ORIGINAL CONTRIBUTION

A Neural Architecture for Visual Motion Perception: Group and Element Apparent Motion

STEPHEN GROSSBERG* AND MICHAEL E. RUDD†

Boston University

Abstract—*A neural network model of motion segmentation by visual cortex is described. The model's properties are illustrated by simulating on the computer data concerning group and element apparent motion, including the tendency for group motion to occur at longer ISIs and under conditions of short visual persistence. These phenomena challenge recent vision models because the switch between group and element motion is determined by changing the timing of image displays whose elements flash on and off but do not otherwise move through time. The model clarifies the dependence of short-range and long-range motion on a spatial scale. Its design specifies how sustained response cells and transient response cells cooperate and compete in successive processing stages to generate motion signals that are sensitive to direction-of-motion, yet insensitive to direction-of-contrast. Properties of beta motion, phi motion, gamma motion, and Ternus motion are explained. A number of prior motion models are clarified, transformed, and unified, including the Reichardt model, Marr-Ullman model, Burt-Sperling model, Nakayama-Loomis model, and NADEL model. Apparent motion and real motion generate testably different model properties. The model clarifies how preprocessing of motion signals by a motion OC Filter is joined to long-range cooperative motion mechanisms in a motion CC Loop to control phenomena such as induced motion, motion capture, and motion after effects. The total model system is a motion Boundary Contour System (BCS) that is computed in parallel with the static BCS of Grossberg and Mingolla before both systems cooperate to generate a boundary representation for 3-D visual form perception.*

Keywords—Neural networks, Visual perception, Long-range motion, Short-range motion, Apparent motion, Boundary Contour System, Visual cortex.

1. INTRODUCTION: WHY ARE STATIC AND MOTION BOUNDARY CONTOUR SYSTEMS NEEDED?

This article further develops a neural network model of motion segmentation by visual cortex, called the motion Boundary Contour System (BCS), that was introduced in Grossberg (1987a, section 32). The results developed herein were reported in abbreviated form in Grossberg and Rudd (1989a, 1989b). This model suggests rigorous computational solutions to a number of long-standing problems concerning the design of a motion segmentation system and its functional role in the preattentive representation of three-dimensional form.

One of the most salient problems concerns the very existence of a motion segmentation system. It is well known that some regions of visual cortex are specialized for motion processing, notably region MT (Albright, Desimone, & Gross, 1984; Maunsell & van Essen, 1983; Newsome, Gizzi, & Movshon, 1983; Zeki, 1974a, 1974b). On the other hand, even the earliest stages of visual cortex processing, such as the simple cells in $V1$, require stimuli that change through time for their maximal activation and are direction-sensitive (De Valois, Albrecht, & Thorell, 1982; Heggelund, 1981; Hubel & Wiesel, 1962, 1968, 1977; Tanaka, Lee, & Creutzfeldt, 1983). Why has evolution gone to the trouble to generate regions such as MT, when even $V1$ is change-sensitive and direction-sensitive? What computational properties are achieved by MT that are not already available in $V1$? It is surprising how many plausible answers to this question do not survive a probing computational analysis.

The monocular Boundary Contour System, theory of Grossberg and Mingolla (1985a, 1985b), schematized in Figure 1, and its binocular generalization (Grossberg, 1987b; Grossberg & Marshall, 1989),

Requests for reprints should be sent to Stephen Grossberg, Center for Adaptive Systems, Boston University, 111 Cummington Street, Boston, MA 02215.

BOUNDARY CONTOUR SYSTEM (BCS)

Cooperative-competitive feedback (CC LOOP)

Oriented-contrast filter (OC FILTER)

KEY

bipole

ON
OFF
hypercomplex

hypercomplex

complex to
hypercomplex
(end stopped)

V1 (Area 17)
simple

LGN
concentric
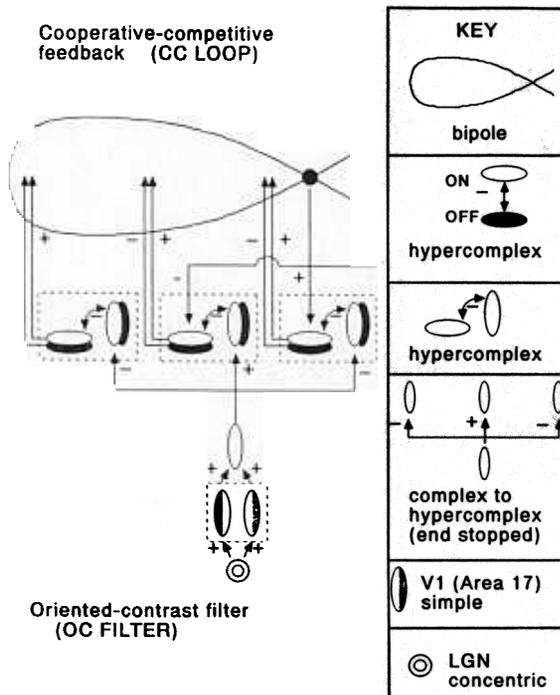
FIGURE 1. The static Boundary Contour circuit described by Grossberg and Mingolla (1985b). The circuit is divided into an oriented contrast-sensitive filter (OC Filter) followed by a cooperative-competitive feedback network (CC Loop). Multiple copies of this circuit are used, one corresponding to each receptive field size of the OC Filter. The depicted circuit has been used to analyze data about monocular vision. A binocular generalization of the circuit has also been described (Grossberg, 1987b; Grossberg & Marshall, 1989).

has successfully modeled many boundary segmentation properties of $V1$ and its prestriate projections. (See Grossberg (1987c, 1988a) for collections of these and related articles). The BCS has thusfar been used to analyze data generated in response to static visual images. Henceforth we therefore call such a BCS a static BCS model. Nonetheless, its model cells can easily be gated by cells sensitive to image transients, such as Y cells (Enroth-Cugell & Robson, 1966; Hoffmann, 1973; Sekuler, 1975; Stone, 1972; Stone & Dreher, 1973; Tolhurst, 1973), to generate receptive fields sensitive to image transients. How does a motion BCS differ from a static BCS whose cells are sensitive to image transients?

2. JOINING SENSITIVITY TO DIRECTION-OF-MOTION WITH INSENSITIVITY TO DIRECTION-OF-CONTRAST

The static BCS consists of two major subdivisions: an oriented-sensitive filter, called the OC Filter, and a cooperative-competitive feedback network, called the CC Loop (Figure 1). The OC Filter models the simple cells and complex cells of $V1$ (Figure 2). Its

WHY IS A MOTION BCS NEEDED?

Simple cells in V1 are already sensitive to change and direction-of-motion. What does MT add?
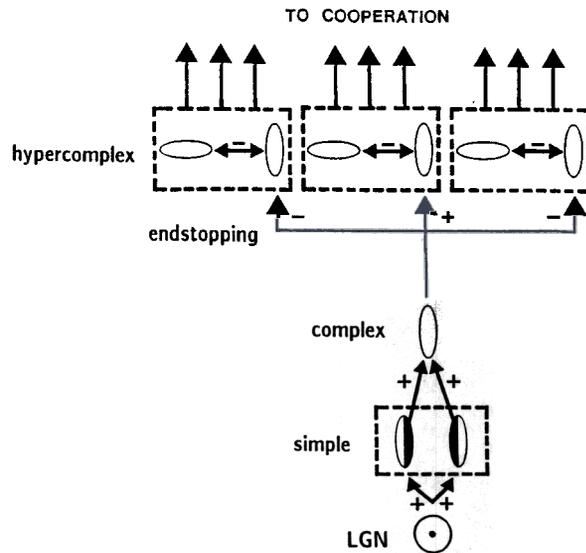
STATIC ORIENTED CONTRAST (OC) FILTER

Insensitive to Direction-of-Contrast

Insensitive to Direction-of-Motion

FIGURE 2. Stages of the OC Filter: Simple cell receptive fields are oriented and sensitive to direction-of-contrast. Rectified outputs from pairs of simple cells sensitive to opposite directions-of-contrast input to complex cells, which are sensitive to amount-of-contrast, but not to direction-of-contrast or to direction-of-motion. Complex cells input to two successive stages of hypercomplex cells. At the first competitive stage, a complex cell excites like-oriented hypercomplex cells at its position and inhibits, via an endstopping operation, hypercomplex cells at nearby positions. At the second competitive stage, hypercomplex cells sensitive to different orientations compete with one another. Inhibition is maximal between perpendicular orientations.

projections to hypercomplex cells form the interface of the OC Filter with the CC Loop. The hypercomplex cells project, in turn, to a cell type called cooperative bipole cells by Grossberg and Mingolla. The bipole cells interact with the hypercomplex cells via the CC Loop.

The OC Filter is a nonlinear filter that multiplexes several different types of image information into a spatially organized representation, or map, across the network of hypercomplex cells. Such a map functions as a compressed code that is capable of reacting selectively to prescribed combinations of image features.

The CC Loop reacts to this multiplexes spatial map by transforming and amplifying those spatial combinations of cell activations whose coded information is mutually consistent, and actively suppress-

ing the rest. The result combines information about image edges, texture, shading, depth, and spatial scale into a coherent, context-sensitive boundary representation.

As shown in Figure 2, although the simple cells of the BCS are sensitive to direction-of-contrast, or contrast polarity, the complex cells of the BCS are rendered insensitive to direction-of-contrast by receiving inputs from pairs of simple cells with opposite direction-of-contrast. Such a property is also true of the simple cells and complex cells in area $V1$ (DeValois, Albrecht, & Thorell, 1982; Poggio, Motter, Squatrito, & Trotter, 1985; Thorell, DeValois, & Albrecht, 1984).

This property is useful for extracting boundary structure that is independent of illumination fluctuations, such as shadows. As a result, the output of the OC Filter is unable to differentiate direction-of-motion. A key property of the motion BCS model presented here is that it possesses a modified OC Filter that multiplexes the property of insensitivity to direction-of-contrast, which is equally useful for the processing of static and moving forms, with sensitivity to direction-of-motion (Grossberg, 1987a). The properties of this motion OC Filter enable us to explain many properties of motion perception, including percepts of apparent motion. When the motion OC Filter is connected to the CC Loop, a much larger body of data, including coherent global motion percepts such as induced motion and motion capture, can also be analysed.

Thus the present article suggests that a fundamental computational property achieved by a motion segmentation system, such as MT, is to generate output signals that maintain insensitivity to direction-of-contrast without sacrificing sensitivity to direction-of-motion. This modification of the static OC Filter enables us to define a motion BCS that is useful to analyze a large body of data concerning motion segmentation. The parsimony of this result provides additional support for both the static BCS model and the motion BCS model by showing that both models may be considered variations on a single neural architectural theme.

3. APPARENT MOTION AS A PROBE OF MOTION MECHANISMS

This conception of the motion BCS was derived from an analysis of motion-related data taken from psychophysics, visual perception, and neurophysiology. Herein some challenging data about apparent motion are analyzed. Apparent motion is a particularly useful probe of motion mechanisms because it describes controllable experimental situations in which nothing moves, yet a compelling percept of motion is generated.

For example, two brief flashes of light, separated in both time and space, create an illusion of movement from the location of the first flash to that of the second when the spatiotemporal parameters of the display are within the correct range. Since this phenomenon—labeled apparent motion by the Gestalt psychologists—was first discovered by Exner in 1879, it has been the object of intensive scientific experimentation. Variants of apparent motion include *phi motion*, or the *phi phenomenon*, whereby a "figureless" or "objectless" motion signal propagates from one flash to the other, analogous to the rapid motion of an object so quickly that its form cannot be clearly identified; *beta motion*, whereby a well-defined form seems to move smoothly and continuously from one flash to the other; and *gamma motion*, the apparent expansion at onset and contraction at offset of a single flash of light (Bartley, 1941; Kolers, 1972). For some Gestalt psychologists who popularized the study of apparent motion in the early 20th century, the observed mental construction of a non-veridical motion path epitomized the constructive aspects of the human mind. Their attempts to construct a model of the brain mechanisms underlying the phenomenon were, however, premature.

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

Due to the difficulty of such problems, the construction of a satisfactory neural model of apparent motion—and the more general category of motion perception phenomena—remains a challenging problem to the present time. A large psychophysical literature exists to indicate complex interdependencies between such physical variables as stimulus contrast, size, luminance, duration, color, and figural organization in determining the perception of movement. In addition, it is clear that the neural networks

which compute motion are concerned with the simultaneous extraction of other information from the visual stimulus as well. The real problem is to discover, then, not just how the brain computes motion, but rather how these computations are embedded in a more general process of generating a 3-D representation of moving objects.

In the present paper we summarize several psychophysical results which illustrate the complex interrelationship between stimulus factors influencing the organization of motion percepts. The motion BCS is then defined and shown capable of explaining these percepts.

4. SPATIOTEMPORAL PARAMETERS FOR GENERATING APPARENT MOTION

In order to produce apparent motion, the spatial and temporal organization of the stimulus must be appropriate (Kolers, 1972). In the simple case of two successive flashes (Figure 3a), there is a correlation

TWO FLASH DISPLAY

Frame 1

Frame 2

Small ISI: Stationarity

Intermediate ISI: Motion

Large ISI: Motion

(a)

TERNUS DISPLAY

Frame 1

Frame 2

Small ISI: Stationarity

Intermediate ISI: Element Motion

Large ISI: Group Motion

(b)
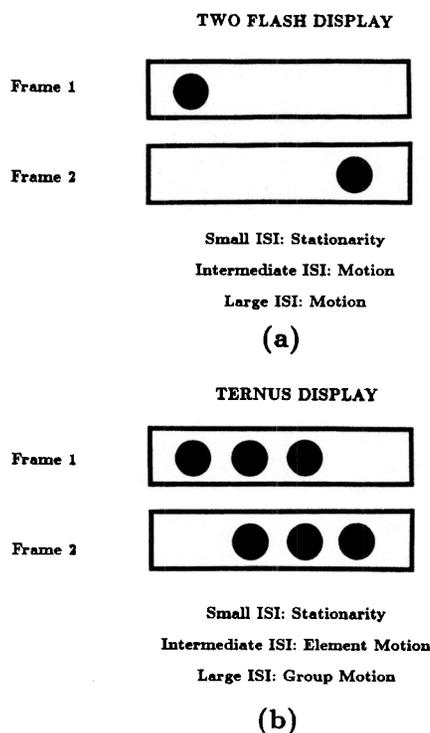
FIGURE 3. Two types of apparent motion displays in which the two frames outline the same region in space into which the dots are flashed at successive times: In (a), a single dot is flashed, followed by an interstimulus interval (ISI), followed by a second dot. At small ISIs, the two dots appear to flicker in place. At longer ISIs, motion from the position of the first dot to that of the second is perceived. (b) In the Ternus display, three dots are presented in each frame such that two of the dots in each frame occupy the same positions. At short ISIs, all the dots appear to be stationary. At longer ISIs the dots at the shared positions appear to be stationary, while apparent motion occurs from the left dot in Frame 1 to the right dot in Frame 2. At still longer ISIs, the three dots appear to move from Frame 1 to Frame 2 as a group.

between the spatial separation of the flashes and the range of the temporal separations of the flashes which give rise to apparent movement. The temporal separation of the two flashes is often defined in two different ways. The stimulus onset asynchrony (or SOA) is a measure of the time between the onsets of the two successive flashes. The interstimulus interval (or ISI) is the time between the offset of the first flash and the onset of the second flash. If we use the symbol SD to indicate the temporal duration of the stimulus, we then have $SOA = SD + ISI$. These terms retain their meaning if the cycle of stimulus flashes is repeated many times, as is often done experimentally. In this case, there may be two ISIs as well as two SDs. In this case we can designate the durations of the two stimuli as SD_1 and SD_2; and the interstimulus interval between the offset of the first stimulus (hereafter S_1) and the onset of the second stimulus (hereafter S_2) as ISI_{12}. Thus, we have $SOA_{12} = SD_1 + ISI_{12}$. Similarly, $SOA_{21} = SD_2 + ISI_{21}$.

In what follows we assume that $SD_1 = SD_2$ and $ISI_{12} = ISI_{21}$ for a continuously cycling display, so we hereafter drop the indices on the symbols SD, ISI, and SOA. Assume for the moment that the stimulus duration of the flashes is fixed for any particular spatial separation. There will be a range of SOAs in which apparent movement is observed along a path connecting the two flashes. At SOAs shorter than the minimum SOA for apparent movement, subjects report that the two flashes flicker in place, with no corresponding sensation of movement between the locations of the two flashes. At SOAs longer than the maximum SOA for apparent movement, the subjects report that the stimuli appear in succession in their respective locations, again with no corresponding appearance of continuous movement from one location to another.

As the spatial separation of the two flashes is increased, the range of SOAs producing perceived movement is decreased, although the SOA corresponding to the midrange remains roughly constant (Burt & Sperling, 1981; Kolers, 1972). If the spatial separation is made sufficiently large, the probability of producing apparent motion at all becomes small.

5. GROUP AND ELEMENT APPARENT MOTION: TERNUS DISPLAYS

A well-known apparent motion display; originally due to Ternus (1926/1950), illustrates the fact that not only the existence of a motion percept, but also its figural identity, may depend on subtle aspects of the display, such as the ISI (Figure 3b). In the Ternus display, a cyclic alternation of two stimulus frames gives rise to competing visual movement percepts. In Frame 1, three black elements are arranged in a

horizontal row on a white background. (The contrast may be reversed without consequence to the discussion which follows.) In Frame 2, the elements are shifted to the right in such a way that the positions of the two leftwardmost elements in Frame 2 are identical to those of the two rightwardmost elements in Frame 1. Depending on the stimulus conditions, the observer will see either of two bistable motion percepts. Either the elements will appear to move to the right as a group between Frames 1 and 2 and then back again during the second half of a cycle of the display or, alternatively, the leftwardmost element in Frame 1 will appear to move to the location of the rightwardmost element in Frame 2, jumping across two intermediate elements which appear to remain stationary. We will refer to the first percept as "group" motion; and the second percept as "element" motion. At short ISIs there is a tendency to observe element motion. At longer ISIs, there is a tendency to observe group motion.

A number of stimulus conditions have been demonstrated that affect the transitional ISI at which the perception of element motion gives way to the perception of group motion. Stated in another way, there exists a range of ISIs for which the figural organization of the perceived motion is determined by the stimulus conditions.

For example, Breitmeyer and Ritter (1986) found that the ISI at which the transitions from element motion to group motion occurs, is a decreasing function of viewing eccentricity, element size, and frame duration. Petersik and Pantle (1979) found that the percentage of group motion responses was an increasing function of frame duration and interframe interval luminance, in addition to the interstimulus interval. They also found that the percentage of group motion responses generally decreased with increasing element background contrast (but there was some crossover; see Petersik and Pantle's Figure 5). In addition, the ISI at which the transition between group and element motion percepts occurred was found to be an increasing function of dark adaptation.

The finding of Breitmeyer and Ritter (1986) concerning the tradeoff between the frame duration and ISI at which the group versus element motion percept transition occurs is consistent with the Petersik and Pantle (1979) discovery that long ISIs and long frame durations each lead to a greater percentage of group movement percepts. Thus, the two factors can be traded off against one another.

It has been suggested that the perception of element motion is a result of the perceived stationarity of the two middle elements (Braddick, 1980; Braddick & Adlard, 1978; Breitmeyer & Ritter, 1986; Pantle & Petersik, 1980). Breitmeyer and Ritter pointed out that the stimulus conditions in their ex-

periment that favored the perception of element motion are also conditions which favor visual persistence and thus the perceived stationarity of the inner elements of the ternus display.

6. THE INSENSITIVITY OF MOTION CORRESPONDENCE TO FIGURAL IDENTITY

It is well-documented that apparent motion percepts depend very little on the figural identity of the elements being matched from frame to frame. Thus motion perception does not utilize a form-dependent matching process. Early in the history of apparent motion research, it was discovered that the motion system was capable of computing motion between locations occupied by shapes which varied greatly between frames (Higginson, 1926; Orlansky, 1940; Wertheimer, 1912). Depending on the shapes presented in each frame and the timing of the display, the interframe shape disparity was resolved by the motion system in a variety of ways. In the case of brief flash durations and relatively short ISIs, the disparity might be resolved by an objectless phi motion. In the case of longer flash durations and ISIs, shapes were observed to continuously deform into one another between frames (van der Waals and Roelofs, 1930, 1931), or even to be transformed in depth (Neuhaus, 1930).

More recent experiments (Burt & Sperling; 1981, Kolers, 1972) have confirmed the relatively unimportance of figural identity in the determination of frame-to-frame element matches in apparent motion. Kolers and Pomerantz (1971) presented pairs of frames, each containing a simple geometrical shape such as an arrow, circle, triangle, or square, and studied the simultaneous effects of manipulating ISI and shape dissimilarity on the probability of perceived motion. Shape dissimilarity was found to account for only 1%–3% of the statistical effect.

After reviewing the literature concerning the relationship between figural identity and apparent motion, Kolers (1972) concluded that there appear to be two parallel subsystems in the human visual system for the computation of motion and the maintenance of figural identity. This view is consistent with recent physiological findings which indicate the presence of parallel visual pathways for motion perception and static form perception, and is clarified by our results concerning the design of parallel static BCS and motion BCS architectures.

7. SPATIAL AND TEMPORAL FORM FACTORS FOR APPARENT MOTION

It is clear from the fact that there is a restricted range of SOAs (or, alternatively, ISIs) which give rise to

apparent motion, that there must be a function which defines the probability of perceived motion as a function of the SOA (or ISI). In this article, this function is called the *temporal form factor* for apparent motion.

Similarly, the probability of producing apparent motion falls off as a function of the spatial separation of the flashes at any fixed temporal separation. Thus, there is also a *spatial form factor* for apparent motion.

In an important article combining experimental with theoretical analysis, Burt and Sperling (1981) demonstrated that these spatial and temporal form factors for apparent motion are independent of one another, at least to first approximation. That is, the probability of perceived motion at a particular ISI and spatial separation pair is given by the product of the probability for seeing motion at that ISI times the probability for seeing motion at that spatial separation.

This result was arrived at by creating displays which gave rise to alternative (and exclusive) apparent motion paths. Thus the various potential perceived motion paths were paired against one another. Particular paths were found to dominate as a function of ISI and element spatial separation. Burt and Sperling found the transition points at which the probabilities of motion along two different potential motion paths were equal. Their critical experimental finding was that these transition points were approximately independent of the observer's viewing distance from the display. Thus, the transition points were found to be approximately *scale invariant*.

From this finding of scale invariance, Burt and Sperling were able to show that the probability of perceived motion along a given path is a separable function of the spatial and temporal distances between the elements along that path. They further concluded that the temporal form factor is approximately of the form

$$S(t) = t^{\alpha\beta}e^{-\beta t}, \tag{1}$$

where $\alpha, \beta > 0$. According to this theory, the probability of perceived movement goes to zero for very short as well as very long ISIs; and is maximal at the temporal separation α.

Burt and Sperling also concluded that the spatial form factor must go to zero as the spatial separation of the display element becomes large, and for small spatial separation this form factor must approach some finite limit. Burt and Sperling defined the spatial form factor, as a function of distance d, by

$$S(d) = d^{-1}\exp(-\gamma d^{-1}). \tag{2}$$

Using (1) and (2), they defined the space-time form factor by

$$S(d, t) = S(t)S(d); \tag{3}$$

Insensitive to Direction-of-Contrast

Sensitive to Direction-of-Motion
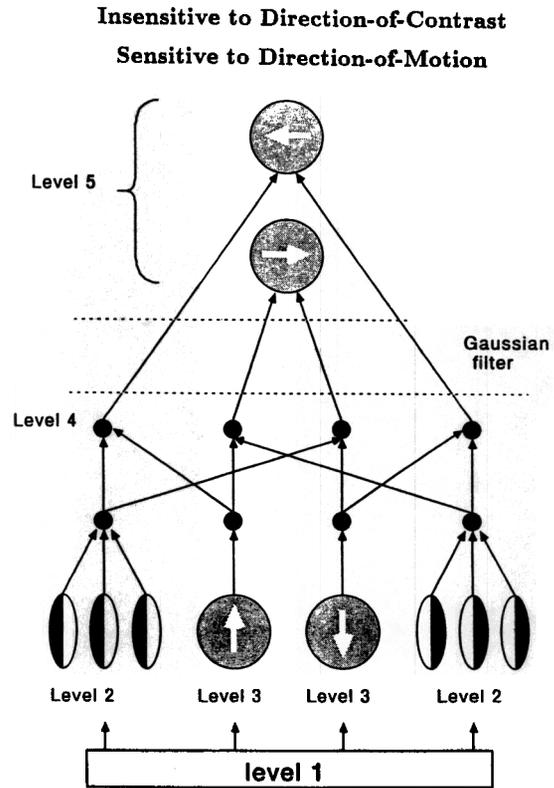
FIGURE 4. The motion OC Filter: Level 1 registers the input pattern. Level 2 consists of sustained response cells with oriented receptive fields that are sensitive to direction-of-contrast. Level 3 consists of transient response cells with unoriented receptive fields that are sensitive to direction-of-change in the total cell input. Level 4 cells combine sustained cell and transient cell signals to become sensitive to direction-of-motion and sensitive to direction-of-contrast. Level 5 cells combine Level 4 cells to become sensitive to direction-of-motion and insensitive to direction-of-contrast.

TABLE 1
Levels of Motion OC Filter

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

that is,

$$S(d, t) = t^{\alpha\beta}e^{-\beta t - \gamma d^{-1}}d^{-1}. \tag{4}$$

They noted that the "correction factor, $\exp(-\gamma d^{-1})$, . . . has little effect at large d but . . . forces s to 0 at $d = 0$. In fact, we tried several different correction factors as we found that the predictions were not sensitive to the form of the correction factor. What matters is that corrected $S(d, t)$ approaches a reasonable limit, not infinity, as d goes to zero" (Burt & Sperling, 1981, p. 187). In addition, Burt and Sperling showed that possible motion paths compete, and that the path with greatest strength determines the motion percept. The motion BCS provides an explanation of a separable space-time form factor that satisfies these constraints.

A single-scale, 1-D version of the motion OC Filter that is sufficient for present purposes is next defined, and summarized in Figure 4 and Table 1.

8. EQUATIONS FOR A MOTION OC FILTER

Level 1: Stimulus Representation

In the 1-D theory, let I_i denote the luminance of the input at position i.

Level 2: Oriented Sustained-Response Filter

In the 2-D theory, spatially aligned arrays of cells are defined with like-oriented receptive fields that are sensitive to the same direction-of-contrast (Figure 4). In the 1-D theory, only horizontal motions are considered. It therefore suffices to consider two types of such cells, one which responds to a light–dark luminance contrast (designated by L, for left) and the other which responds to a dark–light luminance contrast (designated by R, for right), that filter the input pattern I_i. Output pathways from like cells converge (see Figure 4) to generate inputs J_{iL} and J_{iR} at each position i. The activity x_{ik} of the ith target cell obeys a membrane equation

$$\frac{d}{dt}x_{ik} = -Ax_{ik} + (1 - Bx_{ik})J_{ik}, \tag{5}$$

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

Level 3: Unoriented Transient-Response Filter

The outputs of the oriented sustained-response cells are gated by the responses of cells with unoriented receptive fields that are sensitive to transient increments and decrements, respectively, in the input pattern. These transient-response cells may be modeled in a simple way as time derivatives of an unoriented

space-time average x_i of the input pattern I_i, as in

$$\frac{d}{dt}x_i = -Cx_i + (D - Ex_i)\sum_j I_j F_{ji}, \tag{6}$$

where F_{ji} is an unoriented spatial kernel. Then let

$$y_i^+ = \max\left(\frac{d}{dt}x_i - \Gamma, 0\right) \tag{7}$$

and

$$y_i^- = \max\left(\Omega - \frac{d}{dt}x_i, 0\right), \tag{8}$$

where Γ and Ω are constant thresholds. Activity y_i^+ models the response of a transient on-cell, and activity y_i^- models the response of a transient off-cell. This approximation may be improved by a neurally more plausible model in which the time derivative operation is replaced by a feedforward inhibitory interneuron (Grossberg, 1970).

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

Sustained output signals from Level 2 are gated by transient output signals from Level 3 in a manner that renders cells at Level 4 sensitive to direction-of-motion, but also sensitive to direction-of-contrast. Two types of cells are sensitive to local rightward motion: the $(L, +)$ cells that respond to $x_{iL}y_i^+$ and the $(R, -)$ cells that respond to $x_{iR}y_i^-$. Two types of cells are sensitive to local leftward motion: the $(L, -)$ cells that respond to $x_{iL}y_i^-$ and the $(R, +)$ cells that respond to $x_{iR}y_i^+$. All of these cells inherit the sensitivity to direction-of-contrast of their inputs from Level 2 sustained-response cells.

Level 5 cells are also sensitive to direction-of-motion, but are rendered insensitive to direction-of-contrast as a result of a receptive field organization which combines the responses of Level 4 cell types that signal same direction-of-motion but opposite direction-of-contrast. This interaction is the analog within the motion OC Filter of the interaction within the static OC Filter whereby simple cell outputs are combined to form complex cell inputs (Figure 2).

Level 5: Spatial Filtering and Competition

To define this interaction, let

$$r_i = x_{iL}y_i^+ + x_{iR}y_i^- \tag{9}$$

and

$$l_i = x_{iL}y_i^- + x_{iR}y_i^+ \tag{10}$$

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

light contrast moves to the right. Signal l_i increases if either a light–dark or dark–light contrast moves to the left.

These local motion signals are filtered by a long-range Gaussian function

$$G_{ji} = H \exp[-(j - i)^2/2K^2]$$ (11)

on their way to Level 5. The total input signal to the right-motion detector at position i of Level 5 is

$$R_i = \sum_j r_j G_{ji}.$$ (12)

The corresponding left-motion input is

$$L_i = \sum_j l_j G_{ji}.$$ (13)

Contrast-enhancing competitive, or lateral inhibitory, interactions within Level 5 generate the activations at this level that encode motion information. A contrast-enhancing competitive interaction has also been modeled at the complex cell level of the static OC Filter (Grossberg, 1987b; Grossberg & Marshall, 1989). The simplest assumption is that the competition is tuned to select that population whose input is maximal (Grossberg, 1973, 1982, 1988b), as in

$$x_i^{(R)} = \begin{cases} 1 & \text{if } R_i > R_j, j \neq i \\ 0 & \text{otherwise,} \end{cases}$$ (14)

and

$$x_i^{(L)} = \begin{cases} 1 & \text{if } L_i > L_j, j \neq i \\ 0 & \text{otherwise.} \end{cases}$$ (15)

This assumption was made for simplicity in our computer simulations. The functions $x_i^{(R)}$ and $x_i^{(L)}$ change through time in a manner that idealizes parametric properties of the apparent motion phenomena schematized in Figure 3. More generally, we suggest that this competitive process partially contrast-enhances its input pattern and thereby generates a motion signal whose breadth across space increases with the breadth of its inducing input pattern.

9. INTUITIVE EXPLANATION OF CONTINUOUS MOTION PATHS FROM SPATIALLY STATIONARY FLASHES

The model equations listed in section 8 provide an answer to the questions posed in section 3 concerning why individual flashes do not produce a percept of long-range motion, yet long-range interaction between spatially discrete pairs of flashes can produce a spatially sharp percept of continuous motion.

Intuitively, a signal for motion will arise when one or more of the functions $x_i^{(R)}$, $x_{i+1}^{(R)}$, $x_{i+2}^{(R)}$, ..., are activated sequentially through time, or alternatively the functions $x_i^{(L)}$, $x_{i-1}^{(L)}$, $x_{i-2}^{(L)}$, ..., are activated sequentially through time. Each activation $x_i^{(R)}$, or

$x_i^{(L)}$, represents the peak, or maximal activity, of a broad spatial pattern of activation across the network. The broad activation pattern (Figure 5b) is generated by the long-range Gaussian filter G_{ji} in (11) in response to a spatially localized flash to Level 1 (Figure 5a). The sharply localized response function $x_i^{(R)}$ is due to the contrast-enhancing action of the competitive network within Level 5 (Figure 5c). A stationary localized $x_i^{(R)}$ response will be generated in response to a single flashing input every time it occurs.

In contrast, suppose that two input flashes occur with the following spatial and temporal separations. Let the positions of the flashes be $i = 1$ and $i = N$. Let the activity $r_1(t)$ in (9) caused by the first flash start to decay as the activity $r_N(t)$ in (9) caused by the second flash starts to grow. Suppose, moreover, that the flashes are close enough that their spatial patterns $r_1 G_{1i}$ and $r_N G_{Ni}$ overlap. Then the total input

$$R_i = r_1 G_{1i} + r_N G_{Ni}$$ (16)

to the ith cell in Level 5 can change in such a way that the maximum value of the spatial pattern $R_i(t)$ through time, namely $x_i^{(R)}(t)$ in (14), first occurs at $i = 1$, then $i = 2$, then $i = 3$, and so on until $i = N$. A percept of continuous motion from the position of the first flash to that of the second will result.

This basic property of the motion OC Filter is illustrated by the computer simulations schematized in Figures 6–8. Figure 6 depicts the temporal re-

SPATIAL RESPONSE TO A SINGLE FLASH
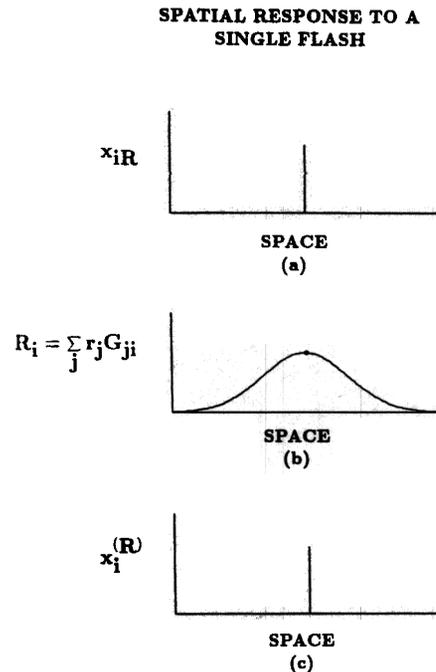
FIGURE 5. Spatial response of the Motion OC Filter to a point input. (a) Sustained activity of a Level 2 cell. (b) Total input pattern to Level 5. (c) Contrast-enhanced response at Level 5.

TEMPORAL RESPONSE TO A
SINGLE FLASH

INPUT

$J_{1R}(t)$

(a)

TIME-AVERAGE

$x_{iR}(t)$

TIME

(b)

GROWTH OF ACTIVATION AS A
FUNCTION OF SPACE AND TIME

$R_i = \sum_j r_j G_{ji}$

SPACE

(c)
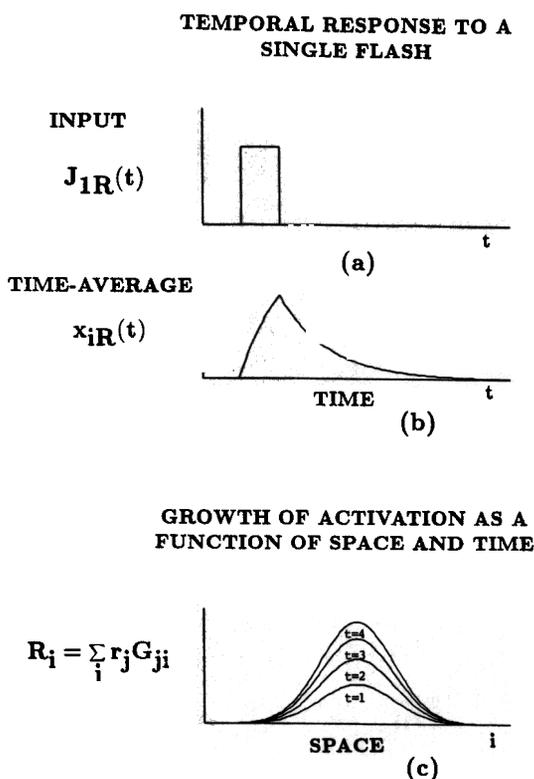
FIGURE 6. Temporal response of sustained response cells to a point input: (a) The input is presented for a brief duration at location 1. (b) The activity of the sustained response cell gradually builds up after input onset, then decays after input offset. (c) Growth of the input pattern to Level 5 through time with transient cell activity held constant. The activity pattern retains a Gaussian shape centered at the location of the input.

TEMPORAL RESPONSE TO
TWO SUCCESSIVE FLASHES

INPUTS

$J_{1R}(t)$

$J_{NR}(t)$

TIME-AVERAGES

$x_{1R}(t)$

$x_{NR}(t)$

TIME

FIGURE 7. Temporal response of the sustained response cells at Level 2 to two successive point inputs. One input is presented briefly at location 1, followed by a second input at location N. For an appropriately timed display, the decaying response at position 1 overlaps the rising response at position N.

sponse to a single flash at position 1 of Level 1. The sustained cell response at position 1 of Level 2 undergoes a gradual growth and decay of activation (Figure 6b), although the position of maximal activation in the input to Level 5 does not change through time (Figure 6c).

Figure 7 illustrates an important implication of the fact that the Level 2 cell activations persist after their Level 1 inputs shut off. If a flash at position 1 is followed, after an appropriate delay, by a flash at position N, then the sustained response to the first flash [e.g., $x_{1R}(t)$] can decay while the response to the second flash [e.g., $x_{NR}(t)$] grows.

Assume for the moment that transient signals are constant and let us examine the way in which the waxing and waning of sustained cell responses control the motion percept. Then the total input pattern R_i to Level 5 can change through time in the manner depicted in Figure 8. Each row of Figure 8a illustrates the total input to Level 5 caused, at a prescribed time t, by $x_{1R}(t)$ alone, by $x_{NR}(t)$ alone, and by both flashes together. Successive rows plot these functions at equally spaced later times. Note that as $x_{1R}(t)$ decays and $x_{NR}(t)$ grows, the maximum value of $R_i(t)$ moves

continuously to the right. Figure 8b depicts the position $x_i^{(R)}(t)$ of the maximum value at the corresponding times.

In summary, the time-averaged and space-averaged responses to individual flashes do not change their position of maximal activation through time (Figure 6c). In this case, "nothing moves." On the other hand, properly phased multiple flashes can generate a temporally and spatially averaged total response whose maximum moves continuously between the positions of the flashes through time (Figure 8).

10. RELATION BETWEEN FLASH SPATIAL SEPARATION AND SPATIAL SCALE

In the classic Neuhaus (1930) study of apparent motion in response to two-flash displays, it was discovered that there is a restricted range of interflash spatial separations over which apparent motion can be induced. This range extends from very close stimulus separations to about 4 degrees of visual angle. The precise upper limit on the spatial separations which give rise to motion depend on both the ob-

LONG-RANGE INTERACTION | SHARP MOTION SIGNAL

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

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


SPACE
(a)

SPACE
(b)
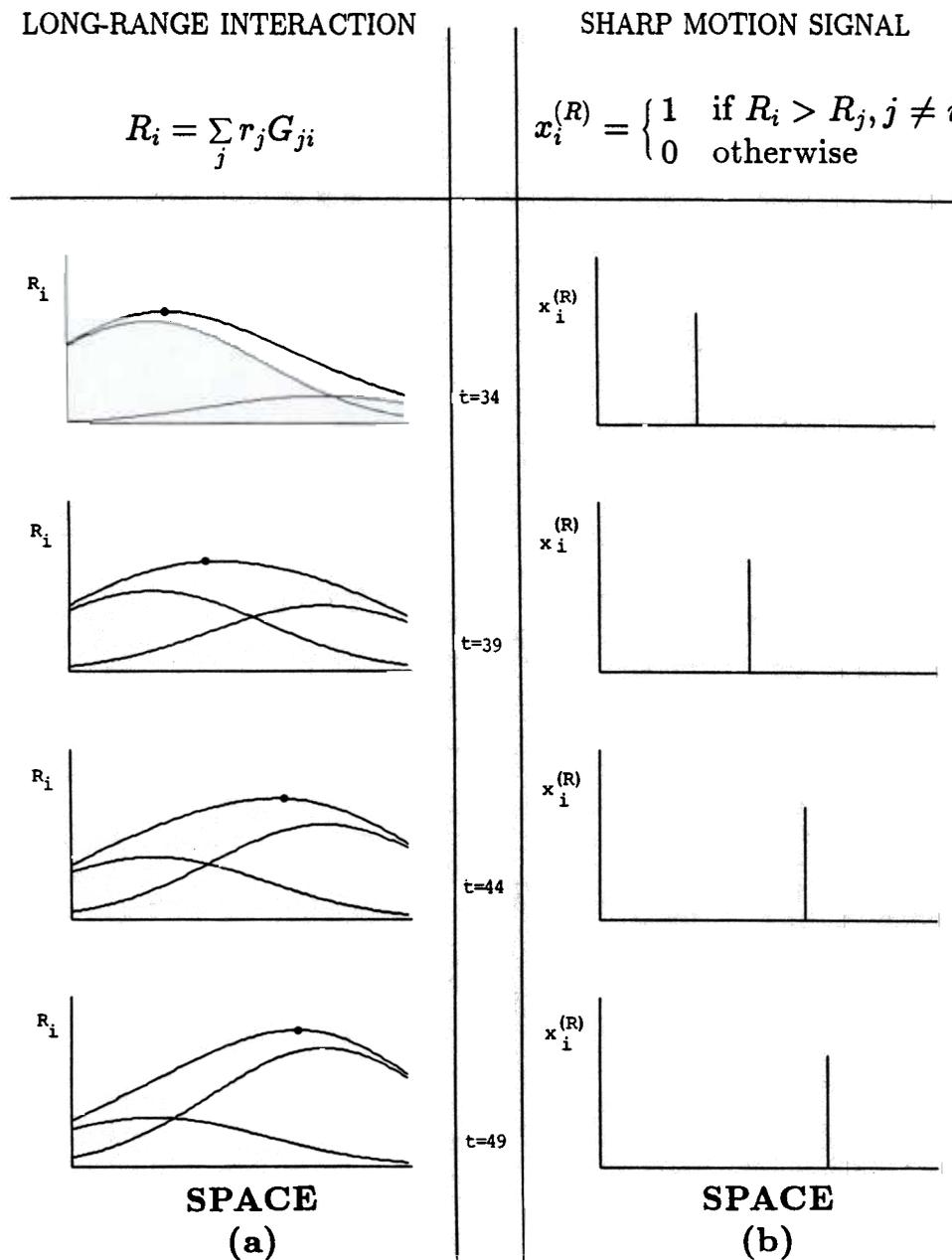
FIGURE 8. Motion OC Filter simulation in response to a two flash display. Successive rows correspond to increasing times: (a) The two lower curves in each row depict the total input to Level 5 caused by each of the two flashes. The input due to the left flash decreases while the input due to the right flash increases. The total input due to both flashes is a traveling wave whose maximum value moves from the location of the first flash to that of the second flash. (b) Position of the contrast-enhanced response at Level 5. Frame 1 offset and Frame 2 onset both occurred at $t = 32$, so that ISI $= 0$. Parameter $A = 0.05$, $B = 0$, $K = 42$, and $i = 1, 2, \ldots, 128$. The transient responses were held fixed at 1. See text for further details.

server and various properties of the stimulus elements.

In the first series of parametric computer simulations (hereafter Series I), our goals were to demonstrate that the model produces apparent motion over a range of flash separations, and that there exists a critical flash separation beyond which the model fails to produce the illusion of movement.

In addition to demonstrating that the model correctly predicts the existence of a critical flash sepa-

ration which limits the spatial extent of the apparent motion process, we analyzed the relationship between the size of the receptive fields (i.e., the Gaussian spatial kernels) in the motion OC Filter and the value of the critical flash separation.

In Figure 9 is shown a spatiotemporal diagram of a typical stimulus configuration employed as input in this simulation series. The stimulus display is represented as a 32 × 32 matrix of luminance values. This matrix describes the evolution over 32 time steps
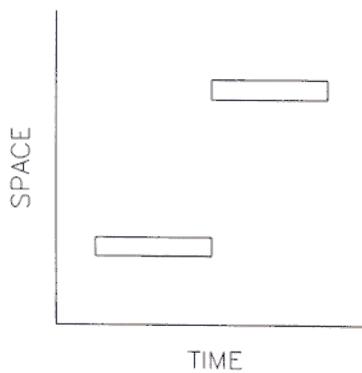
FIGURE 9. Space-time diagram of a two flash apparent motion display. The stimulus is represented as a 32 × 32 matrix of luminance values. The rectangles in the figure represent the spatiotemporal boundaries of a Frame 1 flash of width 3 centered at position s = 3 and presented from time t = 4 to t = 16; and a Frame 2 flash of width 3 centered at s = 24 and presented from t = 16 to t = 28.

of a one-dimensional spatial input image vector of length 32. The x-axis represents the temporal dimension of the stimulus display, and the y-axis represents the spatial dimension.

In the example shown in the figure, a flash of luminance 10.0 and spatial width 3 centered at the spatial coordinate s = 8 is superimposed on a black background beginning at time t = 4. The stimulus flash is then displayed in that spatial position until time t = 16, at which time it is turned off. At the same time that the first flash is turned off, a second flash of luminance 10.0 and spatial width 3 centered at s = 24 is turned on. This second flash remains on in this position for a duration of 12 time units, until it is turned off at t = 28.

In this example, the center-to-center spatial separation of the two flashes is 16 units. Let us denote by the symbol L the spatial separation of the flashes in a given simulation. In Series I, we factorially varied L with values of the spatial spread parameter K of the Gaussian kernel in eqn (11).

For the purposes of this series of simulations, we simplified the front-end (Levels 1–3) of the motion OC Filter by assuming, first, that the sustained cells responded at the center of the stimuli, rather than at their edges; and, second, that the transient cell responses always took on a value of 1. In all Series I runs, the values of the parameters A and B in eqn (5) were set equal to .12 and 0, respectively.

The first of these assumptions had the effect of superimposing the motion paths of the $x_i^{(R)}$ and $x_i^{(L)}$ signals defined in eqns (14) and (15), respectively. The second assumption made it easier to understand the nature of the spatiotemporal interactions of the activities generated by the two separate flashes involved in producing an apparent motion path. In particular, simulations using this approximation show clearly why a single flash, or series of flashes,

occurring in one position does not generate a continuous motion path, yet a properly timed series of two or more flashes in more than one position may generate a continuous motion path. In addition, these simulations provide parametric illustrations of prior analytic computations (see Appendix) which predicted that, for a fixed receptive field size K, the critical spatial separation between flashes, corresponding to the threshold between reports of "No motion" and "Motion," is related to K by the equation

$$L_{crit} = 2K. \qquad (17)$$

Thus, we expected that, for fixed values of K, the model would produce an output that simulated motion for values of $L < L_{crit}$, and an output that simulated stationarity for values of $L > L_{crit}$.

The simulations summarized in Figure 10 illustrate this property of the model. The rows of graphs in Figure 10 each correspond to a particular value of the flash separation L. From bottom to top, the values of L increase from 5 to 25 in steps of 4. The columns of graphs in the figure each correspond to a particular value of the receptive field size parameter K. Across columns, K is varied from 3 to 15 in steps of 4.

As in Figure 9, the rectangular outlines shown in each panel of Figure 10 describes the spatiotemporal extent of each of the two flashes. The x-axis represents the 32 time steps in the simulation, and the y-axis represents the 32 spatial locations in the 1-D stimulus image. In all panels, the temporal parameters and intensity of the stimulus are the same as in the example shown in Figure 9 and described above.

The circles plotted in each panel give the locations i at each time step of the maxima $x_i^{(R)}$ in (14). Large circles represent global maxima of activity, and small circles represent other local maxima. We suggest that it is the paths of these maxima that are phenomenologically observed in reports of apparent motion.

For fixed K and $L < 2K$, the activities R_i generated by the two separate flashes sum to create a single global maximum of activity that moves from the location of the first flash to that of the second as the neural activity resulting from the presence of the first flash decays after the flash is terminated and the activity due to the second flash builds up after the onset of the second flash. Figure 8a show a series of the spatial patterns R_i plotted through time to illustrate the motion trajectory of the maximum value. In this illustration, the decaying trace of an input flash of width 3 centered at $s = 8$ interacts with the increasing neural activity due to the onset of a second flash of width 3 centered at $s = 24$. The total input to position i at Level 5 of the model due to the flash centered at $s = 8$ is thus $\Sigma_{j=7}^{9} r_j G_{ji}$. Similarly, the Level 5 input at position i due to the flash centered at $s = 24$ is $\Sigma_{j=23}^{25} r_j G_{ji}$.

APPARENT MOTION THRESHOLD DEPENDS
ON THE SPATIAL SEPARATION BETWEEN
FLASHES AND THE SCALE OF THE MOTION
OC FILTER

(L$_{crit}$ = 2K Rule)
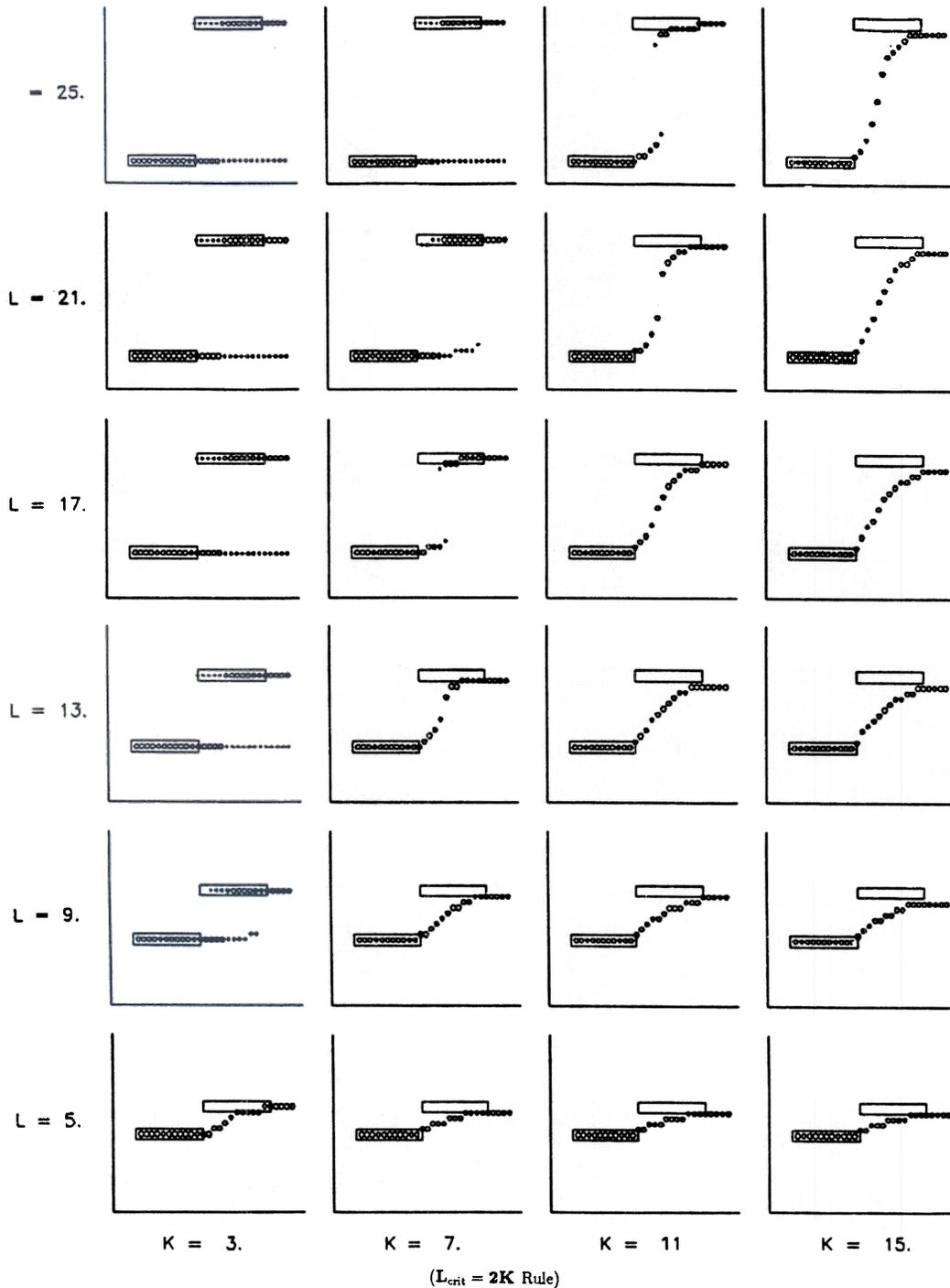
FIGURE 10. Paths of the Level 5 motion signal as a function of flash separation L and Gaussian filter width K. The rectangular outlines in each panel represent the spatiotemporal extent of the two stimulus flashes. Large circles are plotted at the locations of the global maximum of the right-motion signal pattern R_i over 32 time steps. Small circles represent locations of other local maxima of pattern R_i. A continuous motion signal occurs if $L \le 2K$. Parameter $A = .12$ and $B = 0$. Transient responses were fixed at 1 through time. See text for further details.

In contrast, for values $L > 2K$, the neural activities due to the separate flashes do not combine to produce a single global maximum which moves over time. Instead, two local maxima appear at the locations of the separate flashes and remain fixed in place over the time-course of the simulation. We identify the case of the stationary local maxima with reports of lights blinking on and off in place.

11. PARTIAL MOTIONS AND EQUAL TIME MULTIPLE SCALE MOTIONS

The transition from motion to blinking-in-place occurs in the simulations at the spatial separation $L_{crit} = 2K$. For values of L slightly greater than L_{crit}, a transitional percept is produced by the model, as in row 1, column 3 and row 3, column 2 of Figure 10. We suggest that this result corresponds to reports in the literature of "partial movements" (Kolers, 1972; Wertheimer, 1912).

It is of interest to compare motion paths produced by the model at different scales (i.e., different values of the parameter K) in response to an identical stimulus. For the purposes of this comparison we plot the motion paths for different values of K corresponding to a single flash separation ($L = 16$) in Figure 11. In order to avoid a cluttered diagram, we have replaced the circles which previously marked the time-course of the locations of the local and

MOTION PATHS GENERATED BY FILTERS OF DIFFERENT SPATIAL SCALES

TIME
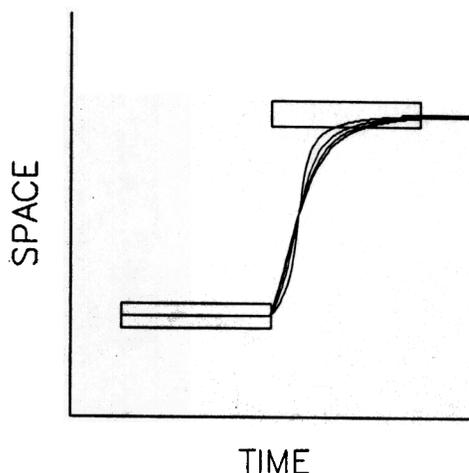
FIGURE 11. Apparent motion paths produced by the motion OC Filters given four different values of spatial scale K. Curves corresponding to different scales cross at a location exactly halfway between the locations of the two flashes (Equal Half-Time Property). The stimulus was represented as a 320 × 320 luminance matrix. Parameter A = .012 (equivalent to A = 0.12 for a 32 × 32 simulation). The curves correspond to K = 90 (steepest sigmoid), 110, 130, and 150 (shallowest sigmoid). As $A \to \infty$, curves corresponding to different K converge.

global maxima with smooth curves. Each curve corresponds to a different K.

As noted above, for fixed L, motion will only be observed at receptive field scales where $K > L/2$. Moreover, the motion paths produced at scales larger than $K = L/2$ are, in general, not straight lines. Paths following straight lines on our space-time diagrams indicate constant velocity motions. Instead, the model predicts that motion accelerates away from the location of the first flash, then decelerates into the location of the second flash (see Appendix). These acceleration effects produced by the model are consistent with the observations of two-flash apparent motion displays. Furthermore, as K approaches the value of $L/2$ from above, the slope of the motion curve becomes very steep, indicating high velocities. Such high velocities may exceed the velocity sensitive range of motion detectors which are activated by the moving peak of neural activity whose path is charted in these diagrams. When this happens there will exist a spatial region in the center of the path of the peak for which the motion will be undetectable, although partial movements will still be registered as accelerations away from the first flash, followed by disconnected decelerations into the second flash. This prediction of the model is, again, consistent with observations of partial movement (Kolers, 1972, p. 9).

Although the motion paths traversed by the moving global maxima within different spatial scales K and distances L do not strictly overlap, they all have the property of passing the point which lies exactly halfway between the locations of the first and second flashes at the same instant of time. This Equal Half-Time Property was proved analytically prior to the simulations (Grossberg, 1977). The analytic derivation of the property is given in the Appendix. The Equal Half-Time Property clarifies how the movement of the peak neural activity can speed up to cover larger distances in the same amount of time when flashes are placed at greater distances L while keeping the ISI constant. Indeed, it has been shown experimentally that "large variations in distance are accommodated within a near-constant amount of time" (Kolers, 1972, p. 26). This property also clarifies how the motion paths generated by multiple spatial scales K can remain approximately in phase to generate a coherent total motion signal.

12. GROUP MOTION IN RESPONSE TO THE TERNUS DISPLAY

In the second series of computer runs (Series II), we simulated the Ternus display. A space-time diagram of the simulated Ternus display is shown in Figure 12. As in the Series I simulations, the stimulus in these runs was represented as a 32 × 32 space-time matrix of luminance values. All parameters except
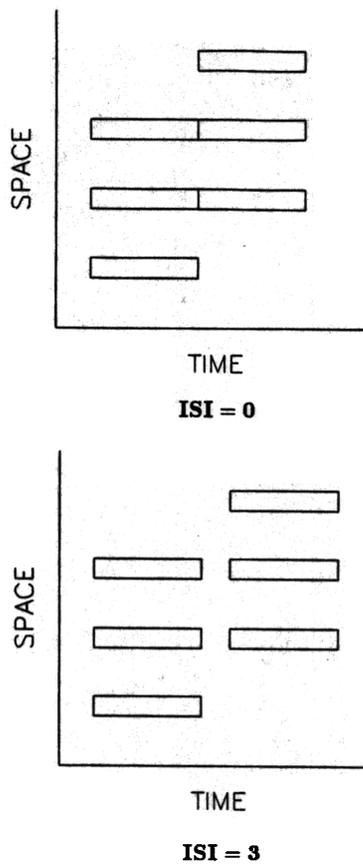
FIGURE 12. Space-time diagrams of two Ternus displays: (a) ISI = 0 (b) ISI = 3. The stimulus was represented as a 32 × 32 luminance matrix. The rectangular outlines represent the spatiotemporal extent of the flashes. See text for spatiotemporal parameters of the display.

K were chosen as in the Series I simulations. Each frame of the display consisted of three simultaneously presented 1-D luminance patches superimposed on a black background. Each patch was made up of 3 pixels of luminance 10.0. In Frame 1, the centers of the three Ternus elements were located at

the equally spaced spatial positions 6, 13, and 20. In Frame 2, the element centers were located at positions 13, 20, and 27. Thus two of the three element locations were the same in both frames.

The duration of each frame was fixed at 12 time units. The first frame was turned on at $t = 4$, and off at $t = 16$. Figure 12a corresponds to the case in which the ISI = 0. In Figure 12b, the ISI of the display equals 3.

In the Series II runs, we manipulated the receptive field size parameter K to show how whole motion could be generated if the value of K is chosen large enough. Across panels from left to right, K is increased from 2 to 8 in steps of 2. The spatiotemporal parameters of the stimulus display remain fixed across panels. When K is small (e.g., 2), the output indicates the presence of local maxima at the spatial locations of each individual element of the Ternus display, which thus appear to remain stationary over time. When K is large (e.g., ≥ 4), only one spatial maximum in the output is observed through time. This single maximum first appears in the center of the three elements presented in Frame 1 at the onset of that frame, and moves smoothly towards the center of the three elements in Frame 2 after the onset of this frame (Figure 13).

13. ELEMENT MOTION IS DUE TO THE GATING OF SUSTAINED MECHANISMS BY TRANSIENT MECHANISMS

The stationary maxima that occur when K is small are due to the fact that the Gaussian-filtered signals in eqns (11)–(13) that are generated by the three elements in each frame of the Ternus display do not significantly interact across space. At sufficiently large values of K, the Gaussian-filtered signals due to the individual elements in each frame summate across space to yield a single global maximum at the

TERNUS DISPLAY:
VARYING THE MOTION OC FILTER
SCALE

K = 2.   K = 4.   K = 6.   K = 8.
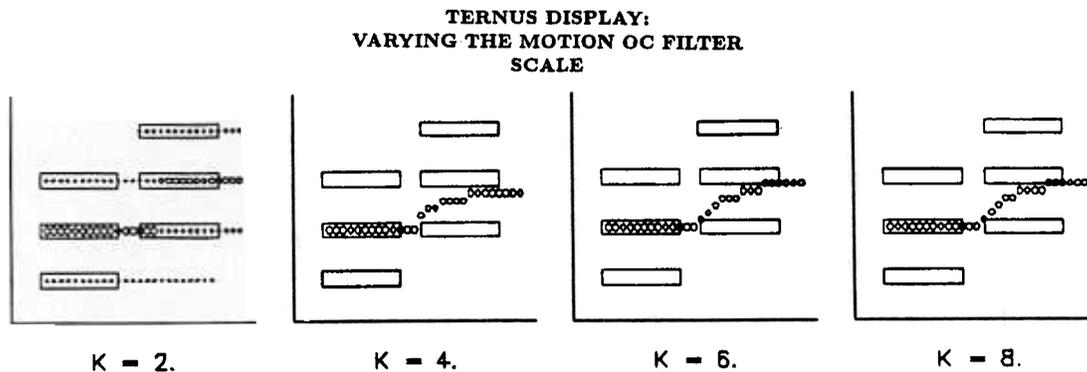
FIGURE 13. Group motion due to a Ternus display: Space-time paths of global maxima (large circles) and other local maxima (small circles) corresponding to Gaussian filters of four widths K. In order to generate apparent motion, K must be large enough so that responses to the single flashes can combine to produce a single moving global peak. Given such a K, group motion always occurs if the transient responses are fixed at a positive constant value. See text for spatiotemporal parameters of the display.

position of the middle element in each frame (Figure 14). If K is chosen sufficiently large for interactions to occur between frames, then as the activation caused by Frame 1 decays and the activation caused by Frame 2 grows, the maximum activation moves smoothly from a location at the middle of Frame 1 to the middle of Frame 2.

This result suggests two conclusions: (a) group motion may result when the interaction length K is chosen sufficiently large to enable the Gaussian-filtered activations caused by the two frames to interact; and (b) at such values of K, the global maximum of activation within each frame occurs at the center of the frame.

Given this result, how can element motion ever be caused? If, in order for *any* motion to occur, K must be chosen large enough to allow interactions between Frames, then surely K will also be large enough to enable the maximum activity to occur at the center of each frame, thereby forcing group motion to occur whenever motion is produced.

Although this result is generated when the model's transient interactions are held constant, it is not produced by the full model. We now trace the occurrence of element motion to the multiplicative

GROUP MOTION IN THE TERNUS PARADIGM

If $L < 2K$ spatially-filtered element edges group to form one global maximum

FIGURE 14. Spatial summation underlying group motion response to the Ternus display: Three sustained response patterns across space in (a) generate Gaussian-distributed inputs to Level 5 that summate into a unimodal total input in (b) whose maximum value in (c) is centered at the middle of the display. If the winner-take-all competition in (c) is replaced by partial contrast-enhancement of the total pattern in (b), then a motion signal is produced whose width covaries with the total separation of the three flashes in each Ternus frame.

interaction of the transient cell mechanism with the sustained cell mechanism, as in eqns (9) and (10). This result supports and explains the intuition derived from experiments that "visual persistence" controls the percept of element motion by causing the perceived stationarity of the inner elements of the Ternus display (section 5).

Such a multiplicative of x_{iL} by y_i^+ to generate a right motion signal and by y_i^- to generate a left motion signal, and of x_{iR} by y_i^- to generate a right motion signal and by y_i^+ to generate a left motion signal, eliminates the ambiguity that exists between direction-of-contrast and direction-of-motion. We hereby reduce the transition from element motion to group motion to a basic property of motion-sensitive receptive fields.

14. COMPARISON WITH THE MARR–ULLMAN MODEL

Before proceeding to this demonstration, its significance may be clarified by a comparison of the motion OC Filter with the well-known Marr–Ullman model (Table 2). Indeed, combining a product of a sustained mechanisms with a transient mechanism was the basic insight of the Marr and Ullman (1981) model of motion detection. In the Marr–Ullman model, the sustained channel was modeled as

$$S = \nabla^2 G * I \qquad (18)$$

and the transient channel was modeled as

$$T = \frac{\partial}{\partial t} (\nabla^2 G * I),$$

where I is the input pattern. This Marr–Ullman model is not, however, sufficient to explain our results for several reasons.

One reason is that the signals J_{ik} in (5), y_i^+ in (7), and y_i^- in (8) need to be rectified, unlike S and T in (18) and (19).

A second reason is that the temporal averaging in eqns (5) and (6) is essential to achieve our results.

TABLE 2
Comparison with the Marr–Ullman Model
$S = \nabla^2 G * I$
$T = \frac{\partial}{\partial t} S$
$S \cdot T$
Differences Between Marr–Ullman Model and Motion BCS
1. Signal rectification
2. Temporal integration
3. Shunting dynamics
4. Long-range spatial interaction
5. Contrast-enhancing competition
6. Rationale

Otherwise, the interaction between the temporal decay of one activation and the temporal growth of another activation centered at a different position could not occur (Figure 7).

A third reason is that the shunting interactions in eqns (5) and (6) are needed, in general, to generate finite maximal activations $1/B$ and D/E, respectively, within which the position and the size of the global maxima $x_i^{(R)}$ and $x_i^{(L)}$ in eqns (14) and (15) can be controlled even if input intensities I_i vary by large amounts.

A fourth reason is that the Gaussian kernel G_{ji} in (11), which determines the spatial extent of the Level 5 receptive fields, is much broader than that envisaged by Marr and Ullman (1981). Their Gaussian G in eqns (18) and (19) corresponds to the local spatial filtering that sets up the individual inputs J_{iL} and J_{iR} in eqn (5).

A fifth reason is that the contrast-enhancing competitive interaction at Level 5 does not exist in their model, because the long-range Gaussian G_{ji} does not exist.

Finally, and perhaps most importantly, the design rationale for the motion OC Filter was not articulated in the Marr–Ullman model. The need for a special design that enables insensitivity to direction-of-contrast to coexist with sensitivity to direction-of-motion does not become salient until one studies a static OC Filter in which insensitivity to direction-of contrast implies insensitivity to direction-of-motion (Figure 2). In the Marr–Ullman model, simplified individual terms such as $S \cdot T$ were modeled, but not their combinations, as in eqns (9) and (10).

The property of insensitivity to direction-of-contrast in the static BCS reflects one of the fundamental new insights of this model of vision. Insensitivity to direction-of-contrast is possible within the BCS because all boundary segmentations within the BCS are perceptually invisible. Visibility is a property of a complementary system, the Feature Contour System, or FCS (Cohen & Grossberg, 1984; Grossberg, 1987a; Grossberg & Mingolla, 1985a; Grossberg & Todorović, 1988), whose computations are sensitive to direction-of-contrast. Thus the Marr–Ullman model could not articulate the heuristics of the motion BCS because it was not based on an understanding of BCS/FCS complementarity.

15. FORMATION TIME, GAMMA MOTION, AND THE SYNTHESIS OF A GLOBAL MOTION SIGNAL FROM LOCAL MOTION SIGNALS

The sustained oriented reactions (x_{iL}, x_{iR}) and the transient on and off reactions (y_i^+, y_i^-) interact to generate some remarkable properties.

Figure 15 illustrates how, in response to a step

SUSTAINED-TRANSIENT GATING
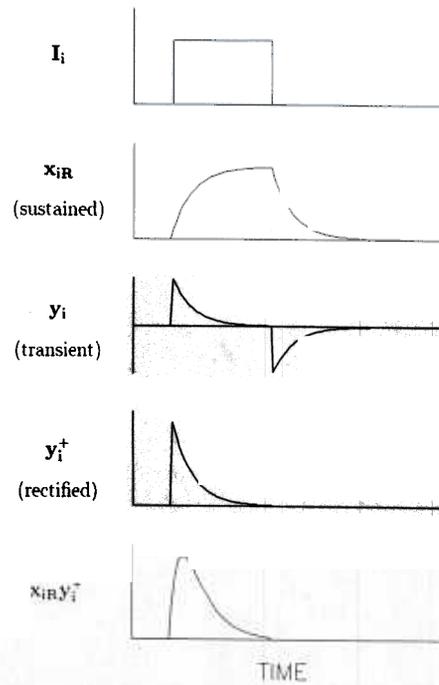
FIGURE 15. Multiplicative gating at Level 4 of sustained responses from Level 2 and transient responses from Level 3 to generate a direction-of-motion sensitive response: Presentation of an input I_i produces both a sustained response x_i and a transient response y_i. The sustained response x_{iR} is gated by the rectified On response y_i^+ to generate an $x_{iR}y_i^+$ response that is sensitive to direction-of-motion and direction-of-contrast. Parameters are $A = .12$, $B = 0$, $C = .12$, $D = .12$, $E = 0$, and $\Sigma_j I_{ji}F_{ji} = J_i$.

input in time, sustained and transient reactions interact at Level 4 to generate a unimodal response function $x_{iR}y_i^+$. Note that flash duration must be sufficient to generate a time interval where both the sustained response x_{iR} and the transient response y_i^+ are large. The existence of such a "figure formation" function has been experimentally reported (Kolers, 1972). Relevant data are summarized in Figure 16.

As a result of the time required in the model for figure formation, the apparent motion percept may be influenced by the flash duration as well as the ISI. Because the SOA is the sum of the flash duration and the ISI (see section 4), the importance of figure formation time in the model may explain why experimentalists have debated whether it is the manipulation of ISI or SOA that is critical in apparent motion studies.

Note that the present model also identifies the figure formation time of Kolers with the temporal form factor for apparent motion analyzed by Burt and Sperling [see section 7, especially eqn (1)].

Figure 17 summarizes how the sustained and transient reactions combine in response to the onset and
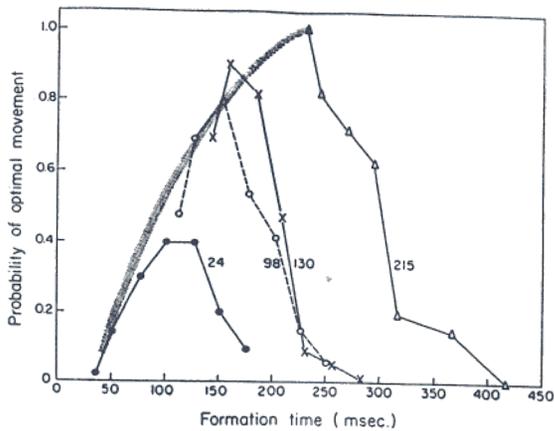
(a)
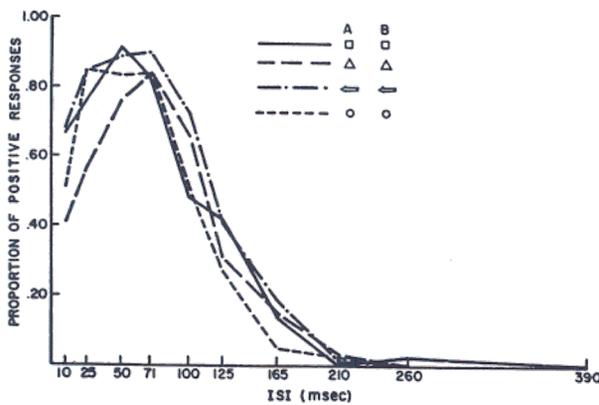
(b)

FIGURE 16. Figure-formation time for perceiving good motion: (a) Probability of seeing beta motion in response to a two flash apparent motion display as a function of the onset-to-onset interval SOA for flashes of duration 24, 98, 130, and 215 ms. (b) Probability of seeing beta motion as a function of ISI for various combinations of dissimilar shapes. The simulated temporal form factor generated in Figure 15 has a shape similar to the empirically derived figure-formation function. Reprinted with permission from P. A. Kolers, *Aspects of motion perception*, 1972 (Figures 3.5 and 4.4), Pergamon Press plc.

offset of a single luminant step on a less luminant background. At stimulus onset, the local left motion signal l_i is activated at the left side of the step, whereas the local right motion signal r_i is activated at the right side of the step. Thus there is a tendency, at least as measured by local motion signals, for stimulus onset to create an impression of stimulus expansion. The opposite is true at signal offset; local motion signals, in themselves, create an impression of stimulus contraction. This property is sometimes called gamma motion (Bartley, 1941; Kolers, 1972).

When two luminance steps at different positions are sequentially activated through time, their local signals r_i interact across space to generate R_i, and their local signals l_i interact to generate L_i.
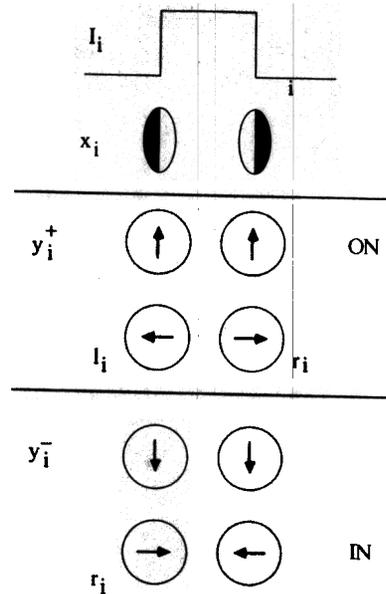
FIGURE 17. Gamma motion due to onset and offset of a single flash: A single bright flash on a dark background gives rise to sustained and transient response at the locations of its edges that produce an apparent outward expansion at flash onset and an inward contraction at flash offset.
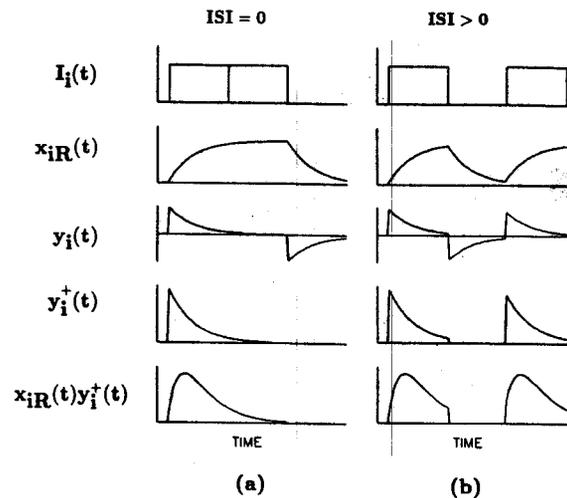

FIGURE 18. Mechanisms for generating group and element motion in the Ternus display. (a) Element motion when ISI is sufficiently small: At display locations 2 and 3, no transient responses y_i are generated at the offset of Frame 1 or the onset of Frame 2. Thus, no contribution to the overall motion signal is made by these locations. Element motion results. (b) Group motion when ISI is sufficiently large (ISI = 40): Gated sustained-transient signals develop at all display locations, including locations 2 and 3, and thus contribute to the overall motion signal. Group motion results. Parameters $A = .09$, $B = 0$, $C = .09$, $D = .09$, and $E = 0$, and $\sum_j l_j F_{ji} = J_i$.

In response to a Ternus display at an ISI = 0 (Figure 12a), the on-transient detectors y_i^+ and the off-transient detectors y_i^- are inactive at the positions of the second and third luminance steps when Frame 1 switches to Frame 2, as in Figure 18a. Thus the local motion signals r_i and l_i equal zero at these positions and do not contribute to the synthesis of R_i and L_i, respectively. Only the first luminance step in Frame 1 and the fourth luminance step in Frame 2 interact across space to generate R_i and L_i. Element motion is then generated between these extremal stimuli of the display by the same mechanisms that generate beta motion between two flashes.

In contrast, suppose that the ISI is chosen sufficiently large that the transient detectors can respond both to the offset and the onset of the second and third luminance steps in Frames 1 and 2, as in Figure 18b. Then each of the four luminance steps generates appropriately timed local motion signals r_i and l_i, which interact across space to form R_i and L_i. Group motion is then generated in the manner described in section 12.

16. SIMULATING THE TRANSITION BETWEEN ELEMENT MOTION AND GROUP MOTION

In the final set of simulations (Series III), we introduced all of the mathematical features of the motion OC Filter defined in section 8. Specifically, the sustained cell activation x_{iL} responds to an increase of contrast to the left, whereas x_{iR} responds to an increase of contrast to the right. Thus we located x_{iL} and x_{iR} at the right and left edges, respectively, of the Ternus display elements. The responses of these sustained mechanisms were gated with the responses of the y_i^+ and y_i^- transient mechanisms to derive local left- and right-motion signals according to eqns (9) and (10). Thus, in these simulations, two sets of space-time output plots were generated for every set of stimulus conditions. These two sets of output corresponded to the $x_i^{(R)}$ and $x_i^{(L)}$ of eqns (14) and (15).

In Figure 19 are plotted the results of varying the ISI of the Ternus display on the motion paths computed by the model. The space-time diagrams shown in column 1 correspond to the $x_i^{(L)}$ (top row) and $x_i^{(R)}$ (bottom row) outputs in response to a Ternus display with ISI = 0. The diagrams in column 2 correspond to a Ternus display with ISI = 14.

These simulations were performed on a 128 × 128 space-time input matrix of luminance to achieve a finer spatiotemporal resolution than in earlier series. As in the previous simulations, the rectangular outlines in the diagrams designate the spatiotemporal boundaries of the elements in the Ternus display. In the simulations, term $\Sigma_j I_j F_{ji}$ in eqn (6) was set equal

to 10.0 at all element edges. Term J_{ik} in eqn (5) was also set equal to 10.0 at the edges.

The elements in the first frames were each 9 pixels wide. In Frame 1, the elements were centered at spatial locations S = 12, 48, and 84. In Frame 2, they were centered at locations S = 48, 84, and 120. The duration of each frame was 56 time units. Frame 1 was turned on at t = 2, and off at t = 58. In the zero ISI simulations, Frame 2 was turned on at t = 58 and off at t = 114. In the simulations in which the ISI = 14, Frame 2 was turned on at t = 72, and off at t = 128.

The spatial scale K was fixed at a value of 60. The parameters in eqns (5) and (6), which determine the temporal response of the front-end sustained and transient mechanisms were set to the values A = 0.05, B = 0, C = 0.05, D = 0.05, and E = 0. The increase in the temporal resolution by a factor of 4 in the Series III simulations implies a redefinition of the time scale, and thus the decay constant. The values of 0.05 chosen here for both decay constants A and C are physically equivalent to the value 0.20 employed in the simulations of Series I and II.

In order to exhibit the properties of the model that produce the transition from the percept of element motion to that of group motion as the ISI of the display is increased, we plot in Figure 20 the neural activities of several component processes of the motion OC Filter in response to Ternus displays with ISI = 0 (first column), and ISI = 14 (second column).

Each row in Figure 20 corresponds to the output of a particular model variable. In rows (A) through (F) are shown the responses of the oriented sustained mechanisms x_{ik}, the time derivatives $d/dt \, x_i$ of the unoriented mechanisms, the activities of the y_i^+ and y_i^- transient mechanisms, the local left-motion signals l_i and right-motion signals r_i, the spatially-convolved left-motion signals L_i and right-motion signals R_i.

The x-axis in each panel represents time, and the y-axis in each panel corresponds to the magnitude of the neural activity at a particular spatial location and level of the model. Each of the spatial locations plotted corresponds to the spatial location of an element edge in the Ternus display. Left edges are designated by the letter "L," and right edges by the letter "R." For the purpose of this plot we have numbered the four spatial locations corresponding to the positions of the elements in the display. Thus, for example, "L1" refers to the spatial location of the left edge of the element centered at s = 12 (element location 1); and "R4" refers to the location of the right edge of the element centered at s = 120 (element location 4).

The activities of the sustained mechanisms build up slowly after the onset of a stimulus and decay

Element Motion
ISI = 0

Group Motion
ISI > 0

(a)

(b)

(c)

(d)
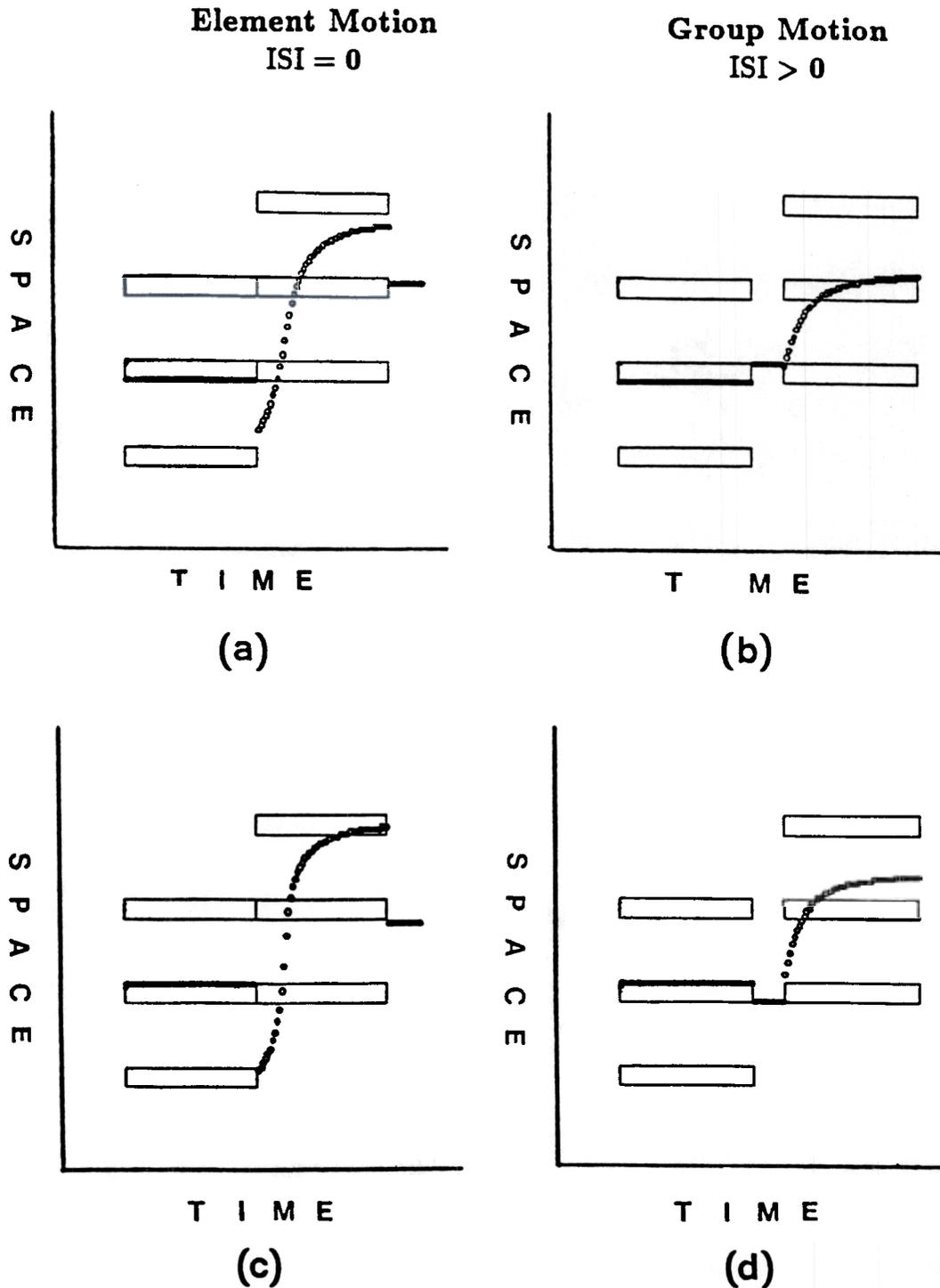
FIGURE 19. Simulated group and element motion to Ternus displays: (a) Element motion response of a left-motion filter when ISI = 0. (b) Group motion response of a left-motion filter when ISI > 0. (c) Element motion response of a right-motion filter when ISI = 0. (d) Group motion response of a right-motion filter when ISI > 0. Circles indicate the locations of the peak neural activity at each of 128 time steps. See section 16 for further details.

gradually after a stimulus offset. The time derivatives of these reactions respond differently to stimulus onset and offset. Onset causes a transient positive reaction. Offset causes a transient negative reaction. When the ISI = 0, neither positive nor negative transients are observed at the onset of Frame 2 nor

the offset of Frame 1 at element locations 2 and 3. When the ISI is increased to 14, positive and negative transients are observed at these locations both at the onset of Frame 2 and the offset of Frame 1, respectively. This difference is the basis of element motion and group motion, respectively, in the model.

ISI=0   ISI>0

LEVEL 2

(A) SUSTAINED RESPONSES

LEVEL 3

(B) TRANSIENT RESPONSES
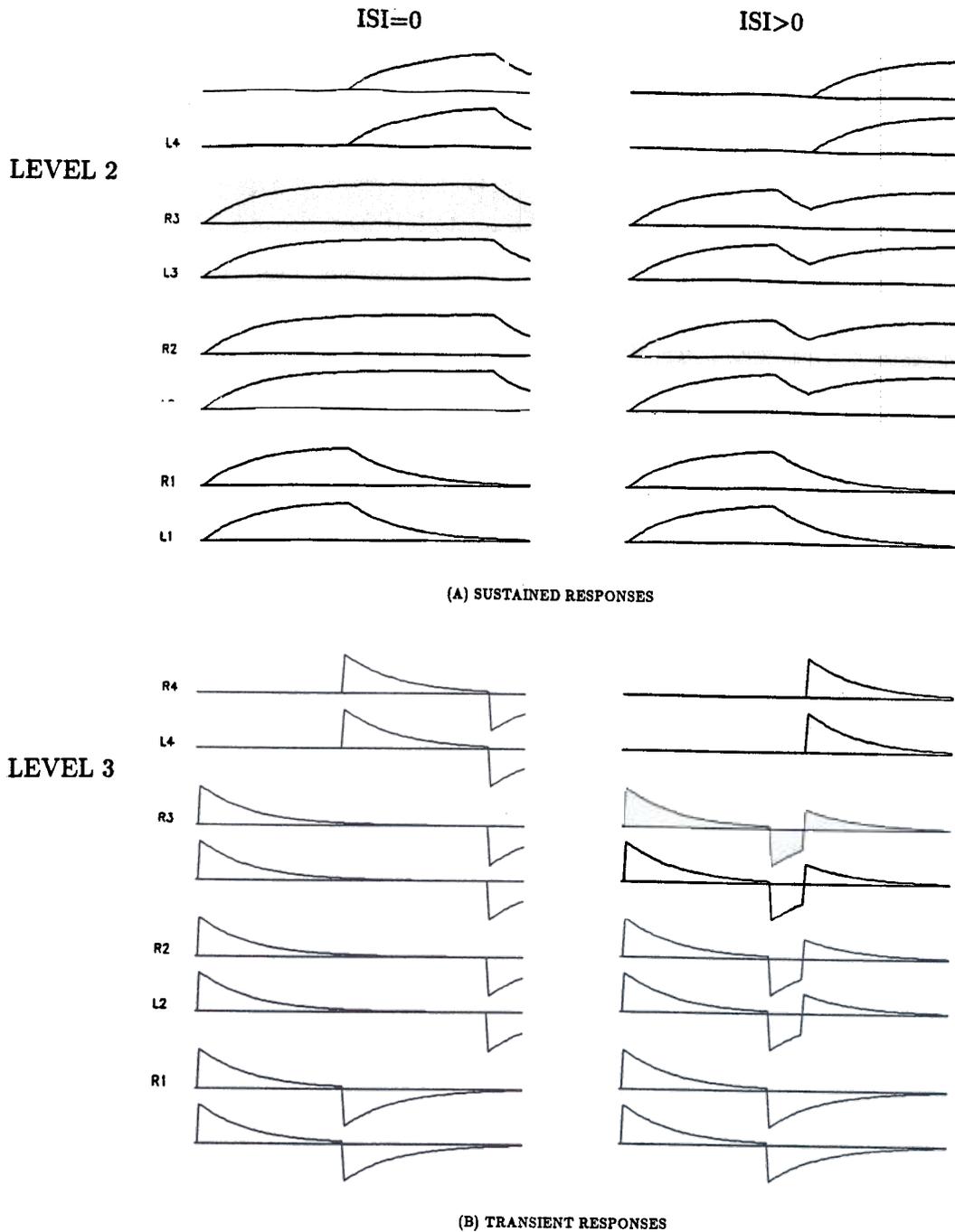
FIGURE 20. Level-by-level analysis of motion OC Filter responses to Ternus displays. The left column corresponds to ISI = 0, and the right to ISI > 0. (A) Level 2 (sustained) responses x_{iL} and x_{iR}. (B) Level 3 (transient) responses $d/dt\, x_i$. (C) Rectified on transient [$y^+(t)$, solid lines] and off transient [$y^-(t)$, dotted lines]. (D) Level 4 local left [l_i, solid lines] and right [r_i, dotted lines] motion signals. (E) Gaussian-convolved input to Level 5; response of left motion filter (F) Gaussian-convolved input to Level 5; response of right motion signal. In all panels, the symbols L1, R1, L2, R2, etc. refer to the locations of the left and right edges, respectively of elements 1, 2 etc. For the stimulus and model parameters employed in the simulation, see section 16 in the text.

ISI=0   ISI>0

LEVEL 3

$$y_i^+(t) =$$
solid lines

$$y_i^-(t) =$$
dotted lines

R4
L4
R3
L3
R2
L2
R1
L1
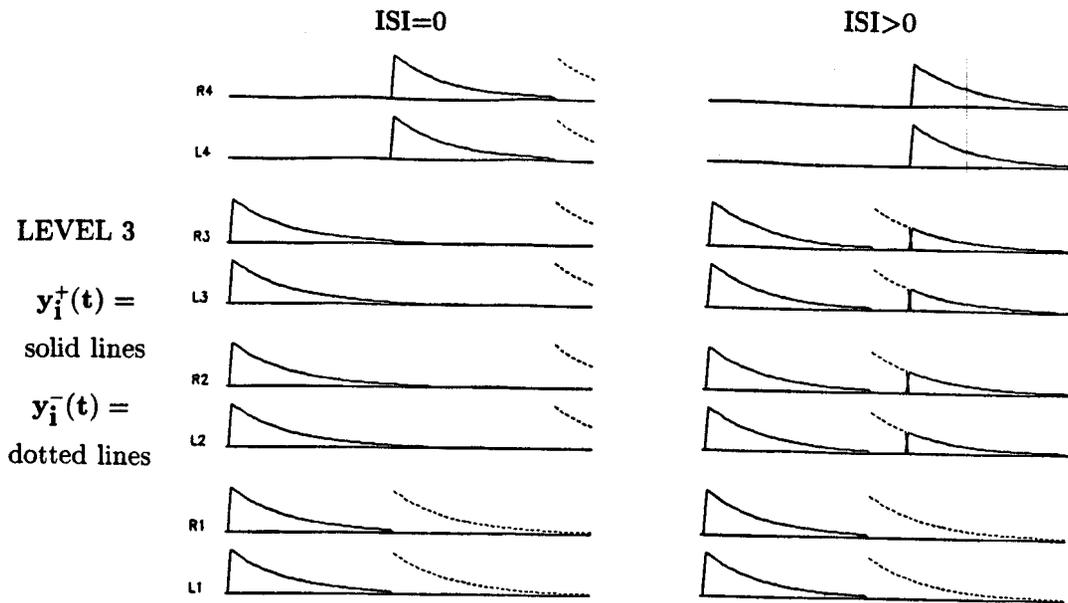
(C) ON AND OFF (RECTIFIED) RESPONSES

LEVEL 4

$$l_i =$$
solid lines

$$r_i =$$
dotted lines

R3
L3
R2
L2
R1
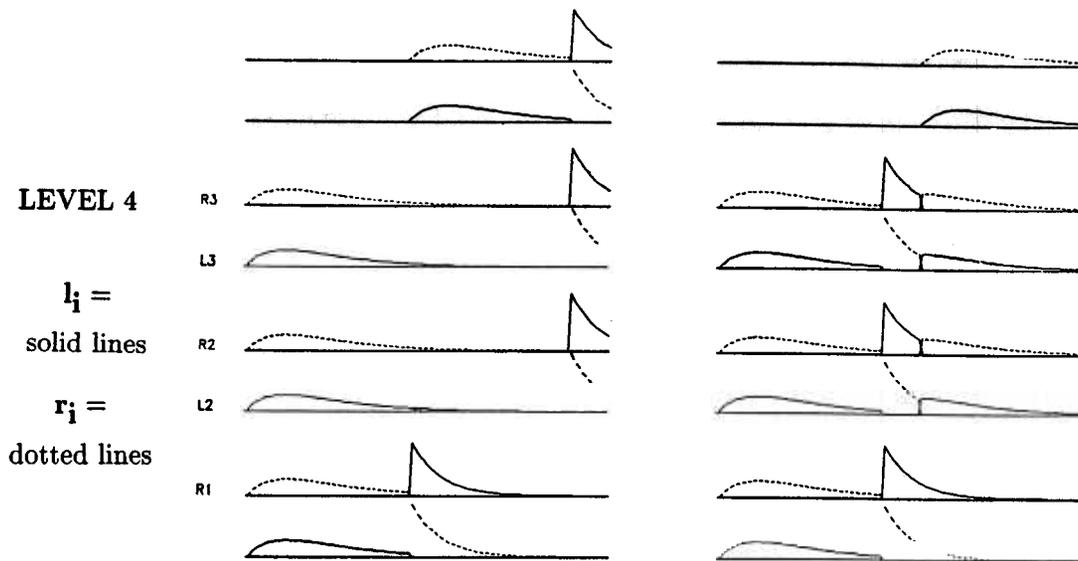
(D) LOCAL LEFT AND RIGHT MOTION SIGNALS
(SUSTAINED-TRANSIENT GATED RESPONSE)

FIGURE 20 (Continued)

ISI=0 ISI>0

INPUT TO
LEVEL 5

$$L_i = \sum_j l_i G_{ji}$$

R4

L4

R3

L3

R2

L2

R1

L1
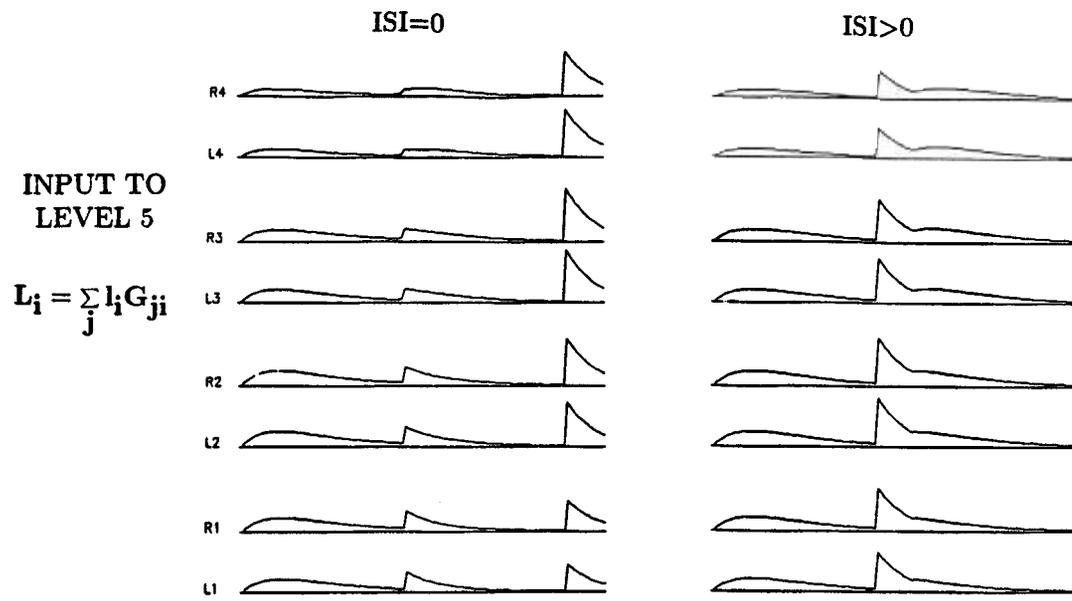
(E) CONVOLVED LEFT MOTION OUTPUT

INPUT TO
LEVEL 5

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

R3

L3

R2

L2

R1

L1
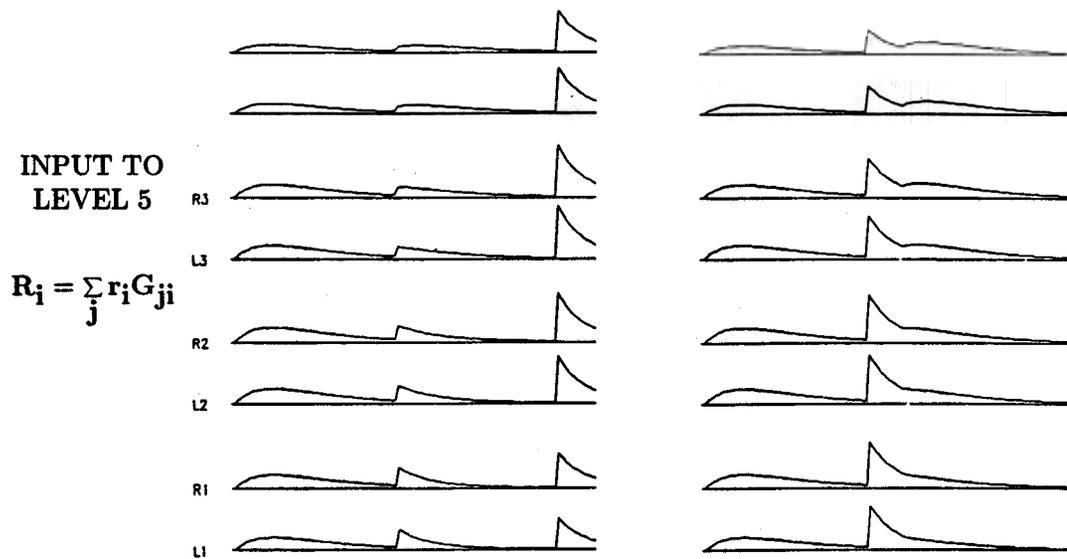
(F) CONVOLVED RIGHT MOTION OUTPUT

FIGURE 20 (Continued)

Physiologically, positive and negative transients in stimulus contrast are responded to by separate subsets of mechanisms: the on transient cells and the off transient cells. In (C), we model transient on and off responses by rectifying and thresholding the transient signals graphed in (B). The solid lines show the response over time of the on transient mechanism y_i^+. The broken lines show the response over time of the off transient mechanism y_i^-.

In these simulations the values of the threshold parameters Γ and Ω in eqns (7) and (8) were set to zero. In general, they can be chosen greater than zero to alter the threshold ISI at which a transition from element motion to group motion will be generated. Thus, the model provides a mechanism by which stimulus factors associated with persistence (i.e., slow decay) can affect the transitional ISI, as discussed in section 5.

In (D) are shown the l_i and r_i signals at each element edge location. When the $ISI = 14$, the transients which occur at the edges of element locations 2 and 3 at the offset of Frame 1 and the onset of Frame 2 produce strong local motion signals which contribute to the overall group motion signal. When the $ISI = 0$, these local motion signals are absent.

The L_i and R_i at element edge locations are plotted in (E) and (F), respectively. The maxima of these activities across all spatial positions corresponds to the $x_i^{(L)}$ and $x_i^{(R)}$ simulated in Figure 19.

In summary, the paths of $x_i^{(L)}$ and $x_i^{(R)}$ demonstrate the same dependence of element and whole motion on ISI as do the results of psychophysical experiments. At short ISIs, element motion is observed. At longer ISIs, group motion is observed.

17. VIRTUAL MOTIONS AND THEIR SUPPRESSION BY ASYMMETRIC INHIBITION AT HYPERCOMPLEX CELLS

An interesting feature of these simulations is that two motion waves $x_i^{(R)}$, $x_{i+1}^{(R)}$, $x_{i+2}^{(R)}$, . . . , and $x_i^{(L)}$, $x_{i+1}^{(L)}$, $x_{i+2}^{(L)}$, . . . , are generated from left to right in response to Frame 1 and Frame 2 flashes in Figure 19. The former wave causes no problem of interpretation because it is a rightward motion composed of local right-motion signals. The latter wave, however, also is a rightward motion, but it is composed of local left-motion signals. We call such a motion wave a *virtual motion*.

Virtual motions seem to be generated only in apparent motion experiments. A continuously moving object does not cause a virtual motion, because it does not cause the two-sided local-motion signals that create gamma motion in response to a stationary flash. This latter fact provides a clue that suggests why virtual motions are not perceived; that is, why higher-motion detectors selectively code the preferred directions of their local-motion input signals.

We suggest that the receptive fields of these higher-order motion cells are adaptively tuned during development by experiences with continuously moving stimuli. During these experiences, the preferred local directions and global direction of motion tend to coincide. As a result of this statistical correlation, the excitatory and inhibitory receptive fields feeding the higher-order motion cells become asymmetrically distributed across space to favor motion in the locally preferred directions and to inhibit motion in the reverse directions.

At what level of processing does such asymmetric inhibition occur? In the motion BCS, the motion OC Filter is assumed to input to a CC Loop just as, in the static BCS, the static OC Filter inputs to a CC Loop (Figure 1). The cells which generate activations $x_i^{(R)}$ and $x_i^{(L)}$ at Level 5 in Figure 4 are analogous to the complex cells of the static OC Filter in Figure 2. A contrast-enhancing competition is assumed to occur at these complex cells (Grossberg, 1987b; Grossberg & Marshall, 1989), just as it is at Level 5 of the motion OC Filter.

The next stage of the static BCS is a level of hypercomplex cells whose receptive fields arise due to an on-center off-surround interaction among like-oriented cells (Figure 2) at the first competitive stage of the CC Loop. The lateral inhibition within the off-surround plays the role of an *end stopping* interaction. We assume that the corresponding off-surround in the motion BCS is asymmetrically organized across space. Consequently, motion is detected in the preferred directions, but it is inhibited in the opposite directions.

18. COMPARISON WITH OTHER MOTION MODELS: GENERALIZED REICHARDT, BURT AND SPERLING, AND NADEL MODELS

The hypothesis that asymmetric inhibition plays an important role in motion perception is a basic feature of the classical Reichardt model (Reichardt, 1957), the elaborated Reichardt model (van Santen & Sperling, 1984, 1985), and its equivalent formulations (Adelson & Bergen, 1985; Watson & Ahumada, 1985). The present model suggests that asymmetric inhibition plays an important role at a relatively high level of processing, as well as at the low level of processing that defines contrast-sensitive receptive fields, whereas an elaboration of the Marr–Ullman type of interaction is operative at an intermediate level of processing (Figure 4). The motion BCS architecture hereby suggests a framework for interpreting and unifying the results of heretofore separately developed motion models.

The motion BCS also clarifies the Burt and Sperling (1981) form factor in eqns (1–4). In particular, the temporal form factor S in (1) is interpreted in

terms of the growth and decay of the gated sustained-transient signals l_i and r_i described in (9) and (10) and illustrated in Figure 15. The property of space-time separability (3) is captured by the multiplication of sustained-transient signals with the Gaussian spatial kernel G_{ij} in (12) and (13). The property that the spatial interaction is finite at zero distance and decays with increasing distance is achieved by the form of the Gaussian kernel in (11). Competition between possible motion paths to select the pathway with greatest strength is achieved by the processes $x_i^{(R)}$ and $x_i^{(L)}$ in (14) and (15).

Seibert and Waxman (1989) have described a Neural Analog Diffusion-Enhancement Layer, or NADEL, model for motion perception that shares some important properties of the motion BCS, but also differs in several basic ways. One similarity occurs in their use of a contrast-enhancing network to sharpen the response to a broad spatial gradient of activation, as in our Level 5. Another similarity occurs in their realization that a mechanism is needed to generate a spatially broad reaction to a localized input flash. The differences occur in how this reaction is generated, and in different analyses of its significance.

In the NADEL model, each input generates an outward diffusion of activation which continues to spread until it dissipates. The contrast-enhancement helps to control the dissipative effect of the diffusion via a feedback pathway. In the motion BCS, a single input does not cause an outward diffusion; rather, it causes the spatially steady growth and decay of a Gaussian-distributed activation profile (Figure 6). Outward motion is produced, not by a single input, but by interaction between the temporally out-of-phase activations of two or more spatially steady Gaussians (Figure 8a). Seibert and Waxman use a nonlinear competitive feedback network to generate spatially sharp motion properties from spatially broad input patterns. They motivate this competitive network by comparison with a cooperative-competitive feedback network, called the CC Loop (Figure 1), which generates spatially sharp segmentations from OC Filter inputs within the static BCS (Grossberg & Mingolla, 1985b; Grossberg, 1987a). A CC Loop is also assumed to receive inputs from the motion OC Filter within the motion BCS (Grossberg, 1987a). Some properties of this interaction are outlined in the next section.

19. GLOBAL MOTION SEGMENTATION BY COOPERATIVE-COMPETITIVE FEEDBACK INTERACTIONS: NAKAYAMA–LOOMIS MODEL

The present article has clarified differences between visual perception of static form and of moving form by articulating computational differences between the static BCS and the motion BCS. Two key differences have thusfar been noted in the motion BCS: (a) coexistence of sensitivity to direction-of-motion and insensitivity to direction-of-contrast in the OC Filter; and (b) asymmetric inhibition at the first competitive stage of the CC Loop.

Both the static OC Filter and the motion OC Filter are assumed to input to a CC Loop for generating a coherent, context-sensitive boundary representation from the outputs of its OC Filter. As noted in Grossberg (1987a), the same CC Loop design is competent to generate coherent properties of both static form and moving form. Using the same CC Loop design for segmentation of static form and moving form clarifies how these disparate sources of visual information may be combined into a final representation of 3-D form.

For example, the CC Loop responds appropriately to moving patterns of random dots (Braddick, 1974; Julesz, 1971; Lappin & Bell, 1976; Nakayama, 1985; Nakayama, Silverman, MacLeod, & Mulligan, 1985; Nakayama & Tyler, 1981). As illustrated in Figures 1 and 4, an oriented receptive field of the CC Loop responds best if a dot density difference that is parallel to the preferred receptive field orientation moves in a direction perpendicular to the preferred receptive field orientation (Nakayama et al., 1985). Local random dot motions superimposed upon such a statistical drift are handled by CC Loop mechanisms. In particular, the second competitive stage of the CC Loop contains a competitive interaction between cells responding to *different* orientations at the *same* position. This competition is assumed to be strongest between perpendicular orientations. Such an orientational competition is capable of amplifying the locally-preferred combination of dot density orientation and direction-of-motion, while suppressing less preferred combinations. Then long-range cooperative interactions within the CC Loop can begin to group together those locally preferred combinations which are consistent across space into a coherent motion segmentation. The cooperative-competitive feedback within the CC Loop which amplifies the consistent grouping also actively suppresses locally preferred combinations which are not consistent with the global grouping.

Both the first and the second competitive stages of the CC Loop were originally derived to explain properties of static form perception (Grossberg, 1987a; Grossberg & Mingolla, 1985a, 1985b). Remarkably, both of these competitive stages are formally analogous to model mechanisms that Nakayama and Loomis (1974) have proposed for the extraction of a figural boundary moving relative to a background. These competitive mechanisms are also consistent with properties of motion-sensitive cells that Frost and Nakayama (1983) discovered

within the intermediate and deeper layers of the pigeon optic tectum. Figure 21 describes two versions of the Nakayama–Loomis model. In both versions, cells that respond to the same direction-of-motion interact via an on-center off-surround network. Given the hypothesis that the preferred direction-of-motion is perpendicular to a cell's preferred orientation, the Nakayama–Loomis model is consistent with the postulate that an on-center off-surround interaction exists among like-oriented cells, as in the first competitive stage of the CC Loop.

As shown in Figure 1, the static CC Loop is organized at the first competitive stage into on-cells and off-cells that are joined by opponent interactions (Grossberg, 1987a; Grossberg & Mingolla, 1985b). In the motion CC Loop, this property implies that opposite directions-of-motion are mutually inhibitory, as in the Nakayama–Loomis model of Figure 21b. In the static CC Loop, these opponent interactions help to explain phenomena such as negative aftereffects of radial patterns and binocular fusion and rivalry (Grossberg, 1987b). In the motion CC
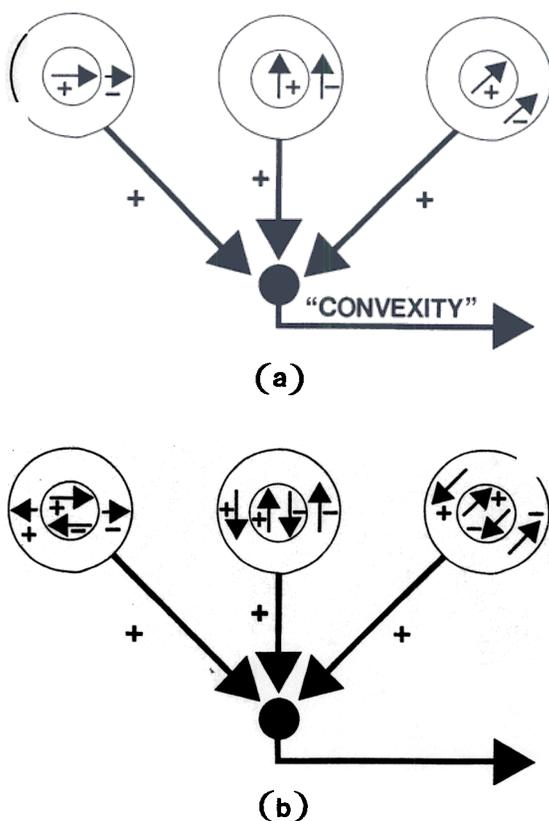
FIGURE 21. Variants of the Nakayama and Loomis (1974) model of "convexity" detecting units for extracting edges of 3-D objects for an observer translated through a rigid environment: (a) Units with like-preferred direction inhibit each other via an on-center off-surround interaction. All orientations at a given position summate their outputs to compute a "convexity" value that measures relative motion near figure-ground boundaries. (b) The model of (a) is augmented with opponent interactions between opposing directions of motion at each position.

Loop, they help to explain phenomena such as motion after effects, notably the waterfall illusion (Sekuler, 1975) and induced motion (Wallach, 1976).

The motion CC Loop also includes properties not posited by the Nakayama–Loomis model. An analog of the second competitive stage does not exist within the Nakayama–Loomis model. As noted above, its orientational competition is needed to suppress non-optimal local random motions and to select a best local direction-of-motion. Also missing in the Nakayama–Loomis model are long-range cooperative interactions which interact with the short-range competitive interactions to impart a coherent global motion representation upon discordant local motion signals, as in the phenomenon of motion capture (Ramachandran & Inada, 1985).

20. SHORT-RANGE VS. LONG-RANGE MOTION AND FORM-COLOR INTERACTIONS

There exist several distinct spatial scales within the motion BCS: the sizes of the sustained cell and transient cell receptive fields at Level 2 and Level 3, respectively, of the motion OC Filter; the breadth of the Gaussian filter from Level 4 to Level 5 of the motion OC Filter; the breadth of the first competitive stage of the CC Loop; and the breadth of the cooperative bipole cells of the CC Loop. Moreover, as in the static BCS, it is assumed that there exist multiple copies of the motion BCS network, each copy corresponding to a different receptive field size in the motion OC Filter. Subsequent spatial interactions within each copy are assumed to be related to receptive field size in a self-similar fashion (Grossberg, 1987b; Grossberg & Marshall, 1989). These relationships among spatial scales enable a variety of difficult motion properties to be explained.

To start this discussion, recall that Wertheimer (1912) made the color of the first flash different from the color of the second flash and found that observers reported that the flashes change color in flight. Van der Waals and Roelofs (1930, 1931), Squires (1931), and later Kolers and von Grünau (1975) confirmed these observations. These results highlight differences between BCS and FCS properties. They support the hypothesis that BCS interactions become independent of direction-of-contrast no later than Level 5 of Figure 4, thereby enabling flashes with different directions-of-contrast to interact across space, as in Figure 8, to generate a motion signal.

Such a motion signal within the BCS generates output signals to the FCS that define a boundary structure which contains color percepts within the FCS. Properties of the FCS clarify how the color can switch in mid-flight before the second flash is reached. These properties include organization of the FCS filter into double opponent color interac-

tions (Grossberg, 1987b), which clarifies how a switch between colors can occur at all, and filling-in of the winning opponent color within the moving boundary structure (Cohen & Grossberg, 1984; Grossberg &Todorović, 1988), which clarifies how the color of the flashes can be perceived at positions between the actual flash locations.

Interactions between the BCS and FCS also help to clarify perceived differences between beta motion and phi motion, since a motion signal can exist within the BCS without necessarily being able to support the full development of seen objects within the FCS.

Anstis and Mather (1985) have provided additional experimental support for the manner in which motion OC Filter circuit shown in Figure 4 becomes independent of direction-of-contrast. They studied the dependence of short-range motion between flashes (7.5 min arc) and longer-range motion between flashes. They also varied the direction-of-contrast of the flashes with respect to the background luminance. For short-range motion, the direction of motion depended upon brightness polarity, with motion only from white flash to white flash and black flash to black flash, as would be expected if successive flashes fell within individual filters at Level 2. For larger separations, motion could jump between white and black flashes, and conversely, as would be expected if successive flashes interacted via the Gaussian filter at Level 5.

This interpreation of the short range vs. long range motion dichotomy also provides an explanation of some results of Pantle and Picciano (1976) concerning Ternus displays. They found that elements of opposite contrast polarity could be matched in Ternus motion, but that interframe polarity changes always resulted in group motion percepts, regardless of the ISI. In the present model, interframe polarity changes will always activate transient mechanisms at all element locations and, thus, result in a group motion percept even when the ISI = 0.

Braddick (1974) originally reported that a constant scale size, D_{max}, controls all percepts of short-range motion. More recent experiments indicate that D_{max} increases with decreasing element density in the stimulus display (Lappin & Bell, 1976; Ramachandran & Anstis, 1983), with increasing field size (Baker & Braddick, 1982; Chang & Julesz, 1983), and with increasing element size (Petersik, Pufahl, & Krasnoff, 1983). These properties are clarified by observations that D_{max} varies with the spatial frequency content of the image (Burr, Ross, & Morrone, 1986; Nakayama & Silverman, 1984, 1985) and that receptive field size varies with spatial frequency (Anderson & Burr, 1987). All such results are consistent with the hypothesis that multiple copies of the motion BCS exist, each copy corresponding to a different receptive field size and subsequent interac-

tions within the copy are related to receptive field size by a property of self-similarity.

21. MULTIPLEXING OF MOTION DIRECTION AND MOTION DEPTH

In the static BCS (Figure 1), it has been shown how cells become binocular at the complex cell level (Grossberg, 1987b; Grossberg & Marshall, 1989). A similar hypothesis is made about the motion BCS; namely, that one role of the Gaussian filter is to combine motion signals from both eyes at the complex cells of Level 5 (Figure 4).

Such a scheme enables us to explain the fact that apparent motion may be perceived when the first flash excites one eye and the second flash excites the other eye (Gengerelli, 1948; Spigel, 1968). This property created great difficulty for early Gestaltist theories of apparent motion.

The Gaussian filter also provides an additional degree of freedom whereby cells at Level 5 can become sensitive to direction-of-motion over a wider range of stimulus orientations than cells at Level 2, whose preferred direction-of-motion is perpendicular to their preferred orientation.

This formal property may be compared with neurophysiological data which have shown that many cells in MT are sensitive to direction-of-motion over a range of stimulus orientations, whereas cells in V1 typically are sensitive to the direction-of-motion perpendicular to their orientational preference (Albright, 1984; Albright et al. 1984; Maunsell & van Essen, 1983). The organization of these cells into hypercolumns whose cells vary with respect to direction-of-motion, rather than orientation, is a subject for future research on the motion BCS design.

Also implicit in the design are properties which may be used to analyze such fundamental motion properties as induced motion, motion capture, and motion aftereffects. These properties will be analyzed in subsequent studies of the motion BCS.

In all, the motion BCS and static BCS architectures provide models as well as a theoretical framework for explaining and predicting an unusually large data base about the visual perception of static and moving form, and for interpreting, modifying, and unifying many of the other motion perception models in the vision research literature.

REFERENCES

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, 2(2), 284–299.

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

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

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

Anstis, S. M., & Mather, G. (1985). Effects of luminance and constrast on direction of ambiguous apparent motion. *Perception*, **14**, 167–179.

Baker, C. L., Jr., & Braddick, O. J. (1982). The basis of area and dot number effects in random dot motion perception. *Vision Research*, **22**, 1253–1259.

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

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)*, **290B**, 137–151.

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

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

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

Burt, P., & Sperling, G. (1981). Time, distance, and feature trade-offs in visual apparent motion. *Psychological Review*, **88**(2), 171–195.

Chang, J. J., & Julesz, B. (1983). Displacement limits for spatial frequency filtered random-dot cinematograms in apparent motion. *Vision Research*, **23**, 1379–1385.

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

De Valois, R. L., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, **22**, 545–559.

Enroth-Cugell, C., & Robson, J. G. (1966). The contrast sensitivity of retinal cells of the cat. *Journal of Physiology*, **187**, 517–552.

Frost, B. J., & Nakayama, K. (1983). Single visual neurons code opposing motion independent of direction. *Science*, **220**, 744–745.

Gengerelli, J. A. (1948). Apparent movement in relation to homonymous and heteronymous stimulation of the cerebral hemispheres. *Journal of Experimental Psychology*, **38**, 592–599.

Grossberg, S. (1970). Neural pattern discrimination. *Journal of Theoretical Biology*, **27**, 291–337.

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: Neural principles of learning, perception, development, cognition, and motor control*. Boston: Reidel Press.

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

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

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

Grossberg, S. (Ed.). (1988a). *Neural networks and natural intelligence*. Cambridge, MA: MIT Press.

Grossberg, S. (1988b). Nonlinear neural networks: Principles, mechanisms, and architectures. *Neural Networks*, **1**, 17–61.

Grossberg, S., & Marshall, J. (1989). Stereo boundary fusion by cortical complex cells: A system of maps, filters, and feedback networks for multiplexing distributed data. *Neural Networks*, **2**, 29–51.

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

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

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

Grossberg, S., & Rudd, M. E. (1989b). *A neural architecture for visual motion perception: Group and element apparent motion*. Proceedings of the International Joint Conference on Neural Networks, June 19, 1989, Washington, DC. Piscataway, NJ: IEEE.

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

Heggelund, P. (1981). Receptive field organization of simple cells in cat striate cortex. *Experimental Brain Research*, **42**, 89–98.

Higginson, G. D. (1926). Apparent visual movement and the Gestalt. *Journal of Experimental Psychology*, **9**, 228–252.

Hoffman, K.-P. (1973). Conduction velocity in pathways from retina to superior colliculus in the cat: A correlation with receptive field properties. *Journal of Neurophysiology*, **36**, 409–424.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, **160**, 106–154.

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, **195**, 215–243.

Hubel, D. H., & Wiesel, T. N. (1977). Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London (B)*, **198**, 1–59.

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

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

Kolers, P. A., & Pomerantz, J. R. (1971). Figural change in apparent motion. *Journal of Experimental Psychology*, **87**, 99–108.

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

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

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

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

Nakayama, K. (1985). Biological image motion processing. A review. *Vision Research*, **25**, 625–660.

Nakayama, K., & Loomis, J. M. (1974). Optical velocity patterns, velocity-sensitive neurons, and space perception: A hypothesis. *Perception*, **3**, 63–80.

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

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

Nakayama, K., Silverman, G. H., MacLeod, D. I. A., & Mulligan, J. (1985). Sensitivity to shearing and compressive motion in random dots. *Perception*, **14**, 225–238.

Nakayama, K., & Tyler, C. W. (1981). Psychophysical isolation of movement sensitivity by removal of familiar position cues. *Vision Research*, **21**, 427–433.

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

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

Orlansky, J. (1940). The effect of similarity and difference in form on apparent visual movement. *Archives of Psychlogy*, **246**.

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

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

Petersik, J. T., & Pantle, A. J. (1979). Factors controlling the competing sensations produced by a bistable stroboscopic display. *Vision Research*, **19**, 143–154.

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

Poggio, G. F., Motter, B. C., Squatrito, S., & Trotter, Y. (1985). Responses of neurons in visual cortex (V1 and V2) of the alert macaque to dynamic random-dot stereograms. *Vision Research*, **25**, 397–406.

Ramachandran, V. S., & Anstis, S. M. (1983). Displacement thresholds for coherent apparent motion in random dot patterns. *Vision Research*, **23**, 1719–1724.

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

Reichardt, W. (1957). Autokorrelation-sauswertung als Funktionsprinzip des Zentralnervensystems. *Z. Naturforsch*, **12b**, 447–457.

Seibert, M., & Waxman, A. M. (1989). Spreading activation layers, visual saccades, and invariant representations for neural pattern recognition systems. *Neural Networks*, **2**, 9–27.

Sekuler, R. (1975). Visual motion perception. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception, volume V: Seeing*. New York: Academic Press.

Spigel, I. M. (1968). Problems in the study of visually perceived movement: An introduction. In R. H. Haber (Ed.), *Contempory theory and research in visual perception* (pp. 103–121). New York: Holt, Rinehart, and Winston.

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

Stone, J. (1972). Morphology and physiology of the geniculocortical synapse in the cat: The question of parallel input to the striate cortex. *Investigative Ophthalmology*, **11**, 338–344.

Stone, J., & Dreher, B. (1973). Projection of X- and Y-cells of the cat's lateral geniculate nucleus to areas 17 and 18 of visual cortex. *Journal of Neurophysiology*, **36**, 551–567.

Tanaka, M., Lee, B. B., & Creutzfeldt, O. D. (1983). Spectral tuning and contour representation in area 17 of the awake monkey. In J. D. Mollon & L. T. Sharpe (Eds.), *Colour vision*. New York: Academic Press.

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

Thorell, L. G., De Valois, R. L., & Albrecht, D. G. (1984). Spatial mapping of monkey V1 cells with pure color and luminance stimuli. *Vision Research*, **24**, 751–769.

Tolhurst, D. J. (1973). Separate channels for the analysis of the shape and the movement of a moving visual stimulus. *Journal of Physiology*, **231**, 385–402.

van der Waals, H. G., & Roelofs, C. O. (1930 and 1931). Optische scheinbewegung. *Zeitschrift für Psychologie und Physiologie des Zinnesorgane*, **114**, 241–288 and **115**, 91–190.

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

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

Wallach, H. (1976). *On perception*. New York: Quadrangle/The New York Times Book Company.

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

Wertheimer, M. (1912). 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.

Zeki, S. M. (1974a). Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *Journal of Physiology (London)*, **236**, 549–573.

Zeki, S. M. (1974b). Cells responding to changing image size and disparity in the cortex of the rhesus monkey. *Journal of Physiology (London)*, **242**, 827–841.

APPENDIX: APPARENT MOTION THEOREMS

The results herein are drawn from an unpublished manuscript (Grossberg, 1977). These results demonstrated key properties of apparent motion, but waited to be published until the development of the static BCS clarified how they could be used to design a motion BCS. The results below concern interactions caused by two point flashes along a one-dimensional network of intervening cells.

Let flashes occur at positions $i = 0$ and $i = L$ of this network. Suppose, as in eqn (5), that

$$\frac{dx_0}{dt} = -Ax_0 + J_0 \tag{A1}$$

and

$$\frac{dx_L}{dt} = -Ax_L + J_L, \tag{A2}$$

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

$$x_0(t) = \int_0^t e^{-A(t-v)}J_0(v)\,dv, \tag{A3}$$

and

$$x_L(t) = \int_0^t e^{-A(t-v)}J_L(v)\,dv. \tag{A4}$$

Let input

$$J_0(t) = \begin{cases} J & \text{if } 0 \le t \le T \\ 0 & \text{if } T < t, \end{cases} \tag{A5}$$

and

$$J_L(t) = \begin{cases} J & \text{if } T + I \le t \le 2T + I \\ 0 & \text{if } 2T + I < t, \end{cases} \tag{A6}$$

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

$$x_0(t) = \frac{J}{A}(1 - e^{-AT})e^{-A(t-T)} \tag{A7}$$

and

$$x_L(t) = \frac{J}{A}(1 - e^{-A(t-T-l)}).$$ (A8)

Let $x_0(t)$ and $x_L(t)$ interact via a Gaussian filter

$$G_{ji} = \exp[-(j - i)^2/2K^2]$$ (A9)

as in eqn (11). 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

$$T(w, t) = x_0(t)\exp\left[\frac{-w^2}{2K^2}\right] + x_L(t)\exp\left[\frac{-(w - L)^2}{2K^2}\right]$$ (A10)

By (A7) and (A8),

$$T(w, t) = \frac{J}{A}\left[\left(1 - e^{-AT}\right)e^{-A(t-T)}\exp\left[\frac{-w^2}{2K^2}\right] \right.$$

$$\left. + (1 - e^{-A(t-T-l)})\exp\left[\frac{-(w - L)^2}{2K^2}\right]\right]$$ (A11)

Our first task is to show under what combinations of parameters the maximum value of $T(w, t)$ moves continuously from position $w = 0$ towards position $w = L$ through time.

Theorem 1 (Apparent Motion)

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

$$L < 2K.$$ (A12)

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

$$\frac{\partial T(w, t)}{\partial w} = 0.$$ (A13)

Such locations obey the equation

$$\frac{e^{A(t-T)} - e^{Al}}{1 - e^{-AT}} = \frac{w}{L - w}\exp\left[\frac{L(L - 2w)}{2K^2}\right]$$ (A14)

The function

$$f(t) = \frac{e^{A(t-T)} - e^{Al}}{1 - e^{-AT}}$$ (A15)

is an increasing function of t. We wish to determine when the postions $w = w(t)$ at which $T(w, t)$ is maximal increase as a function of t. In order for this to happen, the right hand side of (A14), namely function

$$g(w) = \frac{w}{L - w}\exp\left[\frac{L(L - 2w)}{2K^2}\right],$$ (A16)

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

$$w = g^{-1}(f(t))$$ (A17)

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

Function $g(w)$ is monotone increasing if $g'(w) > 0$, which holds if and only if function

$$h(w) \equiv (L - w)\left[1 - \frac{Lw}{K^2}\right] + w$$ (A18)

satisfies

$$h(w) > 0.$$ (A19)

In order for (A17) to hold for all $0 \le w \le L$, the minimum of $h(w)$ for $0 \le w \le L$ must be positive. The minimum of $h(w)$ occurs at $w = L/2$, and equals

$$h\left(\frac{L}{2}\right) = \frac{L}{2}\left(2 - \frac{L^2}{2K^2}\right).$$ (A20)

The number $h(L/2)$ is positive if (A12) holds.
The next result proves that the apparent motion signal reaches

the position $w = L/2$ midway between positions $w = 0$ and w L at a time t_1 that is independent of L and K.

Theorem 2 (Equal Half-Time Property)

The time at which the motion signal reaches position $w = L/2$ is

$$t_1 = T + \frac{1}{A}\ln[e^{Al} + (1 - e^{-AT})].$$ (A21)

Proof: By (A17), we need to compute $t = f^{-1}(g(w))$ when $w = L/2$, namely

$$t_1 = f^{-1}\left(g\left(\frac{L}{2}\right)\right).$$ (A22)

By (A16),

$$g\left(\frac{L}{2}\right) = 1.$$ (A23)

Equation (A21) follows immediately from (A23) and (A15).
The next result quantifies how the motion accelerates and decelerates by comparing the times t_1, t_1, and t_1 at which the motion signal attains positions $w = L/4$, $w = L/2$, and $w = 3L/4$, respectively.
Interestingly, although time t_1 is independent of L, times t_1 and t_1 are not.

Theorem 3 (Accelerate-Decelerate)

$$t_1 - t_1 \ge t_1 - t_1.$$ (A24)

Proof: By (A17),

$$t_1 = f^{-1}\left(g\left(\frac{L}{4}\right)\right)$$ (A25)

and

$$t_1 = f^{-1}\left(g\left(\frac{3L}{4}\right)\right)$$ (A26)

By (A16),

$$g\left(\frac{L}{4}\right) = \frac{1}{3}\exp\left(\frac{L^2}{4K^2}\right)$$ (A27)

and

$$g\left(\frac{3L}{4}\right) = 3\exp\left(\frac{-L^2}{4K^2}\right)$$ (A28)

By (A15),

$$t_1 = T + \frac{1}{A}\ln\left[e^{Al} + (1 - e^{-AT})g\left(\frac{L}{4}\right)\right]$$ (A29)

and

$$t_1 = T + \frac{1}{A}\ln\left[e^{Al} + (1 - e^{-AT})g\left(\frac{3L}{4}\right)\right]$$ (A30)

Thus by (A30) and (A21),

$$t_1 - t_1 = \frac{1}{A}\ln\left[\frac{e^{Al} + (1 - e^{-AT})g\left(\frac{3L}{4}\right)}{e^{Al} + (1 - e^{-AT})}\right]$$ (A31)

By (A29) and (A21),

$$t_1 - t_1 = \frac{1}{A}\ln\left[\frac{e^{Al} + (1 - e^{-AT})}{e^{Al} + (1 - e^{-AT})g\left(\frac{L}{4}\right)}\right]$$ (A32)

Inequality (A24) is thus equivalent to inequality

$$(1 - e^{-AT})\left[1 - g\left(\frac{L}{4}\right)g\left(\frac{3L}{4}\right)\right]$$

$$\le e^{Al}\left[g\left(\frac{L}{4}\right) + g\left(\frac{3L}{4}\right) - 2\right]$$ (A33)

By (A27) and (A28),

$$g\left(\frac{L}{4}\right)g\left(\frac{3L}{4}\right) = 1 \qquad (A34)$$

Hence (A33) reduces to

$$\frac{1}{3}\exp\left[\frac{L^2}{4K^2}\right] + 3\exp\left[\frac{-L^2}{4K^2}\right] \geq 2, \qquad (A35)$$

which is true for all positive values of L and K. In fact,

$$t_1 - t_1 = t_1 - t_1 \qquad (A36)$$

only if

$$\frac{L^2}{4K^2} = \ln(3), \qquad (A37)$$

which corresponds to a limiting case in which

$$t_1 = t_1 = t_1. \qquad (A38)$$

Similar results may be proved when a single pair of flashes is replaced by a cyclic alternation of flashes. One such result shows that the furthest extent of the motion between the flashes may decrease to a limiting value ("motion contraction") on successive flash cycles.