipole cells, followed by a steady input level after habituation takes place. When the input subsequently shuts off, the altered balance of transmitter between the horizontal and vertical channels now favors the vertical channel and permits a rebound of vertical cell activity in response to an internally generated tonic input that equally activates both channels. When this happens, an inhibitory input to the bipole cell occurs. The rebound is transient because transmitters in both channels then gradually equilibrate to equal levels.

In summary, rebound-driven inhibition of the bipole cells selectively limits persistence and smearing at those locations where the image is changing. Some further implications of BCS reset by a transient antagonistic rebound are worth noting here due to its conceptual importance in the model. Psychophysical data that support orientationally opponent aftereffects include the MacKay illusion (MacKay, 1957; Taylor, 1958). Several components of the gated dipole circuit have known cellular correlates in visual cortex, including tonically active cells (such as the cells that feed the habituative transmitters) and polarization from opposite orientations (Creutzfeldt, Kuhnt, and Benevento, 1974; Levitt, Kiper, and Movshon, 1994). In more complex versions of the BCS, both ON cells that are turned on by an input and OFF cells that are turned off by an input are modeled (Gove, Grossberg, and Mingolla, 1995; Grossberg, 1991). In such a network, offset of a horizontal ON cell mediates transient activation of a horizontal OFF cell, as well as the type of onset of a vertical ON cell that is here simulated. Inhibition of horizontal bipole cells may thus be mediated by horizontal OFF cells, rather than by vertical ON cells, as was here assumed for simplicity. In this more general model, any influence of vertical cells on horizontal bipoles could be mediated by horizontal ON or OFF cells. Further neurophysiological experiments are needed to test the cellular substrate of this predicted boundary reset mechanism and, by implication, of orientationally opponent aftereffects.

The rebound-driven reset mechanism shuts off boundary segmentations at locations that lose input support due to image offset or motion. The second reset mechanism helps to prevent image smearing across space. It uses the spatial endstopping competition that occurs among like-oriented hypercomplex cells at the first competitive stage (see Figure 13). Castet (1994) has reported experiments that are consistent with this model prediction. Francis *et al.* (1994) shown that these two mechanisms of the BCS model are sufficient to explain the key parametric properties of visual persistence experiments. Francis (1994) has used the model to simulate the Castet (1994) data.

To explain properties of apparent motion, two characteristics of visual persistence are relevant:

(I) Visual persistence experiments demonstrate that persistence duration decreases in response to image edges as stimulus duration increases (Bowen *et al.*, 1974). BCS simulations of boundary signal persistence agree with these findings (Francis *et al.*, 1994). The strength of the inhibitory rebound in the gated dipole mechanism explains the model's results. As stimulus duration increases, the gate habituates more and the strength of the subsequent rebound increases with the amount of habituation. A longer stimulus hereby generates stronger inhibition at stimulus offset, which speeds up the erosion of boundary signals, and reduces measured persistence.

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

The persistence of BCS output signals will be shown below to establish the upper ISI thresholds for apparent motion. This link will use the above properties of visual persistence of non-moving stimuli to explain data about apparent motion.

16 The MOC Filter

Grossberg and Rudd (1989, 1992) and Grossberg and Mingolla (1993) developed the Motion BCS and its front end, the MOC Filter, to explain psychophysical and neural data about motion perception. The Motion BCS is used to generate boundary segmentations of moving objects. Whereas the boundary segmentations of the Static BCS compute boundaries based on static image *orientations*, the Motion BCS computes boundaries based on moving image *directions*. Both BCS systems generate segmentations whose outputs pool signals from opposite contrast polarities, so that their boundaries can interpolate object boundaries that cross textured and shaded image regions where contrast polarity reverses.

The MOC Filter may be conceptually described in several ways. It is a minimal filter that produces output signals that are insensitive to direction-of-contrast but sensitive to direction-of-motion. It transforms information from multiple orientations and unoriented input signals into directionally-selective output signals. To accomplish these goals the MOC Filter uses a hierarchy of short-range and long-range spatial interactions that help to explain data about short-range and long-range motion within a single system. The qualitative properties of the five MOC Filter processing levels of Figure 2 are summarized below.

Figure 14

Level 1: Preprocess Input Pattern

The image is preprocessed before activating the filter. For example, as in the Static BCS, it is passed through a shunting on-center, off-surround net to compensate for variable illumination, or to "discount the illuminant," and to thereby process ratio contrasts in the image (Grossberg and Todorović, 1988).

Level 2: Sustained Cell Short-Range Filter

Four operations occur here, as illustrated in Figure 14.

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

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

Short-range spatial filter. A spatially aligned array of simple cells with like orientation and direction-of-contrast pool their output signals to activate the next cell level. As shown in Figure 14a, the target cells are pooled in a movement direction that is not necessarily perpendicular to the simple cell's preferred orientation. This spatial pooling plays the role of the short-range motion limit D_{max} (Braddick, 1974). The breadth of spatial pooling scales with the size of the simple cell receptive fields (Figure 14). Correspondingly, D_{max} depends on the spatial frequency content of the image (Anderson and Burr, 1987; Burr, Ross, and Morrone, 1986; Nakayama and Silverman, 1984, 1985; Petersik, Pufahl, and Krasnoff, 1983) and is not a universal constant.

Time average. The target cell time averages the inputs that it receives from its short-range spatial filter. This operation has properties akin to the "visual inertia" during apparent motion that was reported by Anstis and Ramachandran (1987); see Figure 14c.

Level 3: Transient Cell Filter:

In parallel with the sustained cell filter, a transient cell filter reacts to input increments (on-cells) or decrements (off-cells) with positive outputs (Figure 15). These filters use five operations:

Figure 15

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

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

Change detector. The on-cells are activated when the time average increases (Figure 15a). The off-cells are activated when the time average decreases (Figure 15b).

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

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

Level 4: Sustained-Transient Gating Yields Direction-of-Motion and Direction-of-Contrast Sensitivity

Maximal activation of a Level 2 sustained cell filter is caused by image contrasts moving in either of two directions that differ by 180°. Multiplicative gating of each Level 2 sustained cell output with a Level 3 transient cell on-cell or off-cell removes this ambiguity (Figure 16). For example, consider a sustained cell output from vertically oriented dark-light simple cell receptive fields that are joined together in the horizontal direction by the short-range spatial filter (Figure 14a). Such a sustained cell output is maximized by a dark-light image contrast moving to the right or to the left. Multiplying this Level 2 output with a Level 3 transient on-cell output generates a Level 4 cell that responds maximally to motion to the left. Multiplying

it with a Level 3 off-cell output generates a Level 4 cell that responds maximally to motion to the right. Multiplying a sustained cell with a transient cell is the main operation of the Marr and Ullman (1981) motion detector. Despite this similarity, Grossberg and Rudd (1989) described six basic differences between the MOC Filter and the Marr-Ullman model.

Figure 16

Level 5: Long-Range Spatial Filter and Competition

Outputs from Level 4 cells that are sensitive to the same direction-of-motion but opposite directions-of-contrast activate individual Level 5 cells by a long-range spatial filter that has a Gaussian profile across space (Figure 17). This is the Gaussian filter that generates G-waves in Sections 5 and 6. The functional role of this filter is to group together Level 4 cell outputs that are derived from Level 3 short-range filters with the same directional preference but different simple cell orientations or polarities. The long-range filter hereby provides the extra degree of freedom that enables Level 5 cell to function as *direction* cells, rather than as *orientation* cells, and to be able to track object boundaries that cross regions of opposite contrast polarity. Cells in cortical area MT can also respond to a range of orientations that are not perpendicular to their preferred direction-of-motion (Albright, 1984; Albright, Desimone and Gross, 1984; Maunsell and van Essen, 1983; Newsome, Gizzi, and Movshon, 1983).

Figure 17

The long-range spatial filter broadcasts each Level 4 signal over a wide spatial range in Level 5. Competitive, or lateral inhibitory, interactions within Level 5 contrast-enhance this input pattern to generate spatially sharp Level 5 responses. A winner-take-all competitive network (Grossberg, 1973, 1982) can transform even a very broad input pattern into a focal activation at the position that receives the maximal input. More generally, this competitive process may only partially contrast-enhance its input pattern to generate a motion signal whose breadth across space increases with the breadth of its inducing pattern.

These model interactions can generate a continuous G-wave in response to the discrete flashes of an apparent motion display. Grossberg and Rudd (1989, 1992) correlated properties of this traveling peak of activity with properties of apparent motion data, including beta motion, gamma motion, delta motion, split motion, Ternus motion, and reverse-contrast Ternus motion. Grossberg and Rudd (1992) also suggested that the inverse relationship between lower ISI thresholds and stimulus duration, as in Figure 12, is due to a lag in the response time of the Level 3 transient cells. MOC Filter model transient cells respond more quickly to the offset of a long duration flash than to the offset of a short duration flash. Breitmeyer (1984) reviewed studies of transient cells that are consistent with this property. Keeping the ISI constant and increasing the spatial separation of apparent motion stimuli produces weaker G-waves at Level 5. This property clarifies why the lower ISI threshold increases and the upper ISI threshold decreases as a function of spatial separation.

17 Integration of form and motion processing

The MOC filter pathways model luminance-based V1 $\rightarrow$ MT interactions. As a result, the MOC Filter, by itself, cannot generate motion between stimuli defined by other stimulus

characteristics, such as the emergent boundary segmentations that help to define many visual form percepts in response to textured and shaded images. Apparent motion of illusory contours illustrates this competence because the illusory contours are emergent boundaries and their apparent motion depends critically upon the motion system.

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

Figure 18 depicts the Static BCS-to-MOC Filter connections that are simulated in Francis and Grossberg (1996a). Oriented boundary signals from the Static BCS feed into like-oriented sustained cells and unoriented transient cells in the MOC filter that correspond to the same retinal location. This BCS-MOC Filter pathway, which models a V2 $\rightarrow$ MT pathway *in vivo*, renders the MOC Filter sensitive to spatiotemporal changes in form as well as luminance, so that it can generate apparent motion signals in response to illusory contours and other boundary segmentations.

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

Figure 18

18 Simulation of Illusory Contour Apparent Motion

Figure 19 shows the results of simulating the Static BCS and MOC Filter interactions with illusory contour inducers. Figure 19a shows the inputs for a display presenting two sets of illusory contour inducers in sequence. Figure 19b shows the responses of Level 1 cells in the MOC Filter. These activities respond only at the location of the luminance increments. Figure 19c shows the responses of the BCS hypercomplex cells at the second competitive stage (see Figure 13), notably the illusory contour between the two luminance increments. The activities in Figure 19b and 19c feed into the sustained cells at Level 2 and the transient cells at Level 3 of the MOC Filter, whose outputs are multiplied to generate local motion signals at Level 4. Figure 19d shows the responses of the Level 4 cells. The tall spikes indicate the onset of the luminous inducers. The smaller hills mark the offset of different parts of the illusory contour. These responses are pooled by the Gaussian long-range filter to generate the Level 6 activities that are shown in Figure 19e. Due to the strong spatial competition between these cells, only one cell is active at a time. The location of the active

cell shifts continuously from the first stimulus to the second stimulus during the apparent motion display. This G-wave demonstrates apparent motion of the illusory contour.

Figure 19

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

Figure 20 shows the strength of the G-wave for different combinations of stimulus duration and ISI. Also plotted is a threshold value. When the stimulus parameters create a G-wave with a strength above threshold, then the motion is counted as observable. As the stimulus duration increases from 50 to 100 milliseconds, the intersection between the G-wave strength curve and the threshold shifts to a longer ISI. As the stimulus duration increases still further, the intersection between the G-wave strength curve and the threshold shifts to shorter ISIs. This inverted U was qualitatively explained in Section 16. The ISIs that produce intersections in the strength and threshold curves identify the upper and lower ISI values for perceiving apparent motion. Figure 12b plots those threshold ISI values.

Figure 20

As noted in Section 11, the persistence of illusory contours shows a shape qualitatively similar to the ISI thresholds in Figure 12. Contours that persist longer supply strong inputs to the motion system for longer durations so that greater ISIs continue to generate strong motion percepts. Figure 11c plots the upper ISI thresholds from the data and the simulation. The simulation thresholds fall between the thresholds of the two subjects. Moreover, the simulation thresholds are an inverted-U function of stimulus duration. The simulation of the BCS boundary signals to explain these apparent motion thresholds of Mather (1988) used the same parameters and equations as Francis *et al.* (1994) used to explain the visual persistence data of Meyer and Ming (1988). In summary, the model properties responsible for integrating form and motion information explain percepts of illusory contour apparent motion (Ramachandran, 1985) by linking dynamic persistence properties of illusory contour form perception (Meyer and Ming, 1988) to the dynamic properties of apparent motion (Mather, 1988).

19 Generalizations to Motion Grouping, Speed perception, and Morphing Motion

The Motion BCS reviewed above is now being extended and refined in a number of ways (Chey, Grossberg, and Mingolla, 1995a, 1995b; Nogueira, Grossberg, and Mingolla, 1993). One refinement explores whether the oriented sustained cells in Figure 2 can be replaced by a motion filter in which only transient cells occur, in keeping with data about the effects of magnocellular but not parvocellular LGN lesions on motion perception (Ferrera, Nealey, and Maunsell, 1992; Maunsell, Nealey, and Ferrera, 1992; Schiller, Logothetis, and Charles, 1990)

This revision maintains the short-range filters and the long-range Gaussian filters (including the G-wave), but uses very short-range anisotropic inhibition and anisotropic short-range spatial filters to achieve directional sensitivity. Using a range of scales, the model reproduces empirically derived speed discrimination curves and simulates data showing how visual speed perception and discrimination can be affected by stimulus contrast, duration, dot density and spatial frequency (Campbell and Maffei, 1981; De Bruyn and Orban, 1988; Diener, Wist, Dichgans, and Brandt, 1976; Stone and Thompson, 1992; Watamaniuk, Grzywacz, and Yuille, 1993)

The model has been further extended to incorporate a motion grouping mechanism that is a homolog of the grouping network of the Static BCS. Using both the multiple-scale MOC filter and its grouping network, the model simulates data concerning how the perceived speeds of moving lines are affected by their length and angle (Castet, Lorenceau, Shiffrar, and Bonnet, 1993), how the barberpole illusion (Wallach, 1976) is produced, how it can be affected by various configurational changes, and how plaid patterns move both coherently and incoherently. In addressing plaid pattern motion, the model provides explanations of when plaid patters cohere or do not (Adelson and Movshon, 1982; Kim and Wilson, 1993; Lindsey and Todd, 1995), how contrast effects the perceived speed and direction of moving plaids (Stone, Watson, and Mulligan, 1990), and why the movement of so-called Type 2 patterns differs from those of Type 1 patterns (Ferrera and Wilson, 1990, 1991; Yo and Wilson, 1992).

Within this version of the Motion BCS, motion direction and speed both emerge as part of an interactive motion grouping or segmentation process. The model proposes a solution to the global aperture problem by showing how information from feature tracking points, namely locations form which unambiguous motion directions can be computed, can propagate to ambiguous motion direction points, and capture the motion signals there. The model does this without computing intersections of constraints, as in the model of Adelson and Movshon (1982), or parallel Fourier and non-Fourier pathways, as in the model of Wilson, Ferrera, and Yo (1993). Instead, the model uses orientationally-unselective cell responses to activate directionally-tuned transient cells. These transient cells, in turn, activate spatially short-range filters and competitive mechanisms over multiple spatial scales to generate speed-tuned and directionally-tuned cells. Spatially long-range filters and top-down feedback from grouping cells are then used to track motion of featural points and to select and propagate correct motion directions to ambiguous motion points. Top-down grouping can also prime the system to attend a particular motion direction.

In another direction, the Motion BCS has been used to simulate data bout attentional speedup (Hikosaka *et al.*, 1993) and morphing motion (Tse and Cavanagh, 1995). In particular, the motion percept when a spot is followed by a line was explained by Hikosaka *et al.* (1993) as an attentional speed up of information processing near the spot. Tse and Cavanagh (1995) argued that motion percepts when more complex forms change shape (morphing) are controlled more by featural factors and that "line motion occurs after surface parsing". Baloch and Grossberg (1996) have explained all their results and some new ones using available neural models of form and motion processing, such as the models of boundary segmentation, surface filling-in, motion processing, figure-ground separation, form-motion interactions, and spatial attention that are reviewed above. This analysis suggests that different displays engage different combinations of these processes. Simulations include percepts

when the forms start out with the same shape but different color, start out with the same colors but different shape, cause pop-out of an occluding object, are morphed into occluding and occluded objects, are morphed into Kanizsa squares, are morphed into different shapes, and change color periodically.

The Motion BCS model's broad explanatory range makes it a viable alternative to Reichardt-style motion models (Van Santen and Sperling, 1984, 1985) or motion energy models (Adelson and Bergen, 1985; Watson and Ahumada, 1985) which have, until recently, been the most promising candidates for explaining how the brain processes motion information. The close homologs between the Motion BCS for generating directional segmentations of moving forms and the Static BCS for generating orientational segmentations of static forms call attention to the fact that these models form part of a more general theory of preattentive vision that is steadily being developed.

References

Adelson, E.H. and Bergen, J.R. (1985). Spatio-temporal energy models for the perception of motion. *Journal of the Optical Society of America*, **2**, 284–299.

Adelson, E.H. and Movshon, J.A. (1982). Phenomenal coherence of moving visual patterns. *Nature*, **300**, 523–525.

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

Albright, T.D. (1995). My most true mind thus makes mine eye untrue. *Trends in Neurosciences*, **18**, 331–333.

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

Anderson, S.J. and Burr, D.C. (1987). Receptive field size of human motion detection units. *Vision Research*, **27**, 621–635.

Anstis, S.M. and Ramachandran, V.S. (1987). Visual inertia in apparent motion. *Vision Research*, **27**, 755–764.

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

Assad, J.A. and Maunsell, J.H.R. (1995). Neural correlates of inferred motion in primate posterior parietal cortex. *Nature*, **373**, 518–521.

Baloch, A.A. and Grossberg, S. (1996). Neural dynamics of morphing motion. *Investigative Ophthalmology and Visual Science*, **37**(4), March 15, 1996, submitted.

Bartley, S.H. (1941). **Vision, a study of its basis**. New York: D. Van Nostrand.

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

Bergen, J.R. and Landy, M.S. (1991). Computational modeling of visual texture segregation. In M.S. Landy and J.A. Movshon (Eds.) **Computational models of visual processing**. Cambridge, MA: MIT Press, pp. 253–271.

Boring, E.G. (1950). **A history of experimental psychology**. Englewood Cliffs, NJ: Prentice-Hall.

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

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

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

Braddick, O. and Adlard, A. (1978). Apparent motion and the motion detector. In J.C. Armington, J. Krauskopf, and B.R. Wooten (Eds.) **Visual psychophysics and psychology**, New York: Academic Press.

Breitmeyer, B. (1984). **Visual Masking: An Integrative Approach.** New York: Oxford University Press.

Breitmeyer, B.G. and Ritter, A. (1986). Visual persistence and the effect of eccentric viewing, element size, and frame duration on bistable stroboscopic motion percepts. *Perception & Psychophysics*, **39**, 275–280.

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

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

Campbell, F.W. and Maffei, L. (1981). The influence of spatial frequency and contrast on the perception of moving patterns. *Vision Research*, **21**, 713–721.

Castet, E. (1994). Effect of the ISI on the visible persistence of a stimulus in apparent motion. *Vision Research*, **34**, 2103–2144.

Castet, E., Lorenceau, J., Shiffrar, M., and Bonnet, C. (1993). Perceived speed of moving lines depends on orientation, length, speed and luminance. *Vision Research*, **33**, 1921–1936.

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

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

Chey, J., Grossberg, S., and Mingolla, E. (1995a). 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. (1995b). Neural dynamics of motion speed and directional grouping: From aperture ambiguity to plaid coherence. Technical report CAS/CNS-TR-95-031. Boston, MA: Boston University.

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

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

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

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

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

De Bruyn, B. and Orban, G.A. (1988). Human velocity and direction discrimination measured with random dot patterns. *Vision Research*, **28**, 1323–1335.

Diener, H.C., Wist, E.R., Dichgans, J., and Brandt, T. (1976). The spatial frequency effect on perceived velocity. *Vision Research*, **16**, 169–176.

Desimone, R. (1991). Face-selective cells in the temporal cortex of monkeys. *Journal of Cognitive Neuroscience*, **3**, 1–8.

Desimone, R. (1992). Neural circuits for visual attention in the primate brain. In G.A. Carpenter and S. Grossberg (Eds.), **Neural networks for vision and image processing**. Cambridge, MA: MIT Press, pp. 343–364.

Desimone, R., Schein, S.J., Moran, J., and Ungerleider, L.G. (1985). Contour, color, and shape analysis beyond the striate cortex. *Vision Research*, **25**, 441–452.

Desimone, R. and Ungerleider, L.G. (1989). Neural mechanisms of visual processing in monkeys. In F. Boller and J. Grafman (Eds.) **Handbook of neuropsychology, Volume 2**. Amsterdam: Elsevier Publishing, pp. 267–299.

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

Eriksen, C.W. and Murphy, T.D. (1987). Movement of attentional focus across the visual field: A critical look at the evidence. *Perception & Psychophysics*, **42**, 29–305.

Exner, S. (1875). Ueber das Sehen von Bewegungen und die Theorie des zusammengesetzen Auges. *Sitzungsberichte Akademie Wissenschaft Wien*, **72**, 156–190.

Ferrera, V.P., Nealey, T.A., and Maunsell, J.H.R. (1992). Mixed parvocellular and magnocellular geniculate signals in visual area V4. *Nature*, **358**, 756–758.

Ferrera, V.P. and Wilson, H.R. (1990). Perceived direction of moving two-dimensional patterns. *Vision Research*, **30**, 273–287.

Ferrera, V.P. and Wilson, H.R. (1991). Perceived speed of moving two-dimensional patterns. *Vision Research*, **31**, 877–893.

Francis, G. (1994) Cortical dynamics of lateral inhibition: Visual persistence and ISI. Tech Report 94-1. W. Lafayette, IN: Purdue Mathematics and Psychology Program.

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

Francis, G. and Grossberg, S. (1996b). Cortical dynamics of binding and reset: Orientational 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.

Gochin, P. (1990). Pattern recognition in primate temporal cortex: But is it ART? **Proceedings of the international joint conference on neural networks, I**, 77–80. Hillsdale, NJ: Erlbaum Associates.

Gochin, P.M., Miller, E.K., Gross, C.G., and Gerstein, G.L. (1991). Functional interactions among neurons in inferior temporal cortex of the awake macaque. *Experimental Brain Research*, **84**, 505–516.

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

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

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

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

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

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

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

Grossberg, S. (1983). The quantized geometry of visual space: The coherent computation of depth, form, and lightness. **Behavioral and Brain Sciences, 6**, 625–657.

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

Grossberg, S. (1987a). Cortical dynamics of three-dimensional form, color, and brightness perception I: Monocular theory. *Perception & Psychophysics*, **41**, 97–116.

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

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

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

Grossberg, S. and McLoughlin (1995). Cortical dynamics of 3-D surface perception: Binocular and half-occluded scenic images. Tech Report CAS/CNS-TR-95-022. Boston, MA: Boston University.

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 & 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 & Psychophysics*, **53**, 243–278.

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*, Special issue on ATR, in press.

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

Grossberg, S. and Rudd, M. (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 & Psychophysics*, **43**, 241–277.

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

Harries, M.H. and Perrett, D.I. (1991). Visual processing of faces in temporal cortex: Physiological evidence for a modular organization and possible anatomical correlates. *Journal of Cognitive Neuroscience*, **3**, 9–24.

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

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

Kim, J. and Wilson, H.R. (1993). Dependence of plaid motion coherence on component grating directions. Review copy.

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

Kolers, P.A. and von Grünau, M. (1975). Visual construction of color is digital. *Science*, 187:757–759.

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

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

LaBerge, D. and Brown, V. (1989). Theory of attentional operations in shape identification. *Psychological Review*, **96**, 101–124.

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

Lindsey, D.T. and Todd, J.T. (1995). On the relative contributions of motion energy and transparency to the perception of moving plaids. *Vision Research*, in press.

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.

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

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

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

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

Maunsell, J.H.R., Nealey, T.A., and Ferrera, V.P. (1992). Magnocellular and parvocellular contributions to neuronal responses in monkey visual cortex. *Investigative Ophthalmology and Visual Science*, **33**(4), Abstract # 1047.

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

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

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

Miller, E.K., Li, L., and Desimone, R. (1991). A neural mechanism for working and recognition memory in inferior temporal cortex. *Science*, **254**, 1377–1379.

Mishkin, M. (1982). A memory system in the monkey. *Philosophical Transactions Royal Society of London B*, **298**, 85–95.

Mishkin, M. and Appenzeller, T. (1987). The anatomy of memory. *Scientific American*, **256**, 80–89.

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

Nakayama, K. and Silverman, G.H. (1984). Temporal and spatial characteristics of the upper displacement limit for motion in random dots. *Vision Research*, **24**, 293–299.

Nakayama, K. and Silverman, G.H. (1985). Detection and discrimination of sinusoidal grating displacements. *Journal of the Optical Society of America, A, Optics and Image Science*, **2**, 267–273.

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

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.

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

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

Pantle, A.J. and Petersik, J.T. (1980). Effects of spatial parameters on the perceptual organization of a bistable motion display. *Perception & Psychophysics*, **27**, 307–312.

Pantle, A. and Picciano, L. (1976). A multistable movement display: Evidence for two separate motion systems in human vision. *Science*, **193**, 500–502.

Perrett, D.I., Mistlin, A.J., and Chitty, A.J. (1987). Visual cells responsive to faces. *Trends in Neuroscience*, **10**, 358–364.

Pessoa, L., Mingolla, E., and Neumann, H. (1995). A contrast- and luminance-driven multiscale network model of brightness perception. *Vision Research*, **35**, 2201–2223.

Peterhans, E. and von der Heydt, R. (1989). Mechanisms of contour perception in monkey visual cortex, II. Contours bridging gaps. *The Journal of Neuroscience*, **9**, 1749–1763.

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

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

Reichardt, W. (1961). Autocorrelation: A principle for the evaluation of sensory information by the central nervous system. In W.A. Rosenblith (Ed.) **Sensory Communication**, New York: Wiley and Sons.

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

Riches, I.P., Wilson, F.A.W., and Brown, M.W. (1991). The effects of visual stimulation and memory on neurons of the hippocampal formation and the neighboring parahippocampal gyrus and inferior temporal cortex of the primate. *Journal of Neuroscience*, **11**, 1763–1779.

Schiller, P.H., Logothetis, N.K., and Charles, E.R. (1990a). Functions of the colour-opponent and broad-band channels of the visual system. *Science*, **343**, 68–70.

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

Schwartz, E.L., Desimone, R., Albright, T., and Gross, C.G. (1983). Shape recognition and inferior temporal neurons. *Proceedings of the National Academy of Sciences*, **80**, 576–5778.

Squires, P.C. (1931). The influence of hue on apparent visual movement. *American Journal of Psychology*, **43**, 49–64.

Spitzer, H., Desimone, R., and Moran, J. (1988). Increased attention enhances both behavioral and neuronal performance. *Science*, **240**, 338–340.

Stone, L.S. and Thompson, P. (1992). Human speed perception is contrast dependent. *Vision Research*, **32**, 1535–1549.

Stone, L.S., Watson, A.B., and Mulligan, J.B. (1990). Effect of contrast on the perceived direction of a moving plaid. *Vision Research*, **30**, 1049–1067.

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

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

Ternus, J. (1926/1950). Experimentelle Untersuchungen über phänomenale Identität. *Psychologische Forschung*, **7**, 81–136. Abstracted and translated in W.D. Ellis (Ed.), **A sourcebook of Gestalt psychology**. New York: Humanities Press, 1950.

Tse, P. and Cavanagh, P. (1995). Line motion occurs after surface parsing. *Investigative Ophthalmology and Visual Sciences*, **36**(4), March 15, 1995, Abstract #1919, p. S417.

Ungerleider, L.G. and Mishkin, M. (1982). Two cortical visual systems: Separation of appearance 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, 549–586.

van der Waals, H.G. and Roelofs, C.O. (1930). Optische scheinbewegung. *Zeitschrift für Psychologie und Physiologie des Zinnesorgane*, **114**, 241–288.

van der Waals, H.G. and Roelofs, C.O. (1931). Optische scheinbewegung. *Zeitschrift für Psychologie und Physiologie des Zinnesorgane*, **115**, 91–190.

van Santen, J.P.H. and Sperling, G. (1984). Temporal covariance model of human motion perception. *Journal of the Optical Society of America*, **1**, 451–473.

van Santen, J.P.H. and Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America*, **2**, 300–321.

von der Heydt, R., Hänny, P., and Dürsteler, M.R., (1991). The role of orientation disparity in stereoscopic perception and the development of binocular correspondence. In E. Srastyán and P. Molnár (Eds.). **Advanced Physiological Science, Sensory Functions, Vol. 16**. New York, NY: Pergamon Press.

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

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

Wallach, H. (1976). **On perception**. New York, NY: Quadrangle Press.

Watamaniuk, S.N.J., Grzywacz, N.M., and Yuille, A.L. (1993). Dependence of speed and direction perception on cinematogram dot density. *Vision Research*, **33**, 849–859.

Watson, B. and Ahumada, Jr., A.J. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America*, **2**, 322–243.

Wertheimer, M. 1912/1961. Experimentelle studien über das sehen von bewegung. *Zeitschrift für Psychologie*, 61:161–265. Translated in part in T. Shipley (Ed.), **Classics in psychology**. New York: Philosophical Library, 1961.

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

Wilson, H.R. and Kim, J. (1994). A model for motion coherence and transparency. *Visual Neuroscience*, **11**, 1205–1220.

Yo, C. and Wilson, H.R. (1992). Perceived direction of moving two-dimensional patterns depends on duration, contrast, and eccentricity. *Vision Research*, **32**, 135–147.

FIGURE CAPTIONS

Figure 1. The Ternus display. (a) Three spots are presented in each frame in such a way that the two leftwardmost spots in Frame 2 occupy the same positions as the two rightwardmost spots in Frame 1. The two frames are repeatedly cycled with ISIs inserted between them. At very short ISIs, all dots appear to flicker in place. At longer ISIs the dots at shared positions appear to remain stationary, while apparent motion occurs between the leftwardmost spot in Frame 1 and the rightwardmost spot in Frame 2 ("element motion"). At still longer ISIs, the three dots appear to move from Frame 1 to Frame 2 and back as a group ("group motion"). (b) When the dots in successive frames have opposite contrast with respect to the frame, only group motion occurs at the ISIs where element motion occurred in (a). [Reprinted with permission from Grossberg and Rudd (1992).]

Figure 2. A one-dimensional MOC Filter: The input pattern at Level 1 is spatially and temporally filtered by sustained response cells at Level 2. The sustained cells have oriented receptive fields that are sensitive to the direction-of-contrast in the image, either dark-to-light or light-to-dark. Level 2 cells play the role of a short-range spatial filter. Spatial and temporal averaging are also carried out by transient response cells at Level 3. The transient cells have unoriented receptive fields that are sensitive to the direction of contrast change in the cell input. The upward arrow denotes transient on-cells that are activated by a transition from dark to light. The downward arrow denotes transient off-cells that are activated by a transition from light to dark. Level 4 cells combine sustained and transient cell signals multiplicatively and are thus rendered sensitive to both direction-of-motion and direction-of-contrast. Level 5 cells sum across space via a long-range Gaussian spatial filter, and across the two types of Level 4 cells. Level 5 cells are thus sensitive to direction-of-motion but insensitive to direction-of-contrast. [Reprinted with permission from Grossberg and Rudd (1992).]

Figure 3. Spatial responses at various levels of the MOC Filter to a point input. (a) Sustained activity of a Level 2 cell. (b) Total input pattern to Level 5 after convolution with a Gaussian kernel. (c) Contrast-enhanced output of Level 5 centered at the location of the input maximum. [Reprinted with permission from Grossberg and Rudd (1992).]

Figure 4. Temporal response of the MOC Filter to a point input. (a) The input is presented at a brief duration at location 1. (b) Sustained cell activity at 1 gradually builds after the input onset, then decays after offset. (c) Growth of the input pattern to Level 5 with transient cell activity held constant. The activity pattern retains a Gaussian shape centered at the location of the input, that waxes and wanes through time without spreading across space. The figure illustrates the waxing phase. [Reprinted with permission from

Grossberg and Rudd (1992).]

Figure 5. Temporal response of the sustained cells $x_i(t)$ at Level 2 to two brief successive point inputs $J_i(t)$ at locations $i = 0$ and $i = L$. For an appropriately timed display, the decaying response at position 0 overlaps in time the rising response at position L. [Reprinted with permission from Grossberg and Rudd (1992).]

Figure 6. Simulated MOC Filter response to a two-flash display. Successive rows correspond to increasing times following the Frame 1 offset. (a) The two lower curves in each row depict the total input R_i at position i of Level 5 due to each of the two flashes. The input due to the left flash decreases while the input due to the right flash increases. The summed input due to both flashes is a traveling wave whose maximum value across space moves continuously between the two flash locations. (b) Position over time of the contrast-enhanced Level 5 response. [Reprinted with permission from Grossberg and Rudd (1992).]

Figure 7. Motion paths generated by MOC Filters with different Gaussian filter kernel widths K in (9). The simulated motion paths are plotted in a space-time diagram wherein each rectangle indicates the spatiotemporal boundaries of one flash in a two flash display. All the motion paths intersect at a point halfway between the two flash locations. [Reprinted with permission from Grossberg and Rudd (1992)].

Figure 8. Schematic diagram of anatomical connections and neuronal selectivities of early visual areas in the macaque monkey. LGN = lateral geniculate nucleus (parvocellular and magnocellular divisions). Divisions of V1 and V2: blob = cytochrome oxidase blob regions; interblob = cytochrome oxidase-poor regions surrounding the blobs; 4B = lamina 4B; thin = thin (narrow) cytochrome oxidase strips; interstripe = cytochrome oxidase-poor regions between the thin and thick stripes; thick = thick (wide) cytochrome oxidase strips; V3 = visual area 3; V4 = visual area(s) 4; MT = middle temporal area. Areas V2, V3, V4, MT have connections to other areas not explicitly represented here. Area V3 may also receive projections from V2 interstripes or thin stripes. Heavy lines indicate robust primary connections, and thin lines indicate weaker, more variable connections. Dotted lines represent observed connections that require additional verification. Icons: rainbow = tuned and/or opponent wavelength selectivity (incidence at least 40%); angle symbol = orientation selectivity (incidence at least 20%); spectacles = binocular disparity selectivity and/or strong binocular interactions (V2) (incidence at least 20%); arrow = direction of motion selectivity (incidence at least 20%). [Adapted with permission from DeYoe and van Essen (1988).]

Figure 9. The black and white squares of the background group into long vertical and horizontal boundaries that cross the figure, even though the black–white and white–black

edges have opposite contrast polarity, or direction–of–contrast. Likewise, the intermediate gray of the disks creates alternating black–gray and white–gray edges that have opposite direction–of–contrast, yet group into unitary disk boundaries. Thus the boundary system pools signals from opposite directions–of–contrast to bridge textured and shaded contrast polarity reversals. [Adapted from Kanizsa (1979).]

Figure 10. (a) Macrocircuit of monocular and binocular interactions of the boundary contour system (BCS) and the feature contour system (FCS): Left eye and right eye monocular preprocessing stages (MP_L and MP_R) send parallel pathways to the BCS (boxes with vertical lines, designating oriented responses) along pathways 1, and to the FCS (boxes with three pairs of circles, designating opponent colors) along pathways 2. The monocular signals BCS_L and BCS_R activate simple cells which, in turn, activate pathways, labeled 3 and 4, to generate a binocular boundary segmentation using the complex, hypercomplex, and bipole cell interactions of Figure 13. The binocular segmentation generates output signals to the monocular Filling-In Domains, or FIDOs, of the FCS via pathways labeled 6. This interaction captures and fills-in those monocular FCS signals from pathways 5 that are consistent with the binocular boundaries, and suppresses binocularly inconsistent FCS signals. In this way, the monocular FCS signals are separated into distinct surface depths. Reciprocal FCS $\rightarrow$ BCS interactions along pathways 7 enhance consistent boundaries and suppress boundaries corresponding to more distant surfaces. The monocular FCS signals are also binocularly matched via pathways 8 at the binocular FIDOs. In addition, the surviving FCS signals at the monocular FIDOs use pathways 9 to inhibit redundant FCS signals from pathways 8 at more distant surfaces than their own. The surviving FCS signals interact with augmented binocular BCS signals from pathways 10 to fill-in a multiple-scale surface representation of Form-And-Color-And-DEpth, or FACADE. Processing stages MP_L and MP_R are compared with LGN data; the simple-complex cell interaction with V1 data; the hypercomplex-bipole interaction with V2 and (possibly) V4 data, notably about Interstripes; the monocular FCS interaction with Blob and Thin Stripe data; and the FACADE representation with V4 data (see Figure 8). [Adapted from Grossberg (1994) with permission.]

Figure 11. (a) Contour maps depicting the percentage of trials in which apparent motion of an illusory contour was reported as a function of stimulus duration and ISI. black: 85 - 100%; mid-grey: 70 - 85%; light grey: 55 - 70%; blank, less than 55%. [Used with permission from Mather (1988).] (b) Computer simulation of upper and lower ISI thresholds as a function of stimulus duration. (c) Upper ISI thresholds for perceiving illusory contour apparent motion for subjects and the simulation. The simulation thresholds fall between the data of the subjects. For each subject and the simulation, maximal ISI takes a peak value at an intermediate stimulus duration. [Reprinted with permission from Francis and Grossberg (1996a).]

Figure 12. Upper and lower ISI thresholds as a function of spatial separation for three stimulus durations. Increasing flash duration decreases the threshold ISI values. The upper ISI threshold decreases with spatial separation and the lower ISI threshold increases with

spatial separation. (a) Psychophysical data. [Redrawn from Kolers (1972) after data from Neuhaus (1930).] (b) Simulated ISI threshold values. [Reprinted with permission from Francis and Grossberg (1996a).]

Figure 13. Boundary Contour System with embedded gated dipoles. See text for details. [Reprinted with permission from Francis and Grossberg (1996a).]

Figure 14. The sustained cell short-range filter. Inputs are spatiotemporally filtered by sustained cells with individual oriented receptive fields and temporal filtering characteristics that are determined by the dynamics of a shunting membrane equation. The output of each sustained cell is rectified and thresholded. The outputs of a spatially aligned array of cells with like orientation, direction-of-contrast, and direction-of-motion are pooled. The breadth of the spatial pooling scales with the size of the simple cell receptive fields, as in Panels a and b. [Reprinted with permission from Grossberg and Rudd (1992).] (c). Visual inertia in apparent motion measured by Anstis and Ramachandran (1987). (Ambiguous apparent motion was biased by priming dots, and the degree of bias [inertia] was measured as a function of the interstimulus interval [ISI] between the priming dot and test. The bias induced by the priming dots was about 12% at short ISIs and fell monotonically to about 7% for ISIs exceeding 500 ms. [Reprinted with permission from Anstis and Ramachandran (1987).]

Figure 15. Responses over time of transient on and off cells. (a) On-cell responses are formed from the positive-rectified and thresholded time derivative of a spatiotemporally filtered image. The spatial filter has an unoriented on-center, off-surround receptive field. The temporal filter is based on the dynamics of a shunting membrane equation that time averages the spatially filtered input. The on-cell thus produces a time-averaged response to an increment in the input. (b) Off cells are formed from the negative-rectified and thresholded time-averaged response to a decrement in the input. [Reprinted with permission from Grossberg and Rudd (1992).]

Figure 16. Transient cell gating of sustained cell activities to produce directionally sensitive responses. The short-range filter, which is constructed from like-oriented simple cells, responds ambiguously to a contrast pattern (dark-light in the illustration) moving either to the right or to the left. This ambiguity of motion direction is eliminated by gating the short-range filter response with either a transient on-cell response (to produce a left-motion signal) or a transient off-cell response (to produce a right-motion signal). [Reprinted with permission from Grossberg and Rudd (1992).]

Figure 17. Combination of like direction-of-motion activities across space by a long-range Gaussian filter. Local direction-sensitive responses of opposite direction-of-contrast, over a range of orientations, are gated by transient cells of opposite types to produce like

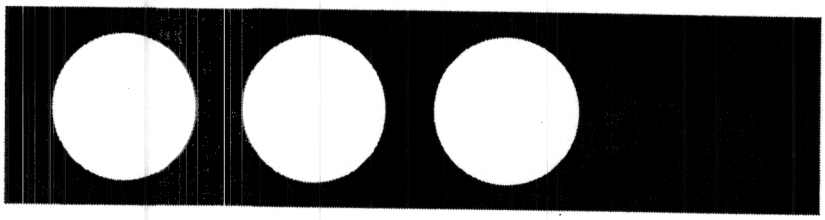
direction-of-motion signals. These local signals are combined by a long-range Gaussian spatial kernel to produce a spatially broad pattern of activity across the Level 5 network. This broad pattern is then contrast enhanced by a competitive, or lateral inhibitory, interaction. The contrast enhancement restores positional information. [Reprinted with permission from Grossberg and Rudd (1992).]

Figure 18. Model of form and motion integration. Oriented boundary signals in the BCS feed into like-oriented sustained cells and unoriented transient cells in the MOC Filter. [Reprinted with permission from Francis and Grossberg (1996a).]

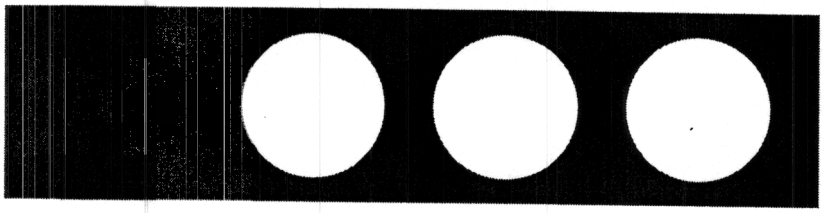
Figure 19. Computer simulation of how the MOC Filter with BCS input responds to an illusory contour apparent motion display. (a) The stimuli consists of two sets of illusory contour inducers. (b) Level 1 activities in the MOC Filter do not create illusory contours. (c) Boundary signals generated by the BCS create an illusory contour between the two inducers. (d) Local motion signals at Level 4 cells. The tall spikes are responses to the luminous inducers, while the smaller curves are produced by the illusory contours generated by the BCS. (e) Response of Level 5 global motion cells. The activity shifts continuously via a G-wave from the location of the first illusory contour to the location of the second illusory contour. [Reprinted with permission from Francis and Grossberg (1996a).]

Figure 20. G-wave strength for illusory contours as stimulus duration and ISI varies. Intersections between G-wave strength curves and the threshold mark upper and lower ISI thresholds. [Reprinted with permission from Francis and Grossberg (1996a).]

Frame 1

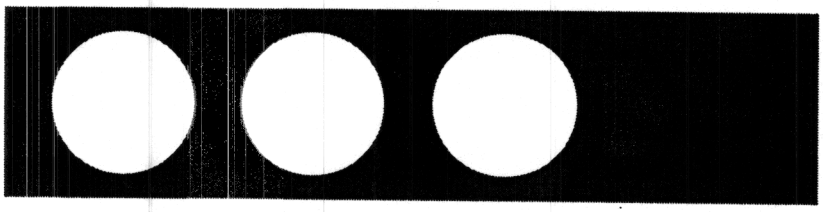


Frame 2

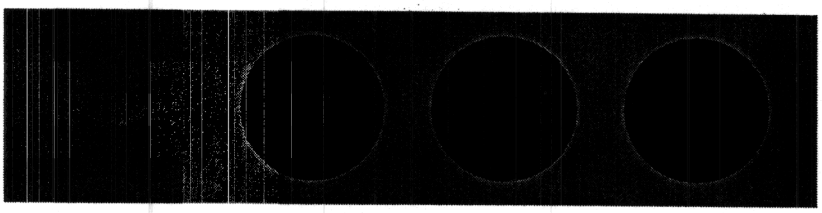


(a)

Frame 1

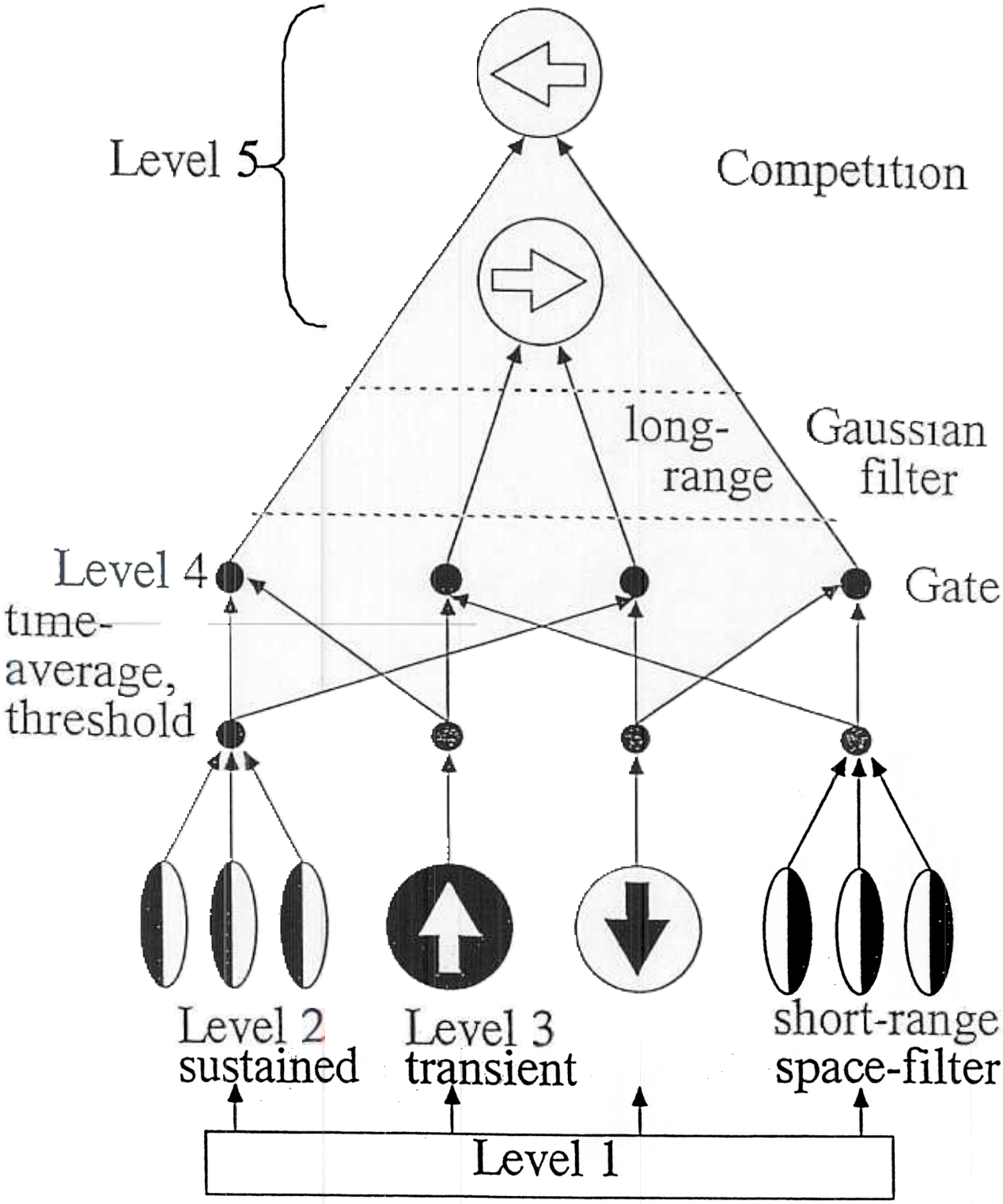


Frame 2



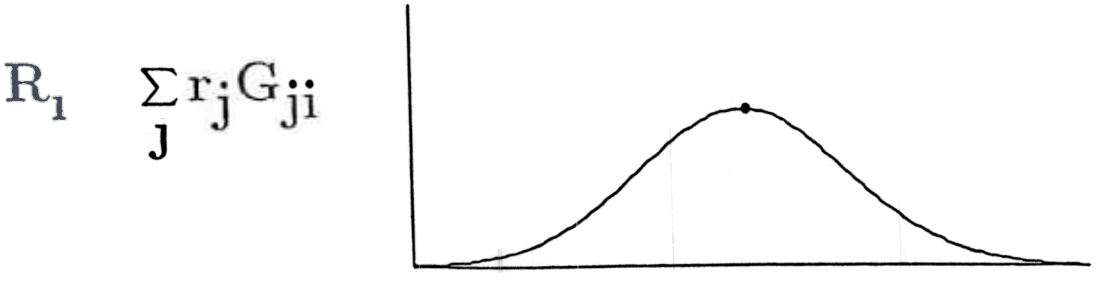
(b)

Figure 1

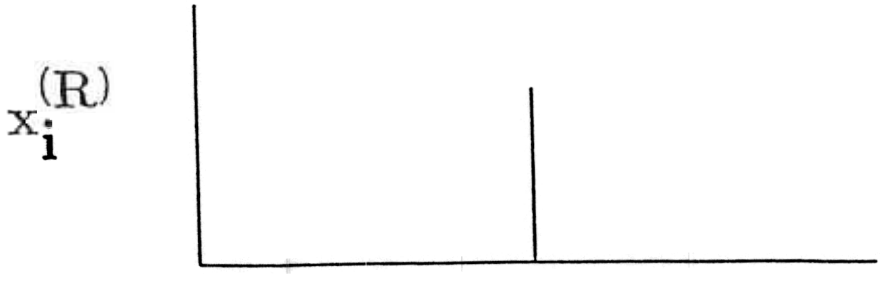

Level 5
Competition
long-
range
Gaussian
filter
Level 4
time-
average,
threshold
Gate
Level 2
sustained
Level 3
transient
short-range
space-filter
Level 1

Figure

(a)



(b)



(c)

Figure

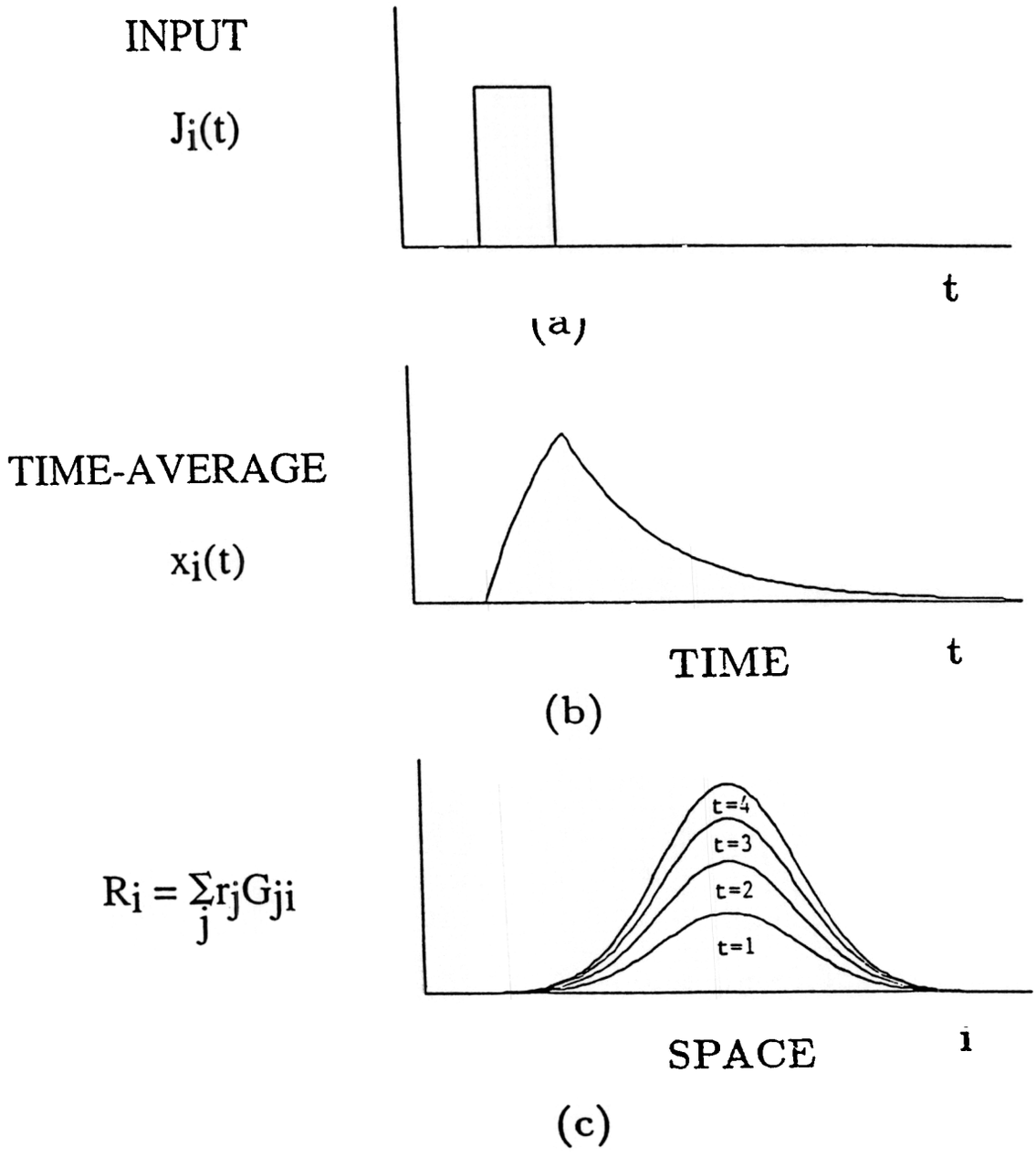

INPUT
$J_i(t)$
t
(a)
TIME-AVERAGE
$x_i(t)$
TIME
t
(b)
$R_i = \sum_j r_j G_{ji}$
t=4
t=3
t=2
t=1
SPACE
i
(c)

Figure 4

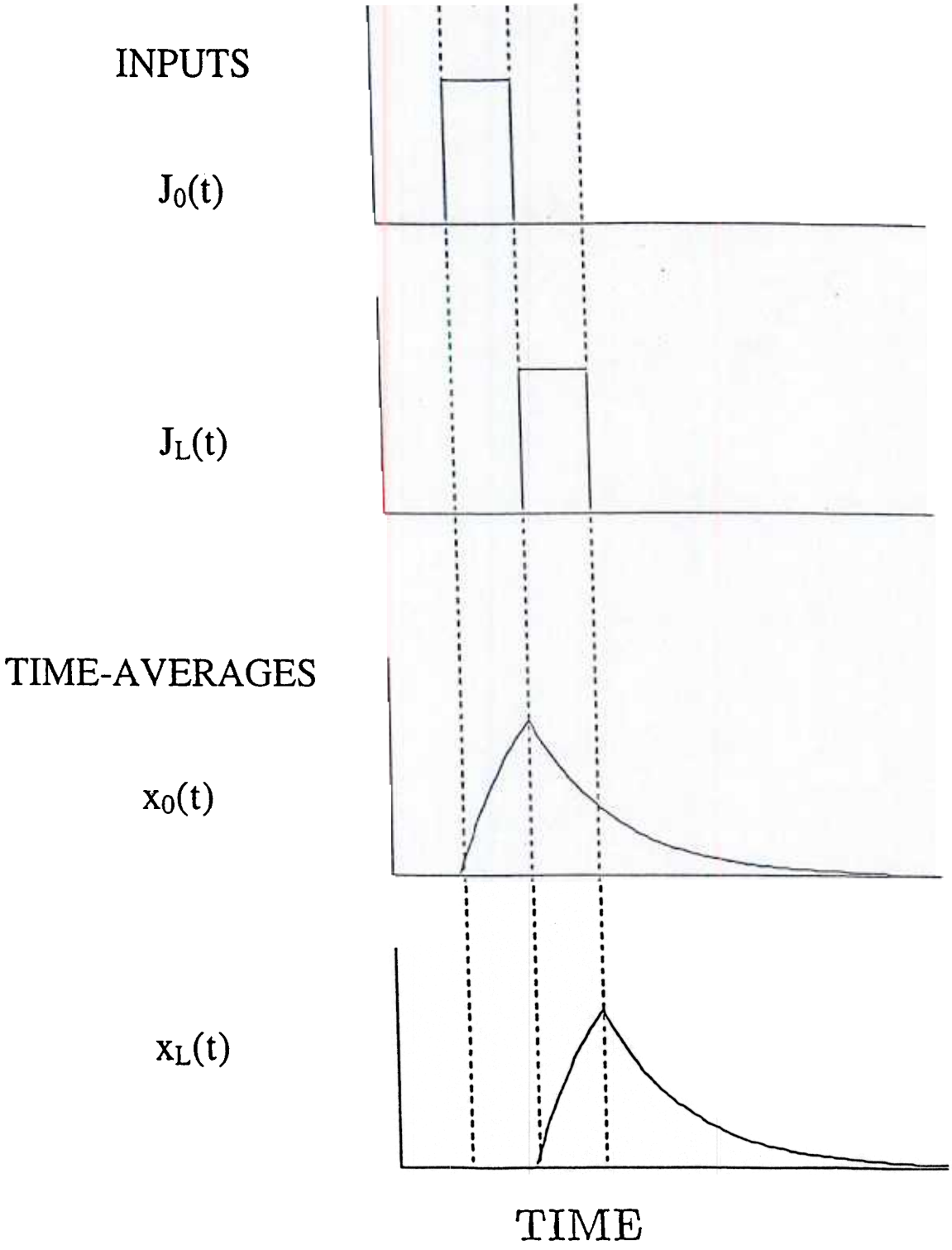


Figure 5

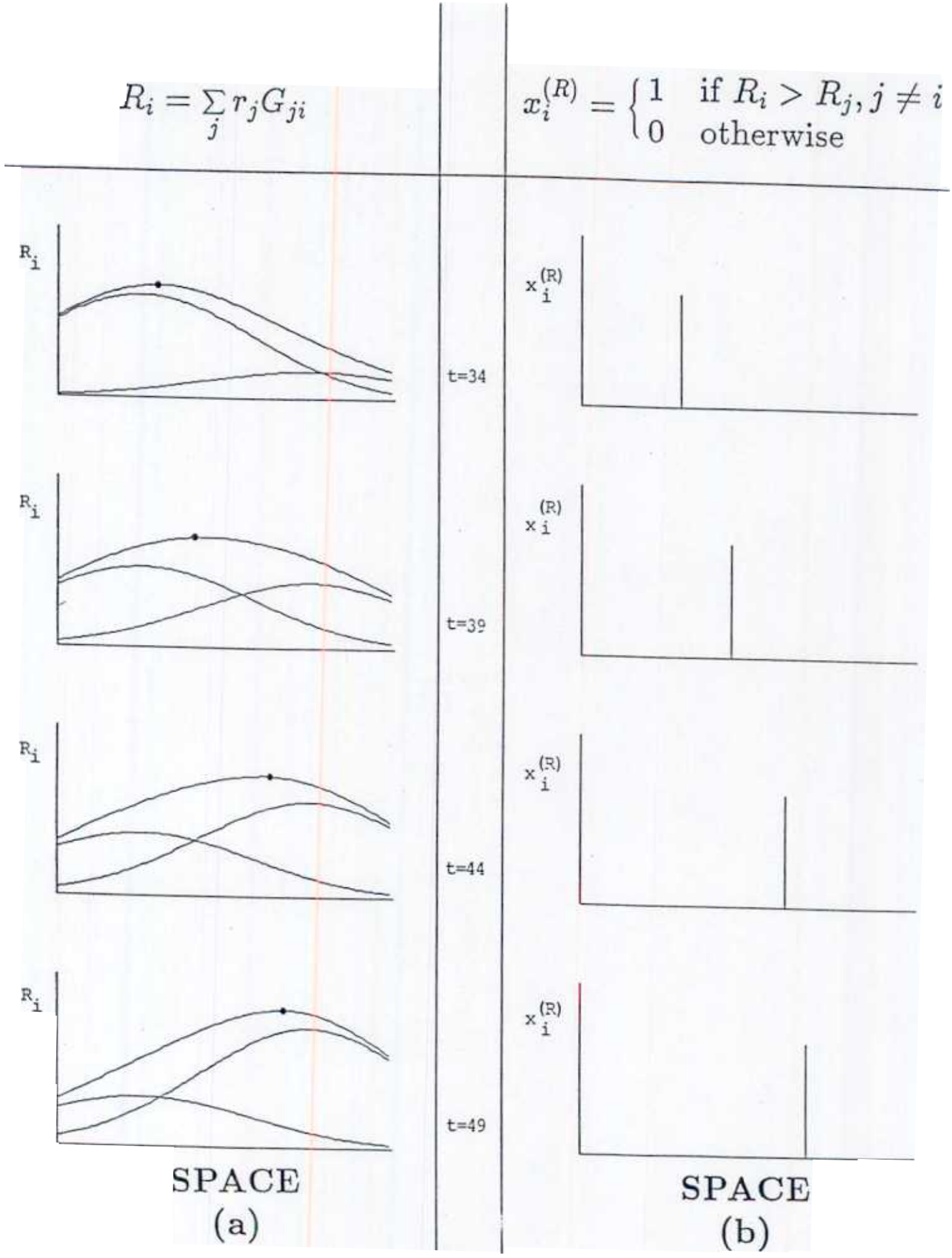
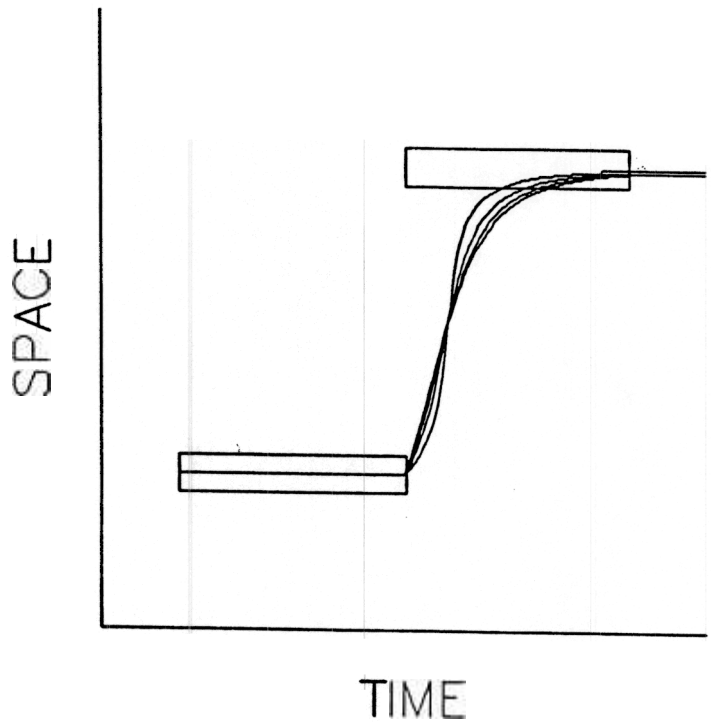

$R_i = \sum_j r_j G_{ji}$
$x_i^{(R)} = \begin{cases} 1 & \text{if } R_i > R_j, j \neq i \\ 0 & \text{otherwise} \end{cases}$
R_i
$x_i^{(R)}$
t=34
t=39
t=44
t=49
SPACE
(a)
SPACE
(b)

Figure 6



Figure

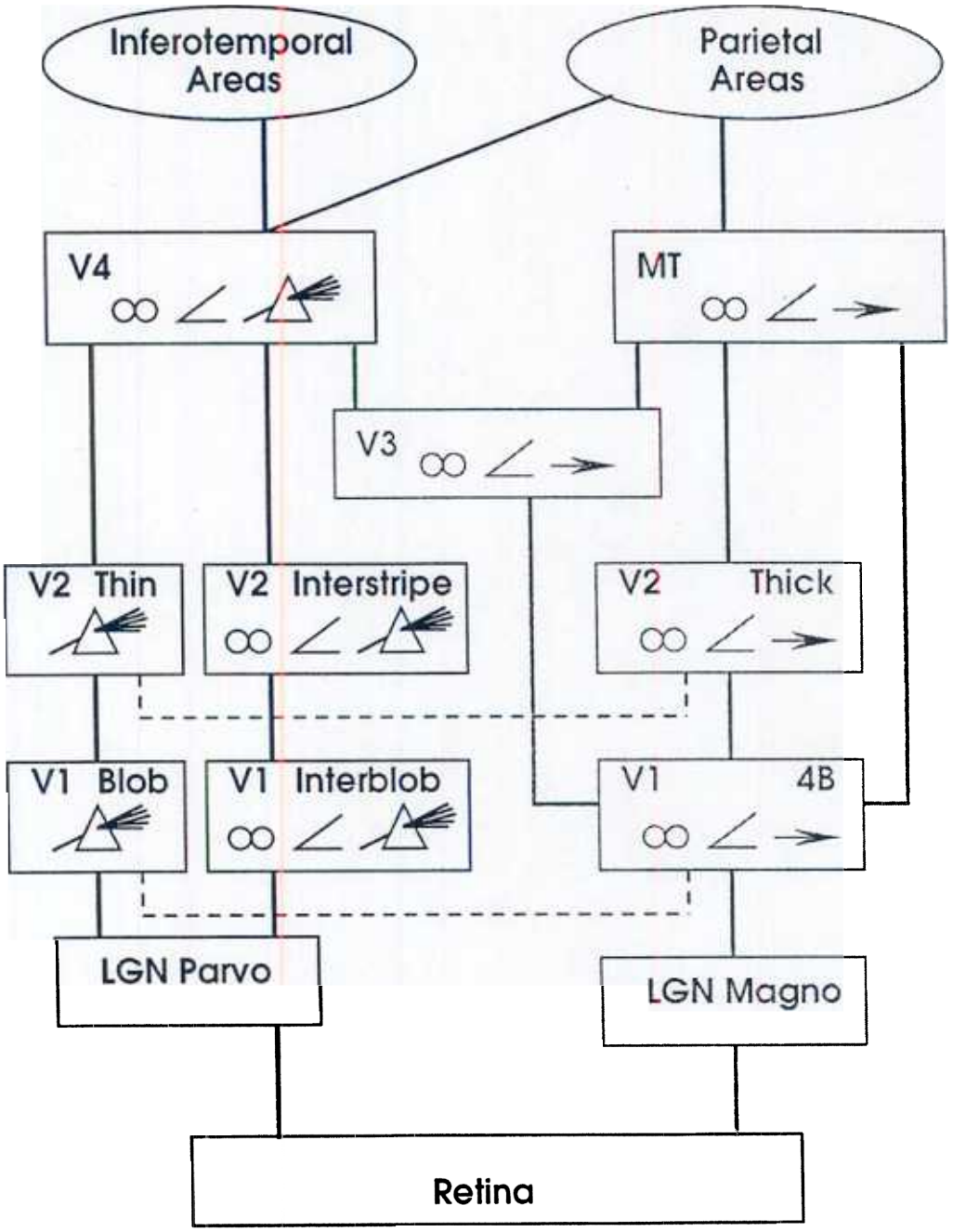
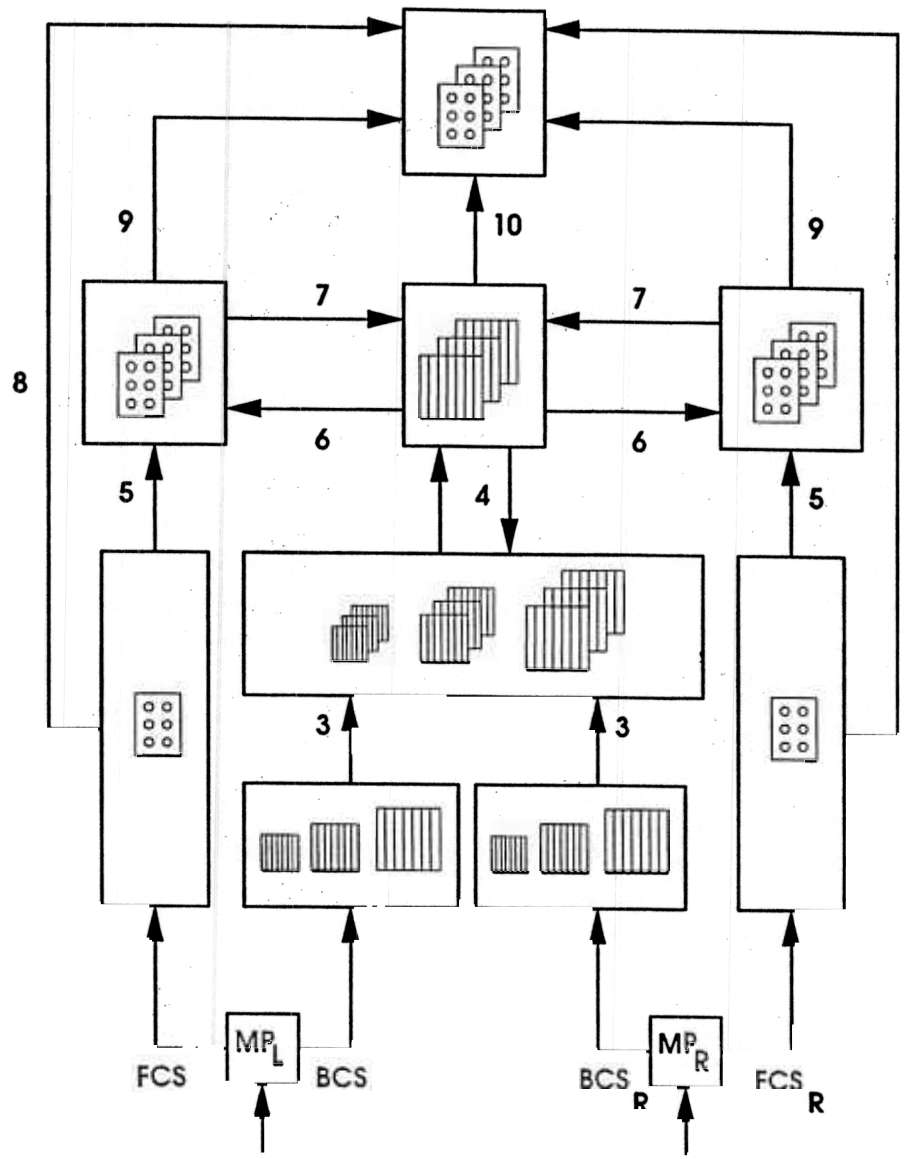
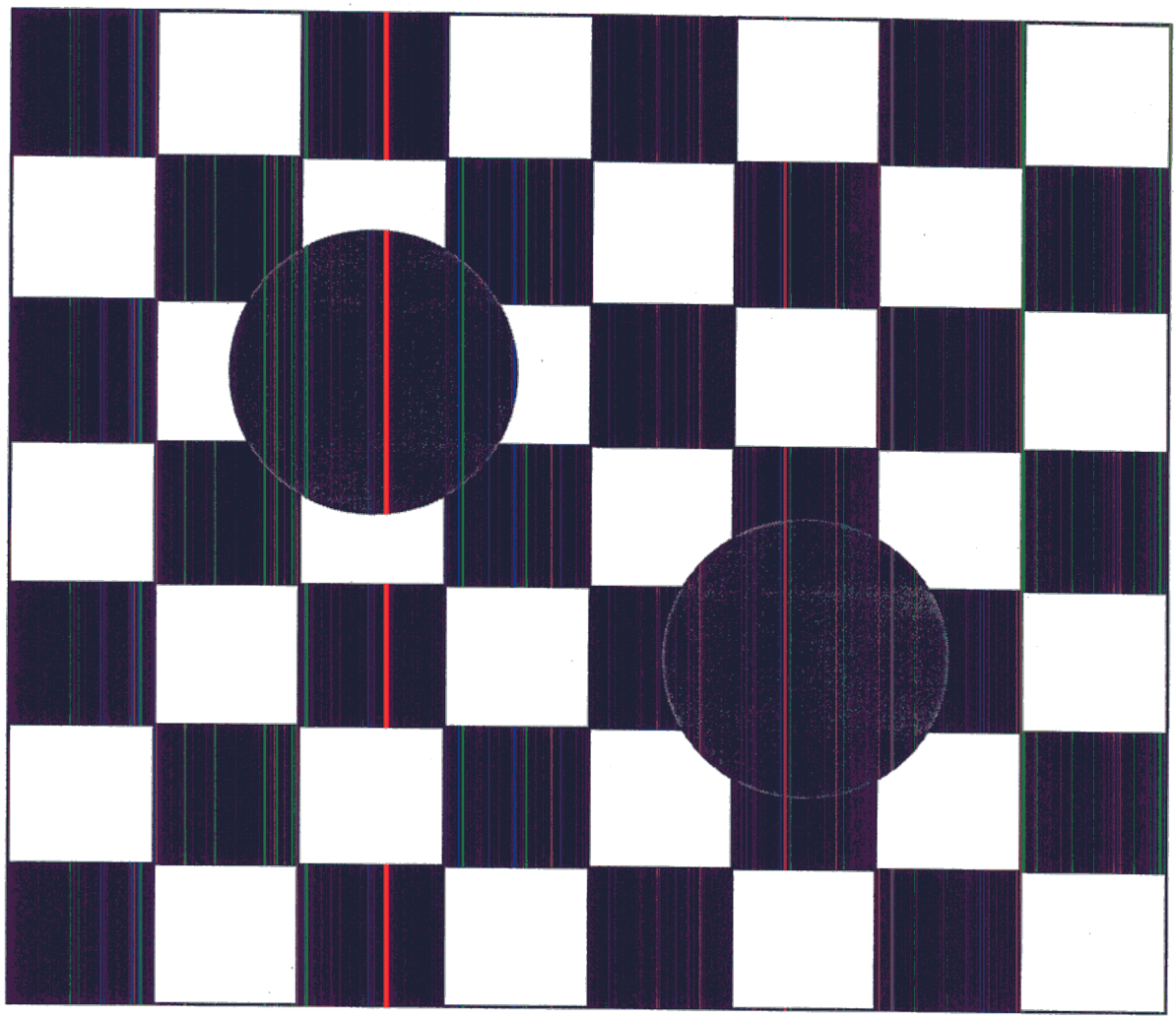

Inferotemporal Areas
Parietal Areas
V4
MT
V3
V2 Thin
V2 Interstripe
V2 Thick
V1 Blob
V1 Interblob
V1 4B
LGN Parvo
LGN Magno
Retina

Figure 8



igure



Fi ju 9

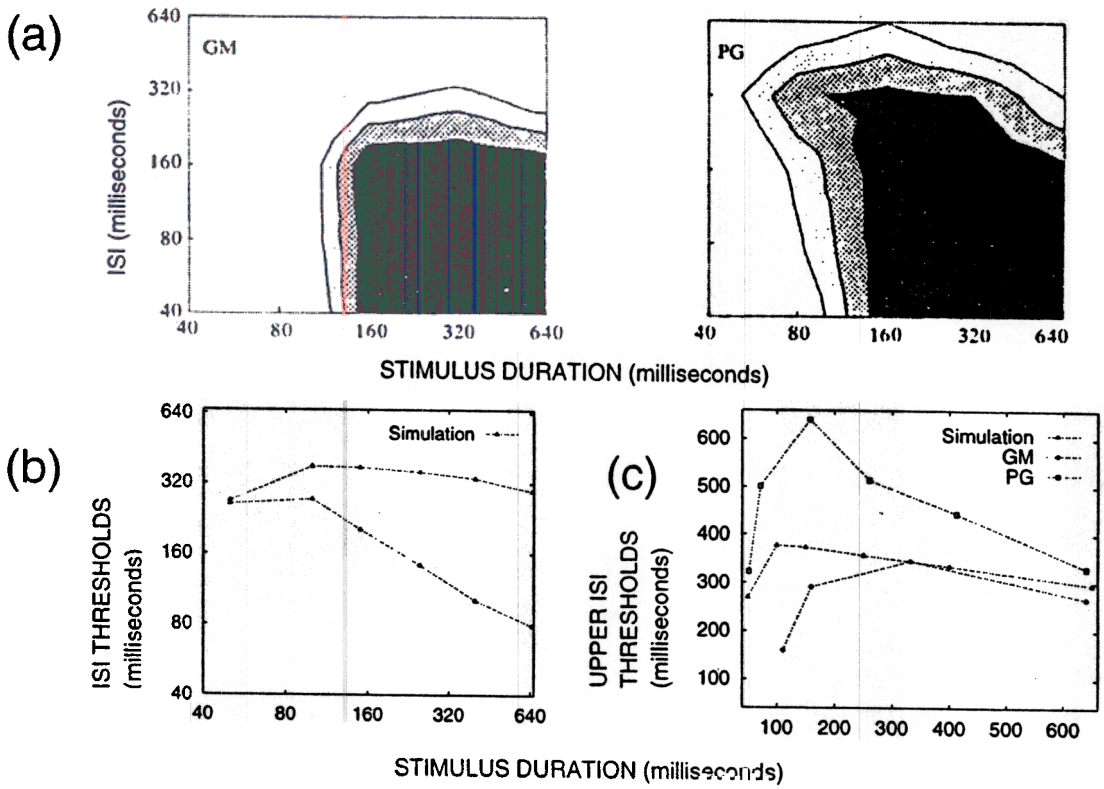
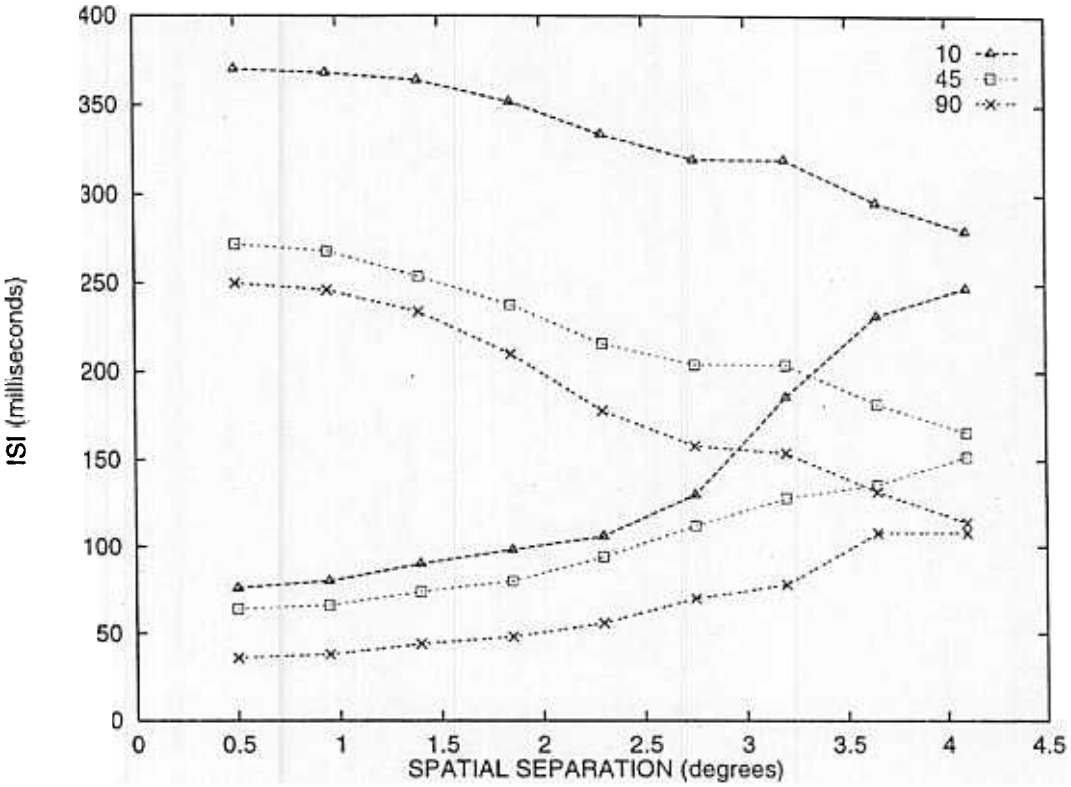

(a)
GM
PG
ISI (milliseconds)
640
320
160
80
40
40
80
160
320
640
STIMULUS DURATION (milliseconds)
(b)
Simulation
ISI THRESHOLDS
(milliseconds)
(c)
Simulation
GM
PG
UPPER ISI
THRESHOLDS
(milliseconds)
600
500
400
300
200
100
100
200
300
400
500
600
STIMULUS DURATION (milliseconds)

Figure 11


ISI (milliseconds)
SPATIAL SEPARATION (degrees)
10
45
90
0
50
100
150
200
250
300
350
400
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5

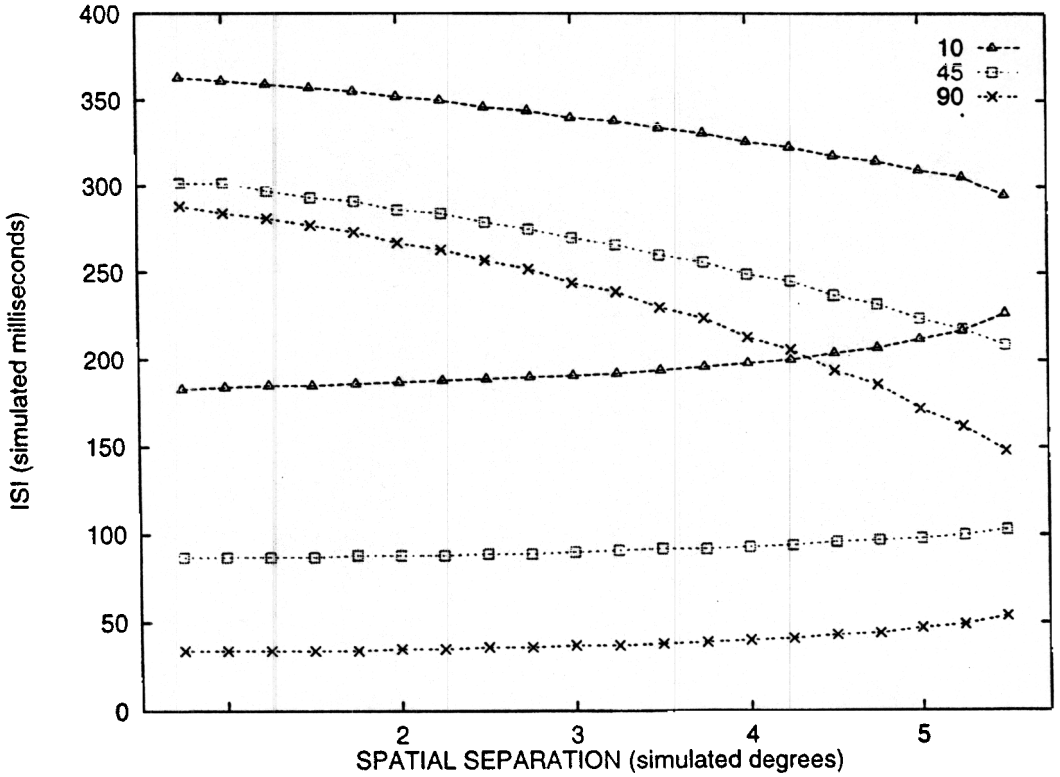

ISI (simulated milliseconds)
SPATIAL SEPARATION (simulated degrees)
10
45
90
0
50
100
150
200
250
300
350
400
2
3
4
5

Figure 12

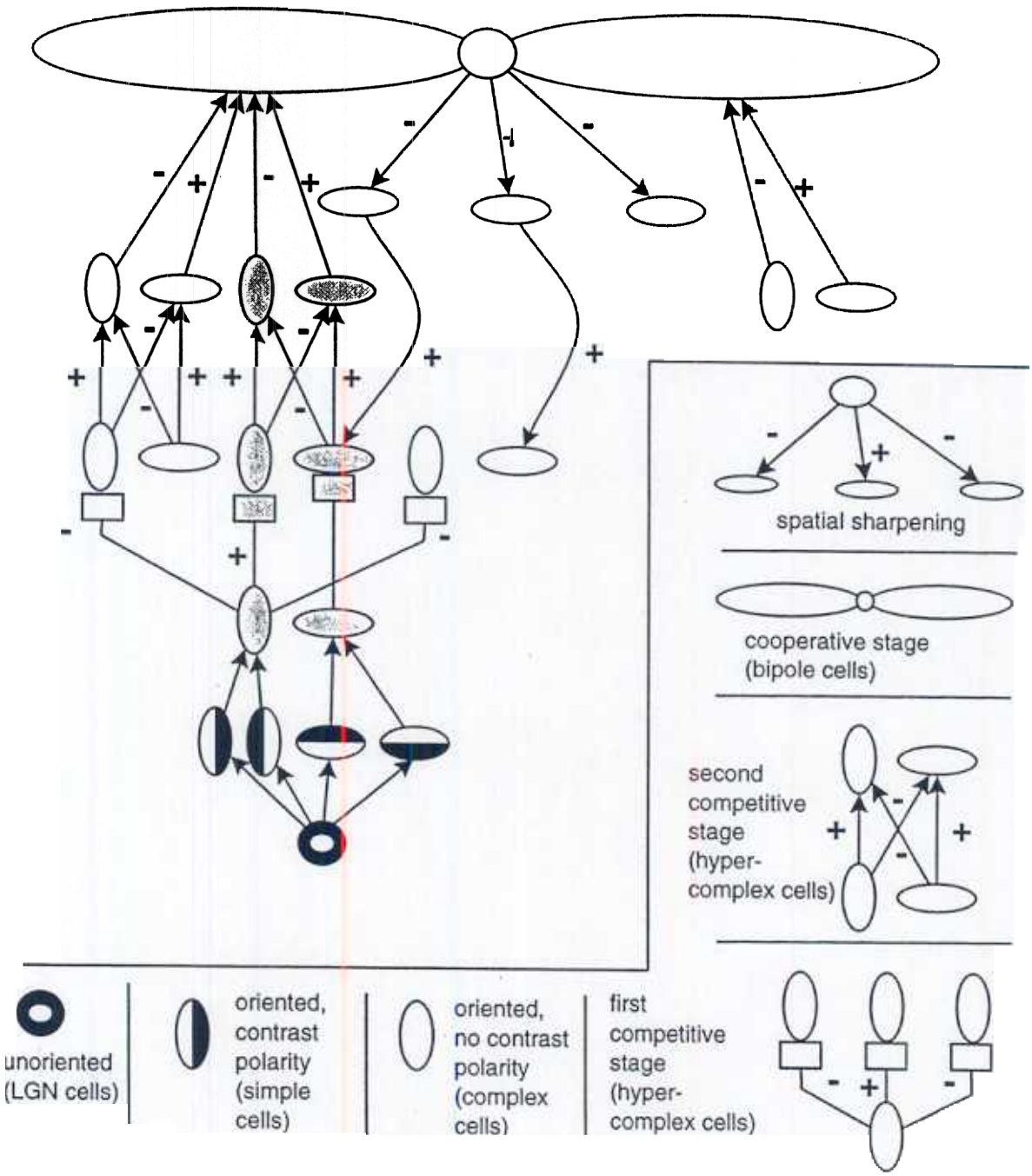

spatial sharpening
cooperative stage
(bipole cells)
second
competitive
stage
(hyper-
complex cells)
unoriented
(LGN cells)
oriented,
contrast
polarity
(simple
cells)
oriented,
no contrast
polarity
(complex
cells)
first
competitive
stage
(hyper-
complex cells)

Figure 13

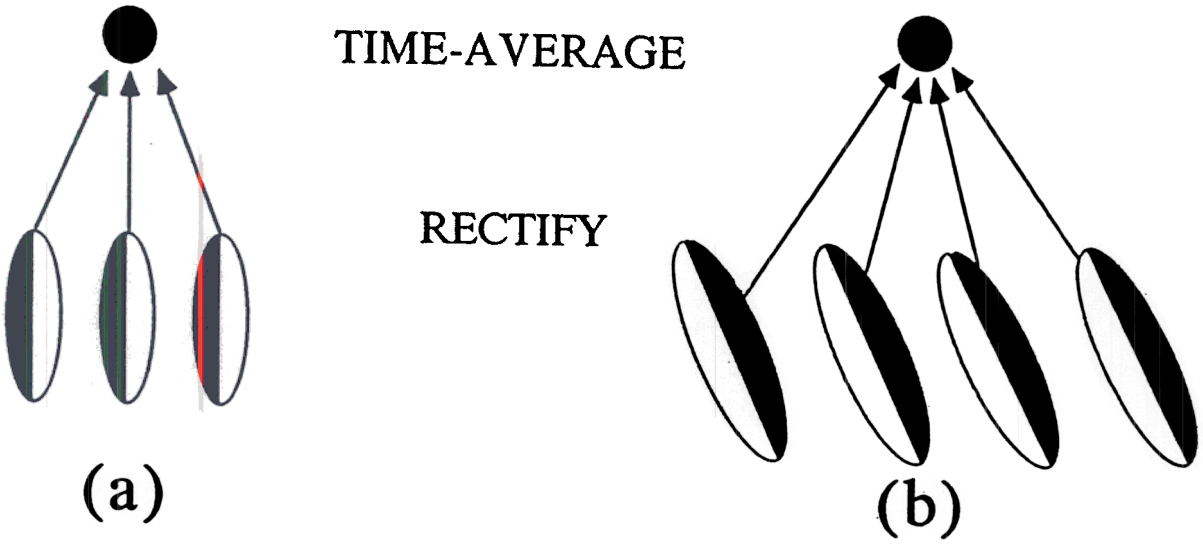


Figure 14 a & b

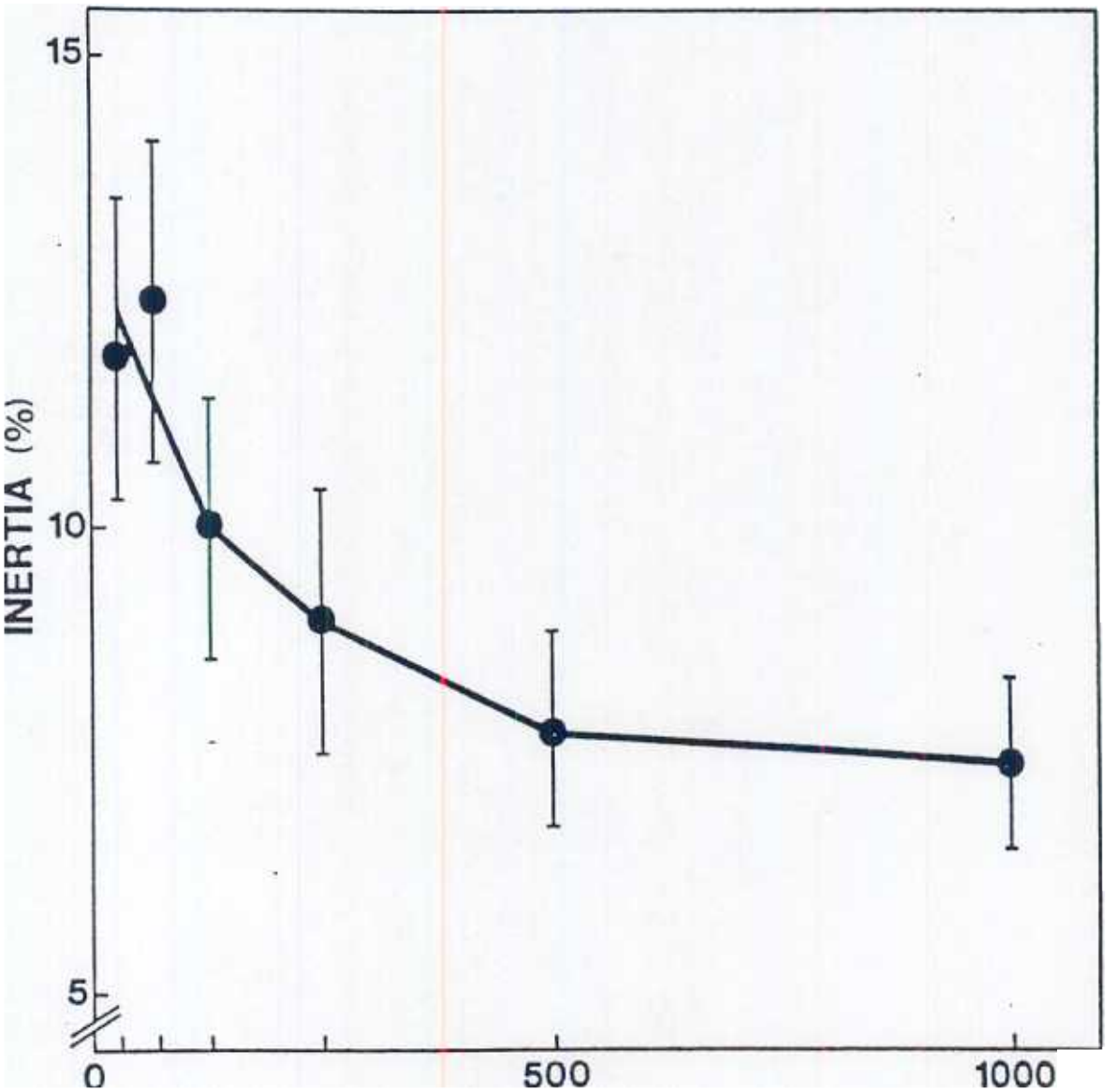


Figure 14 C

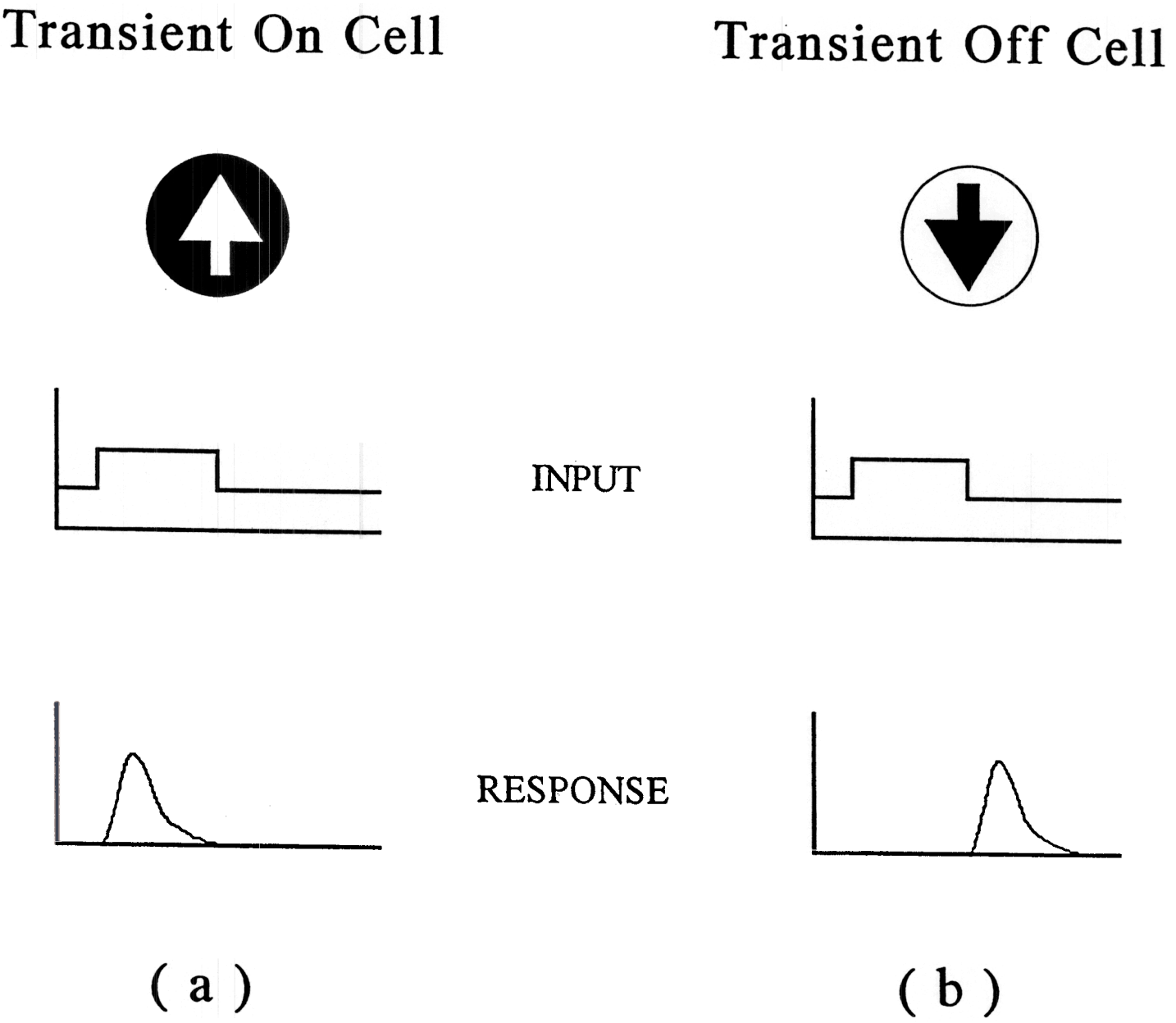


Figure 15

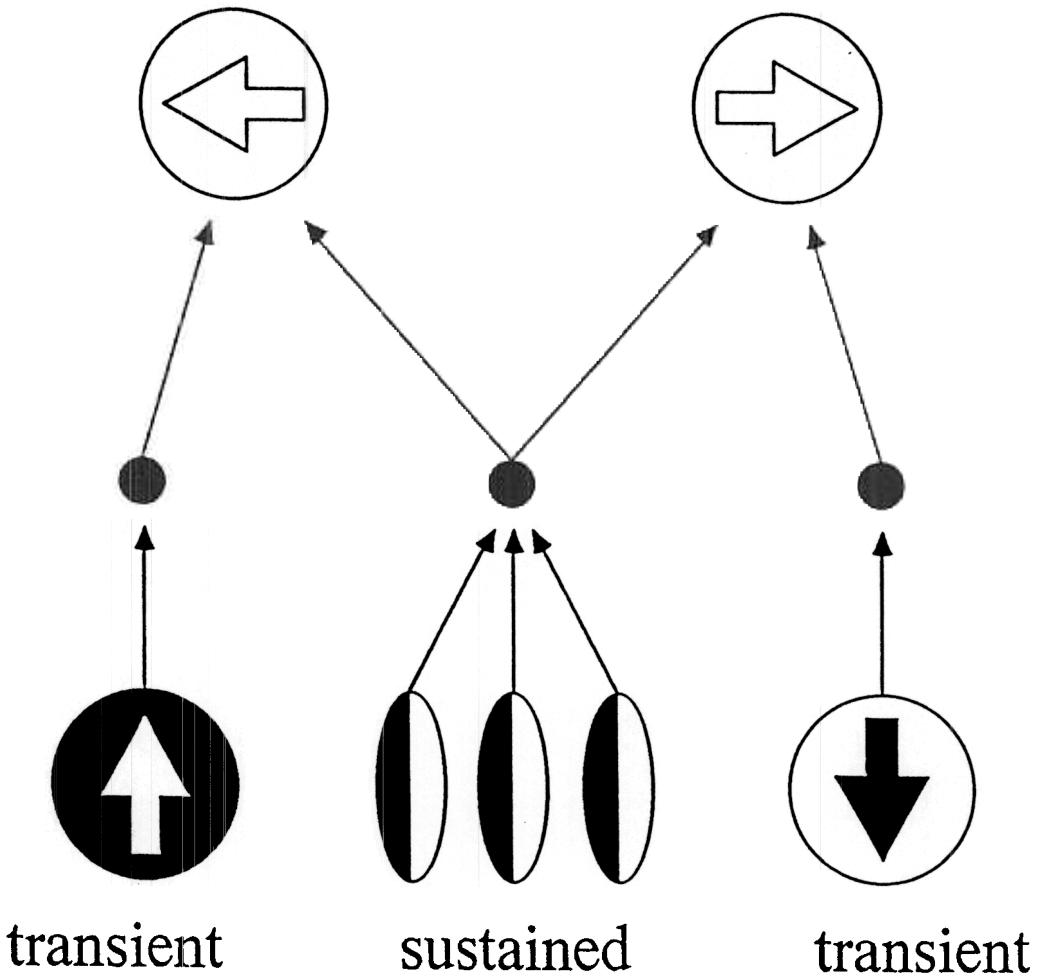


Figure 16

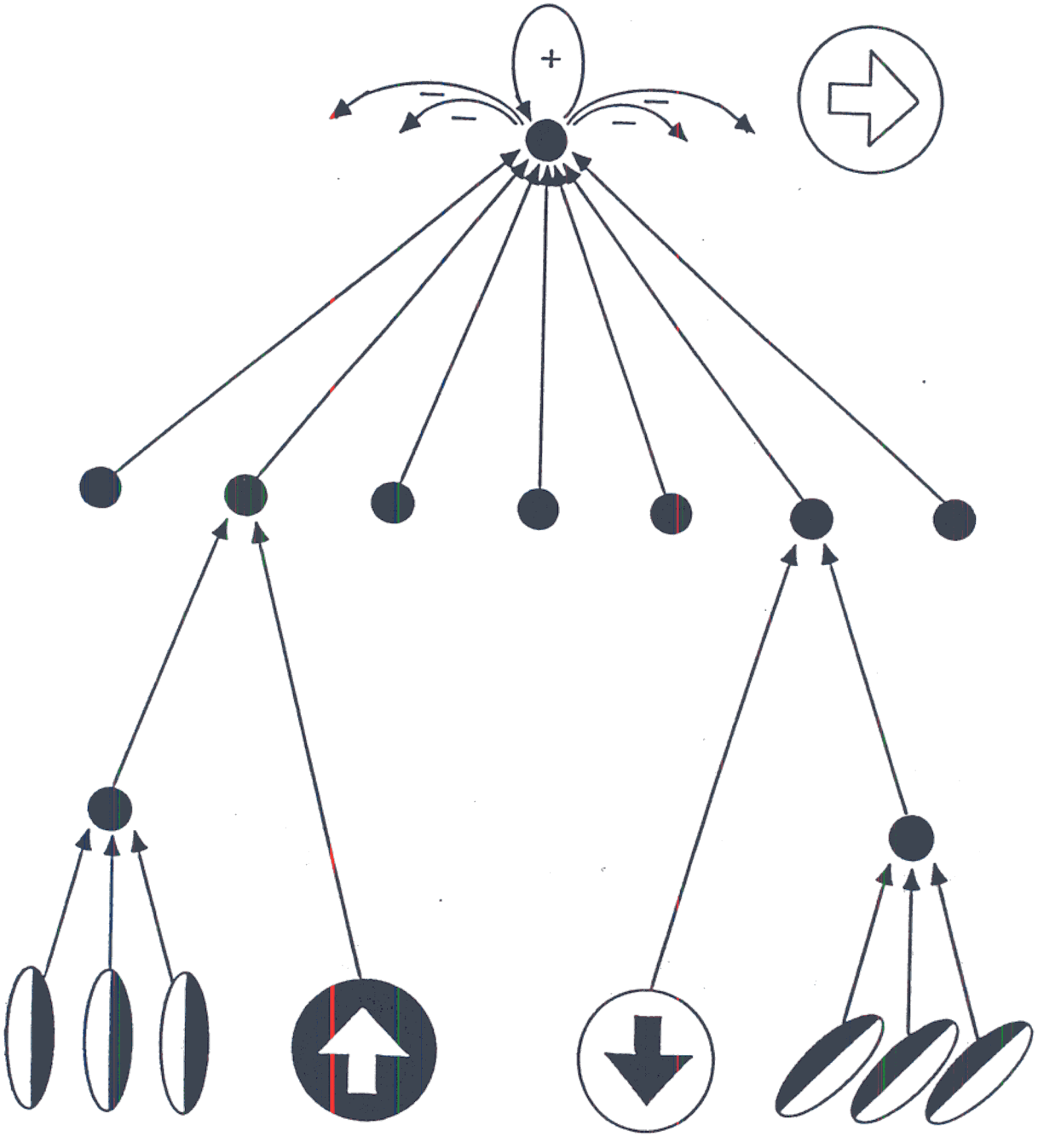
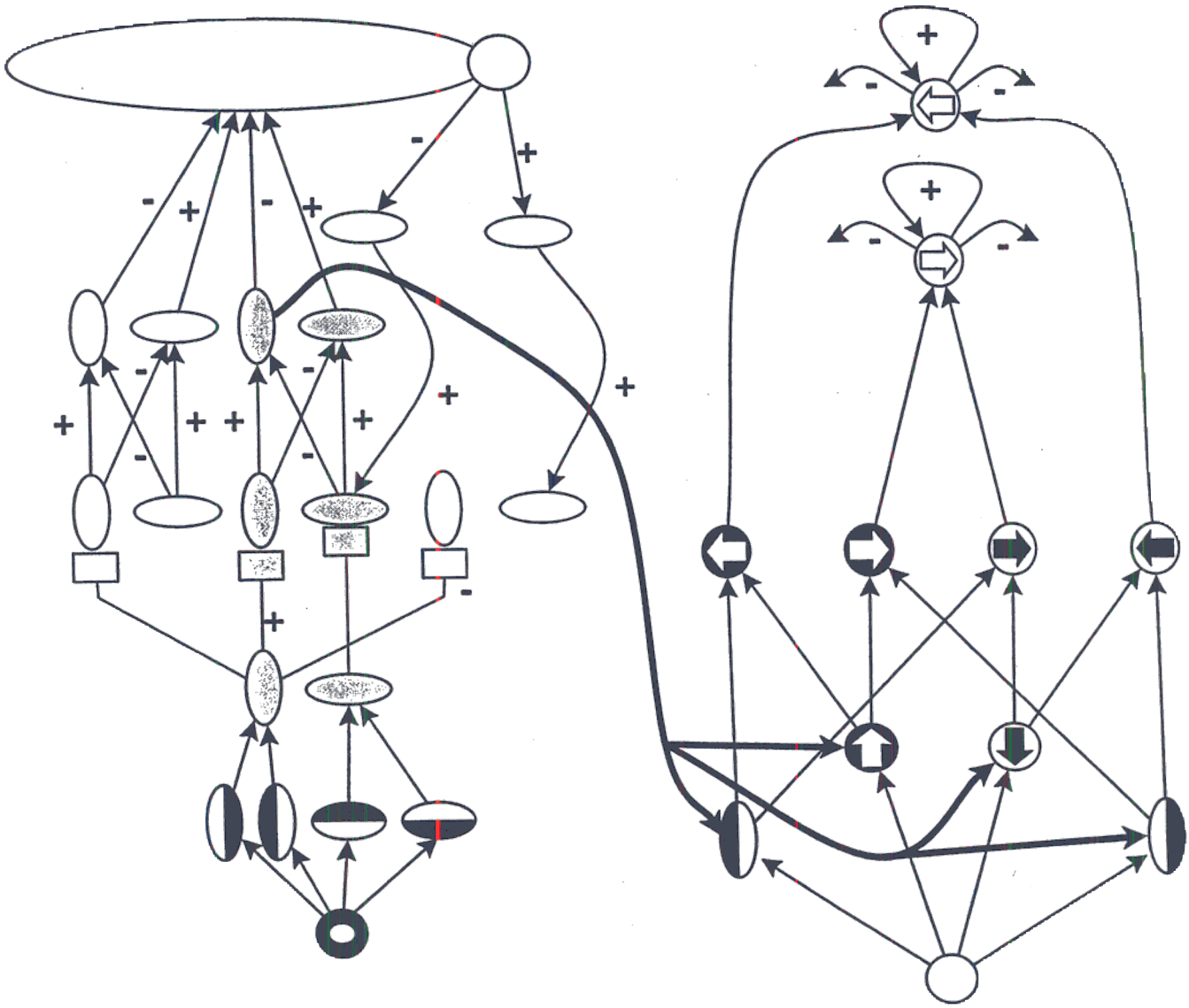


Figure 17



Figure

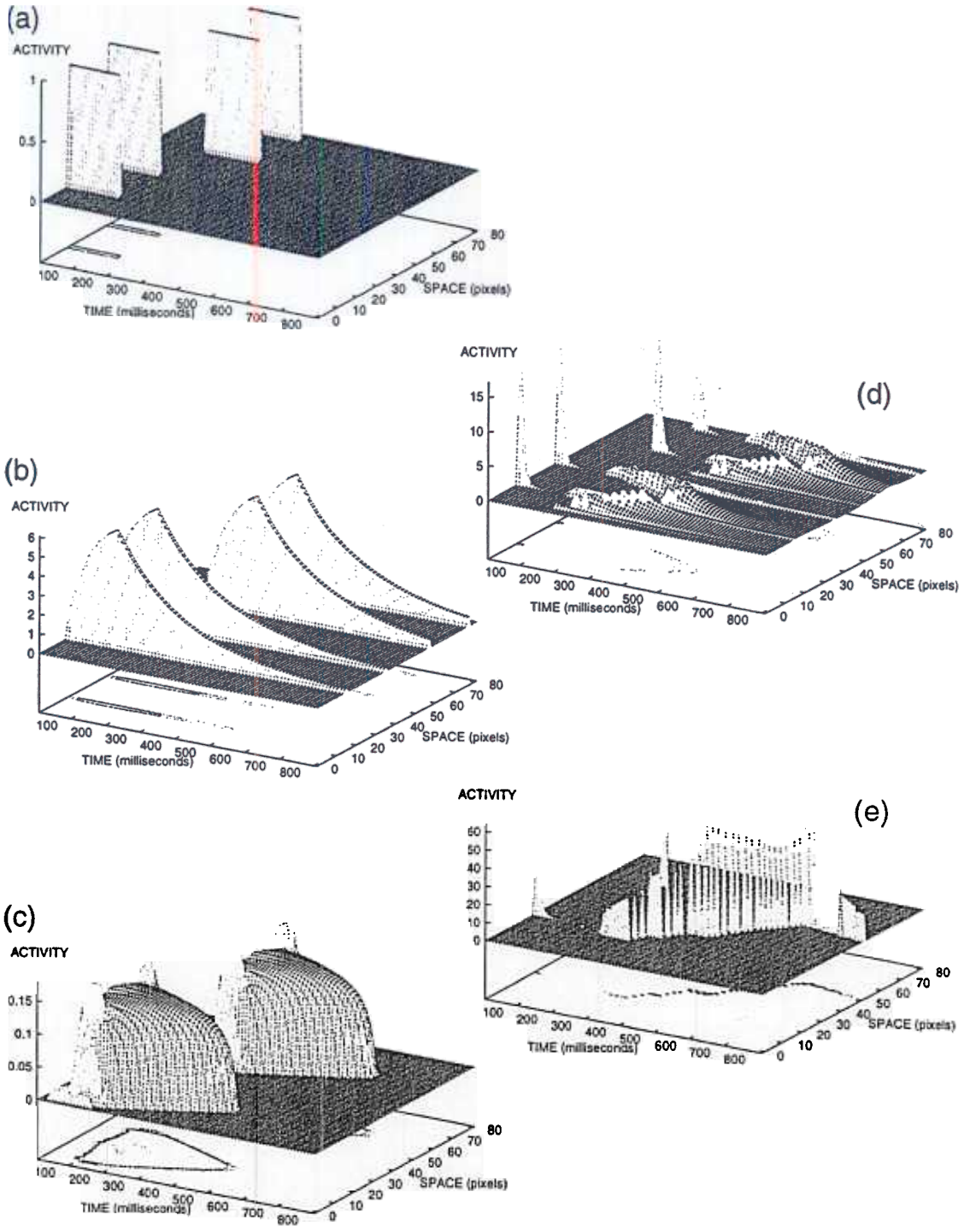

(a)
ACTIVITY
TIME (milliseconds)
SPACE (pixels)
(b)
ACTIVITY
TIME (milliseconds)
SPACE (pixels)
(c)
ACTIVITY
TIME (milliseconds)
SPACE (pixels)
(d)
ACTIVITY
TIME (milliseconds)
SPACE (pixels)
(e)
ACTIVITY
TIME (milliseconds)
SPACE (pixels)

Figure 19

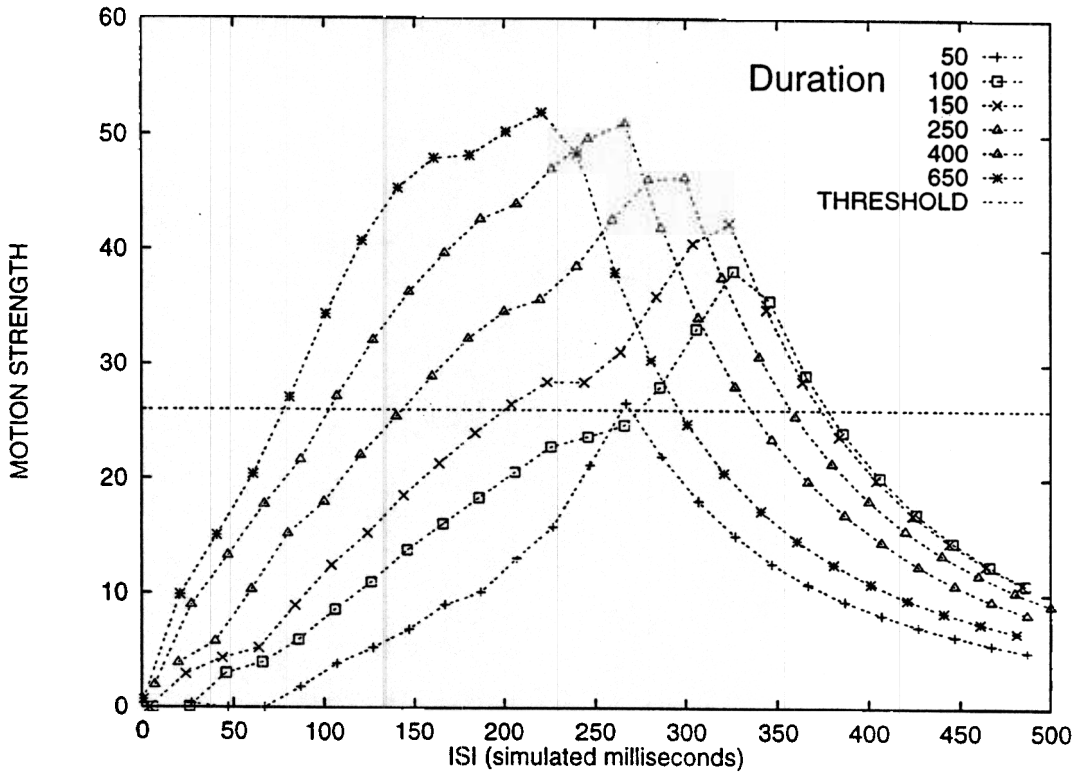

MOTION STRENGTH
ISI (simulated milliseconds)
Duration
50
100
150
250
400
650
THRESHOLD
0
10
20
30
40
50
60
0
50
100
150
200
250
300
350
400
450
500

Figure 20