

1991-04

# Cortical Dynamics of Visual Motion Perception: Short-Range and Long-Range Apparent Motion

---

<https://hdl.handle.net/2144/2073>

*"Downloaded from OpenBU. Boston University's institutional repository."*

**CORTICAL DYNAMICS OF VISUAL MOTION PERCEPTION:  
SHORT-RANGE AND LONG-RANGE APPARENT MOTION**

Stephen Grossberg and Michael E. Rudd

April, 1991

Technical Report CAS/CNS-91-018

Permission to copy without fee all or part of this material is granted provided that: 1. the copies are not made or distributed for direct commercial advantage, 2. the report title, author, document number, and release date appear, and notice is given that copying is by permission of the BOSTON UNIVERSITY CENTER FOR ADAPTIVE SYSTEMS AND DEPARTMENT OF COGNITIVE AND NEURAL SYSTEMS. To copy otherwise, or to republish, requires a fee and/or special permission.

Copyright © 1991

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



**CORTICAL DYNAMICS OF VISUAL MOTION PERCEPTION:  
SHORT-RANGE AND LONG-RANGE APPARENT MOTION**

Stephen Grossberg†  
Center for Adaptive Systems  
and  
Department of Cognitive & Neural Systems  
Boston University  
111 Cummington Street  
Boston, MA 02215  
and  
Michael E. Rudd‡  
Department of Psychology  
Johns Hopkins University  
Baltimore, MA 21218

August 1990  
Revised April 1991

Correspondence concerning this article should be addressed to:  
Professor Stephen Grossberg  
Center for Adaptive Systems, Room 244  
Boston University  
111 Cummington Street  
Boston, MA 02215

---

† Supported in part by the Air Force Office of Scientific Research (AFOSR 90-0175), the Army Research Office (ARO DAAL-03-88-K0088), DARPA (AFOSR-90-0083), and Hughes Aircraft Company (S1-903136).

Acknowledgements: We wish to thank Cynthia E. Bradford, Kelly Dumont and Carol Yanakakis Jefferson for their valuable assistance in the preparation of the manuscript.

‡ Supported in part by the Army Research Office (ARO DAAL-03-88-K0088).

## TABLE OF CONTENTS

1. WHY DO PARALLEL CORTICAL SYSTEMS EXIST FOR THE PERCEPTION OF STATIC FORM AND MOVING FORM? . . . . .	1
2. THE MOC FILTER: JOINING SENSITIVITY TO DIRECTION-OF-MOTION WITH INSENSITIVITY TO DIRECTION-OF-CONTRAST . . . . .	4
3. APPARENT MOTION AS A PROBE OF NEURAL MOTION MECHANISMS . . . . .	6
4. SPATIOTEMPORAL PARAMETERS FOR GENERATING APPARENT MOTION . . . . .	8
5. SPACE-TIME SEPARABILITY OF THE MOTION STRENGTH FUNCTION . . . . .	9
6. THE SHAPE OF THE SPATIAL COMPONENT OF THE MOTION STRENGTH FUNCTION AND ITS DEPENDENCE ON ELEMENT SIZE . . . . .	12
7. THE TEMPORAL MOTION STRENGTH FUNCTION AND ITS DEPENDENCE ON FLASH DURATION . . . . .	15
8. THE INSENSITIVITY OF THE MOTION CORRESPONDENCE PROCESS TO FIGURAL IDENTITY . . . . .	18
9. GROUP AND ELEMENT APPARENT MOTION: TERNUS DISPLAYS . . . . .	20
10. MOTION VERSUS VISUAL PERSISTENCE . . . . .	22
11. DESIGN OF A MOC FILTER . . . . .	22
12. GAMMA MOTION: THE APPARENT EXPANSION OF A SPOT AT ONSET AND ITS CONTRACTION AT OFFSET . . . . .	29
13. CONTINUOUS MOTION PATHS GENERATED BY STATIONARY FLASHES . . . . .	31
14. RELATIONSHIP BETWEEN FLASH SPATIAL SEPARATION AND SPATIAL SCALE: MOTION SPEED-UP AND PARTIAL MOTIONS . . . . .	35
15. EQUAL TIME MULTIPLE SCALE MOTIONS . . . . .	38
16. 'THE LESS YOU SEE IT, THE FASTER IT MOVES' . . . . .	38
17. SPLIT APPARENT MOTION . . . . .	41
18. TERNUS DISPLAY: GROUP MOTION . . . . .	41
19. TERNUS DISPLAY: ELEMENT MOTION AND TRANSIENT CELL GATING OF SUSTAINED CELLS . . . . .	43
20. SIMULATING THE TRANSITION BETWEEN ELEMENT AND GROUP MOTION: SAME DIRECTION-OF-CONTRAST . . . . .	45
21. A CRUCIAL TEST: SIMULATING GROUP MOTION AT SHORT ISI'S WITH REVERSE CONTRAST STIMULI . . . . .	47
22. DELTA MOTION: MOTION FROM THE SECOND FLASH TO THE FIRST FLASH . . . . .	49
23. APPARENT MOTION THRESHOLDS: THE JOINT EFFECTS OF SPATIAL SEPARATION, FLASH DURATION, AND INTERSTIMULUS INTERVAL . . . . .	50
24. SHUNTING CASCADES AND HABITUATING TRANSMITTER GATES . . . . .	53
25. THE MOTION STRENGTH FUNCTION OF FLASH DURATION, ISI, AND FIGURAL IDENTITY . . . . .	59
26. TERNUS DISPLAY: THE EFFECTS OF FLASH DURATION AND ELEMENT SIZE . . . . .	62
27. SHORT-RANGE VS. LONG-RANGE MOTION AND FORM-COLOR INTERACTIONS . . . . .	64
28. MULTIPLEXING OF MOTION DIRECTION AND MOTION DEPTH . . . . .	65
29. CONCLUDING REMARKS: TOWARDS A UNIFIED THEORY OF BIOLOGICAL VISION . . . . .	65
REFERENCES . . . . .	67
APPENDIX: DEPENDENCE OF THRESHOLD ISI AND SOA ON FLASH SEPARATION AND DURATION . . . . .	75

## ABSTRACT

This article describes further evidence for a new neural network theory of biological motion perception. The theory clarifies why parallel streams  $V1 \rightarrow V2$ ,  $V1 \rightarrow MT$ , and  $V1 \rightarrow V2 \rightarrow MT$  exist for static form and motion form processing among the areas  $V1$ ,  $V2$ , and  $MT$  of visual cortex. The theory suggests that the static form system (Static BCS) generates emergent boundary segmentations whose outputs are insensitive to direction-of-contrast and insensitive to direction-of-motion, whereas the motion form system (Motion BCS) generates emergent boundary segmentations whose outputs are insensitive to direction-of-contrast but sensitive to direction-of-motion. The theory is used to explain classical and recent data about short-range and long-range apparent motion percepts that have not yet been explained by alternative models. These data include beta motion; split motion; gamma motion and reverse-contrast gamma motion; delta motion; visual inertia; the transition from group motion to element motion in response to a Ternus display as the interstimulus interval (ISI) decreases; group motion in response to a reverse-contrast Ternus display even at short ISIs; speed-up of motion velocity as interflash distance increases or flash duration decreases; dependence of the transition from element motion to group motion on stimulus duration and size; various classical dependencies between flash duration, spatial separation, ISI, and motion threshold known as Korte's Laws; dependence of motion strength on stimulus orientation and spatial frequency; short-range and long-range form-color interactions; and binocular interactions of flashes to different eyes.

**Key Words:** vision, neural network, motion perception, visual cortex, Boundary Contour System, apparent motion, form perception, MT, V1, V2.

## 1. Why Do Parallel Cortical Systems Exist for the Perception of Static Form and Moving Form?

How do we see things move? Our everyday percepts of moving objects are so immediate and compelling that the synthetic nature of the perceptual processes which generate these percepts are not easily understood. The task of rapidly detecting a leopard leaping from a jungle branch under a sun-dappled forest canopy illustrates the subtlety and vigor of these processes. Consider how spots on the leopard's coat move as its limbs and muscles surge. Imagine how patterns of light and shade play upon the leopard's coat as it leaps through the air. These luminance and color contours move across the leopard's body in a variety of directions that do not necessarily point in the direction of the leopard's leap. Indeed, the leopard's body generates a scintillating mosaic of moving contours that could easily prevent its detection. Remarkably, our perceptual processes are able to actively reorganize such a scintillating mosaic into a coherent object percept with a unitary direction-of-motion. The leopard as a whole then seems to quickly "pop out" from the jungle background and to draw our attention. Such a perceptual process clearly has a high survival value for animals who possess it.

This description of the leaping leopard emphasizes that the process of motion perception is an active one. It is capable of transforming a motion signal that is generated by a luminance contour into a different motion percept. In this sense, our percepts of moving objects are often percepts of *apparent motion*, albeit an adaptive and useful form of apparent motion. The task of understanding how we see "real" motion thus requires that we also understand "apparent" motion. The present article explains a large body of classical and recent data about apparent motion to further support a new theory of motion perception that was described in Grossberg and Rudd (1989). Most of these data have not yet been explained by alternative theories of motion perception.

This new theory of motion perception grew out of an earlier theory of static form perception (Cohen and Grossberg, 1984; Grossberg, 1983, 1987a, 1987b; Grossberg and Marshall, 1989; Grossberg and Mingolla, 1985a, 1985b, 1987; Grossberg, Mingolla, and Todorović, 1989; Grossberg and Todorović, 1988). A key new insight of the static form theory can be summarized by the paradoxical phrase that "all boundaries are invisible". An illustration of this property is provided by the percept of a reverse-contrast Kanizsa square (Figure 1), whose significance for perceptual psychology was first emphasized by Cohen and Grossberg (1984), Grossberg and Mingolla (1985a, 1985b), Prazdny (1983), and Shapley and Gordon (1985). In this percept, a square boundary emerges between the four pac man inducers. The vertical components of this boundary join together dark-light vertical contrasts with light-dark vertical contrasts. Thus the boundaries can form between opposite directions-of-contrast. Another way of saying this is that the output of the boundary completion process is insensitive to direction-of-contrast, even though it is sensitive to amount-of-contrast. A process whose output does not distinguish between dark-light and light-dark cannot carry a visible signal. Hence "all boundaries are invisible."

This boundary completion process has been called the Boundary Contour System, or BCS, in order to emphasize that its boundaries emerge from contrast-sensitive processes. The boundaries formed by the BCS are not created only in response to edges. Rather, they may be generated in response to combinations of edge, texture, shading, and stereo information at multiple size scales. That is why the term "boundary completion" rather

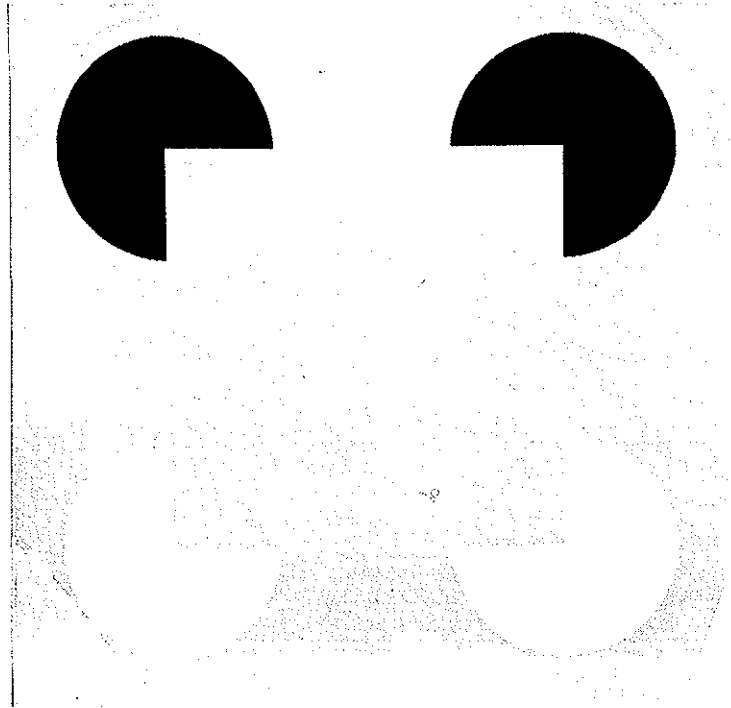


Figure 1. A reverse-contrast Kanizsa square.

than “edge detection” is used. These form-sensitive boundary structures have been called *boundary webs* by Grossberg (1987a) and Grossberg and Mingolla (1987).

Since the BCS does not represent visible percepts, another process than boundary completion must also exist that does generate visible percepts. This process has been suggested to discount the illuminant, or compensate for variable illumination conditions, and to fill-in surface properties of brightness, color, and depth using the discounted signals. It has been called the Feature Contour System, or FCS, because it generates the visible percepts that scientists had earlier attributed to “feature detectors,” and it does so using a contrast-sensitive process.

What is the relationship between the contrast-sensitive processes of the BCS and the FCS? Remarkably, these processes obey laws that are computationally complementary (Grossberg, Mingolla, and Todorović, 1989). Figure 2 summarizes three of the dimensions along which BCS and FCS processes are complementary. The BCS and FCS overcome the limitations of their complementary processes by interacting with one another through both serial and parallel pathways undergoing both feedforward and feedback interactions (Grossberg, 1987a, 1987b). These interactions give rise to a visual representation that is called a FACADE representation because it suggests how properties of Form-And-Color-And-Depth are combined in a visual percept. The theory that explains how BCS and FCS interactions generate these representations is called FACADE Theory.

FACADE representations are predicted to occur in prestriate area  $V4$  of the visual cortex (Grossberg, 1987b). More generally, BCS and FCS processes have been used to explain and predict perceptual and neurobiological data about the regions  $V1$ ,  $V2$ , and  $V4$  of visual cortex, notably the cortical stream  $V1 \rightarrow V2 \rightarrow V4$  that has been linked to perceptual

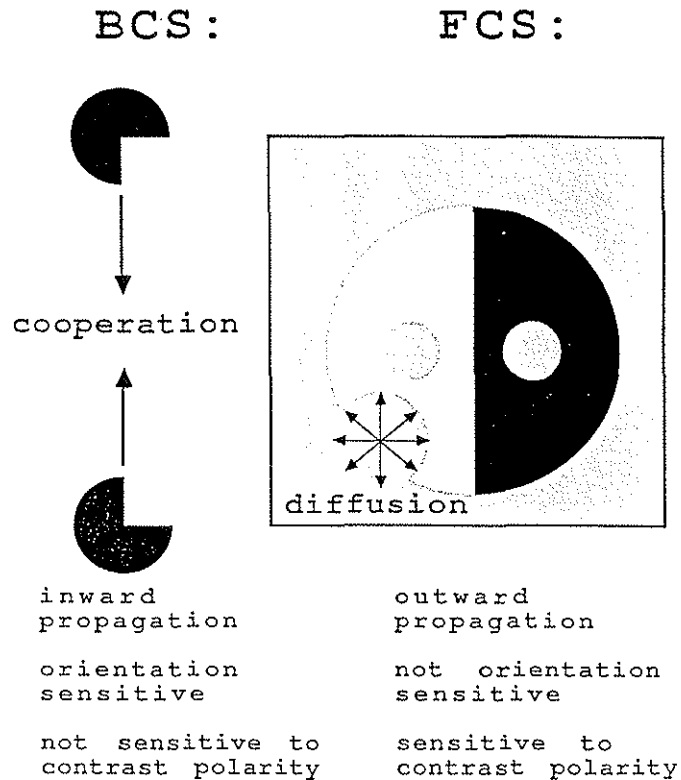


Figure 2. Some complementary computational properties of the Boundary Contour System and the Feature Contour System.

properties of static form, color, and depth. In keeping with these properties, the BCS is called the *Static BCS* in order to differentiate it from the *Motion BCS* that is the subject of this article.

Indeed, a parallel cortical stream  $V1 \rightarrow MT$  exists from cortical area  $V1$  to area  $MT$ . Cells in area  $MT$  are sensitive to properties of motion (Albright, 1984; Albright, Desimone, and Gross, 1984; Maunsell and van Essen, 1983; Newsome, Gizzi, and Movshon, 1983). Why has Nature needed to evolve parallel cortical streams  $V1 \rightarrow V2$  and  $V1 \rightarrow MT$  for the processing of static form and moving form? This is a nontrivial question if only because the first processing stage in  $V1$ , the simple cells, are already sensitive to direction-of-motion and to changes in stimulus intensity (DeValois, Albrecht, and Thorell, 1982; Heggelund, 1982; Hubel and Wiesel, 1962, 1968, 1977; Tanaka, Lee, and Creutzfeldt, 1983). Why has evolution needed to generate region  $MT$  when even the simple cells of  $V1$  are already direction-sensitive and change-sensitive? What computational properties are achieved by  $MT$  that are not already available in  $V1$  and its prestriate projections  $V2$  and  $V4$ ?

A precise answer to this question has come into view through an analysis of why the Static BCS is not adequate for motion processing (Grossberg, 1987b, 1991). This inadequacy of the Static BCS is a consequence of the fact that "all boundaries are invisible." The scientific explication of this paradoxical statement has, in fact, forced a pervasive shift in theoretical perspective that underlies much of the enhanced explanatory power of FACADE Theory.

In order to understand why the Static BCS is inadequate for motion processing, we review below how the process which makes the output signals of the Static BCS insensitive to direction-of-contrast also makes them insensitive to direction-of-motion. A perceptual

system whose output is insensitive to direction-of-motion is certainly not well suited to be a motion processor.

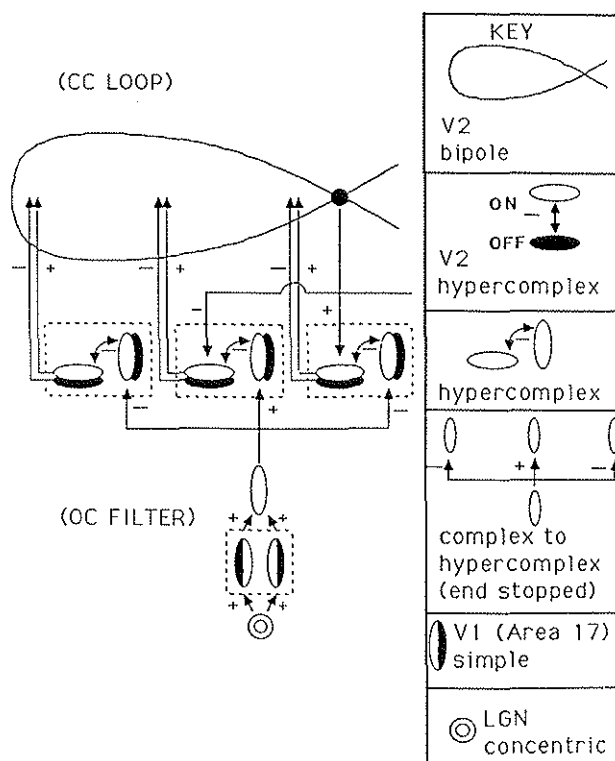
This observation led us to the following theoretical question: What is the minimal change of the Static BCS with which to fashion a Motion BCS whose output signals are insensitive to direction-of-contrast—which is just as important for processing static images as moving images—yet are sensitive to direction-of-motion? The Motion BCS that was hereby derived has been used here and elsewhere (Grossberg, 1991; Grossberg and Mingolla, 1990a, 1990b, 1990c; Grossberg and Rudd, 1989a, 1989b, 1989c) to explain a large data base about motion perception. In addition, as a result of this approach, the Static BCS and the Motion BCS can be viewed as variations of one another. Prior to this observation, data about static perception and motion perception had typically been studied as parts of separate scientific enterprises. The present synthesis allows them to be explained as variations of a common design for the architecture of visual cortex.

Grossberg (1991) has further developed this theme by predicting that the Static BCS and Motion BCS are parallel subsystems of a single total BCS system. This prediction suggests that this total BCS system arises during cortical development as an expression of a global symmetry principle, called FM Symmetry (F=form, M=motion). Manifestations of this symmetry principle are familiar to us in our daily perceptual experiences. Of special interest is the theory's explanation of why the geometries of static and motion form perception differ; for example, why the opposite orientation of a static vertical is a static horizontal—a difference of 90 degrees—whereas the opposite direction of motion upward is motion downward—a difference of 180 degrees.

## **2. The MOC Filter: Joining Sensitivity to Direction-of-Motion with Insensitivity to Direction-of-Contrast**

This section summarizes how the output of the Static BCS becomes insensitive to direction-of-motion due to the interactions that render it insensitive to direction-of-contrast. We noted in Section 1 that simple cells of the visual cortex, which are modelled as the earliest stage of the Static BCS, are sensitive to direction-of-contrast. The Static BCS provides a new computational rationale, as well as a model of the neural circuits governing classical cortical cell types such as simple cells, complex cells, and hypercomplex cells. The theory also predicted the existence of a new type of cell, called a bipole cell (Cohen and Grossberg, 1984; Grossberg, 1984; Grossberg and Mingolla, 1985a) whose properties have been supported by subsequent neurophysiological experiments (von der Heydt, Peterhans, and Baumgartner, 1984; Peterhans and von der Heydt, 1989).

The output of the Static BCS becomes independent of direction-of-contrast due to the way in which its model simple cells combine their outputs to model complex cells. As indicated in Figure 3, each simple cell has an oriented receptive field, represented by an elliptical region. In the illustrated receptive fields, inputs to the white side add up to activate the cell, whereas inputs to the black side add up to inhibit the cell. The cell fires, or emits an output signal, only if the net cell activity exceeds a threshold. Thus the output signal is rectified. Further increments in cell activity give rise to proportional increments in the output signal. As a result of this rule, the model simple cells are sensitive to direction-of-contrast. In particular, a cell with a white-black vertical receptive field can fire to a white-black (nearly) vertical image contrast, but not to a black-white contrast.

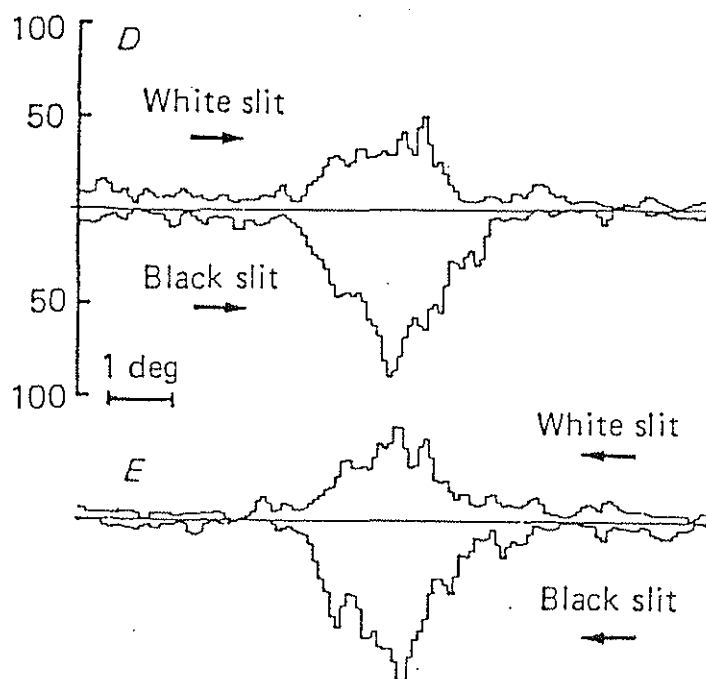


**Figure 3.** The Static Boundary Contour System circuit described by Grossberg and Mingolla (1985b). The circuit consists of an oriented contrast-sensitive filter (OC filter) followed by a cooperative-competitive feedback network (CC Loop). Multiple copies of this circuit are used, one corresponding to each receptive field size of the OC Filter. In the present paper, we refer to the OC Filter as a SOC Filter (for Static OC Filter), in order to distinguish it from the analogous Motion OC (MOC) Filter described in the present paper. 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 and Marshall, 1989).

Figure 3 illustrates how these simple cells combine their output signals to activate complex cells that are sensitive to the *amount* of image contrast, but not to the *direction* of image contrast. In particular, a pair of vertically oriented simple cells are shown inputting to a single complex cell. The complex cell response is insensitive to direction-of-contrast because it adds output signals from a pair of simple cells which are sensitive to opposite directions-of-contrast.

This construction also renders the complex cells insensitive to direction-of-motion. Inspection of Figure 3 shows that a vertically oriented model complex cell could respond, say, to a black-white vertical edge moving to the right or left *and* to a white-black vertical edge moving to the right or left. Thus the process whereby complex cells become insensitive to direction-of-contrast has rendered them insensitive to direction-of-motion. This combination of properties of cortical complex cells has been reported by several laboratories. Figure 4 summarizes data of Foster, Gaska, Nagler, and Pollen (1985, also discussed in Pollen, Gaska and Jacobsen, 1989) that illustrate both properties. Our construction of the Motion BCS focuses upon how oriented receptive fields that are sensitive to direction-of-contrast, such as those of simple cells, can be combined to give rise to cells that are not sensitive to direction-of-contrast, such as complex cells, without causing these cells to lose their sensitivity to direction-of-motion.

## Complex cell



**Figure 4.** Responses of a complex cell to a thin bar drifting first in one direction and then in the opposite direction. The response to the light bar is plotted above the dividing line. The response to the dark bar is plotted below the dividing line. Reprinted from Foster, Gaska, Nagler, and Pollen (1985) with permission.

Figure 3 illustrates that the Static BCS consists of two successive subsystems: an oriented contrast-sensitive filter (OC Filter) and a cooperative-competitive feedback network (CC Loop). The OC Filter contains the processing stages whereby simple cells interact with complex cells and hypercomplex cells. We call this OC Filter a Static OC Filter, or SOC Filter, to distinguish it from the Motion OC Filter, or MOC Filter, that is analysed herein. The SOC Filter preprocesses quasi-static images (the eye never ceases to jiggle in its orbit), whereas the CC Loop performs a grouping function that generates boundary webs in response to the filtered signals. The Motion BCS also includes a CC Loop, which is discussed in Grossberg (1987b), Grossberg and Mingolla (1990a, 1990b, 1990c), and Grossberg and Rudd (1989c). Only the MOC Filter of the Motion BCS will be considered herein. The modification of the SOC Filter which leads to a MOC Filter introduces an extra degree of computational freedom into the Filter that achieves several important properties at once. These properties are summarized in Table 1. All of these properties will be discussed below.

TABLE 1

The remainder of the article describes the apparent motion data that will be analyzed herein and the MOC Filter properties that will be used to explain them.

### 3. Apparent Motion as a Probe of Neural Motion Mechanisms

*Apparent motion* is a label that was given by Gestalt psychologists in the first half of this century to the percept of motion generated by a display in which nothing actually moves. For example, two briefly displayed flashes of light separated by the proper spatiotemporal

interval will result in a compelling illusion of movement between the two flashes. Experiments designed to reveal the nature of the underlying process have been carried out for over 100 years, beginning with the accidental discovery of the basic phenomenon by the physiologist Exner in 1875 (Exner, 1875; Boring, 1950). In light of this fact, it is perhaps surprising that there is still no generally accepted model of the neural mechanisms responsible for even the most basic of apparent motion phenomena.

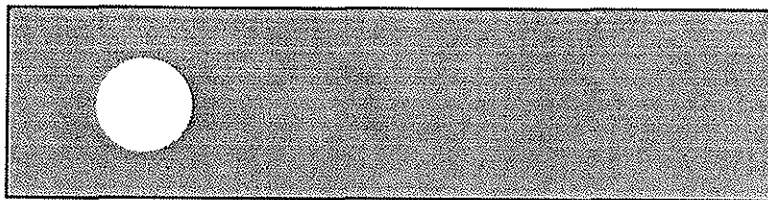
Experimental variants of apparent motion abound in the literature. Early investigators chose to label many of these with letters of the Greek alphabet. Thus, a 'figureless' or 'objectless' motion that is observed to occur between the two flashes when the spatiotemporal parameters of the display are suboptimal is referred to as *phi motion*, or the *phi phenomenon*; the smooth and continuous movement of a perceptually well-defined form is called *beta motion*; a reverse motion which occurs when the luminance of the second flash is much brighter than that of the first is called *delta motion*; and the apparent expansion at onset of a single flash, or its contraction at offset, is referred to as *gamma motion* (Bartley, 1941; Kolers, 1972).

None of these phenomena have been satisfactorily explained by other theories. Any satisfactory theory needs to explain how a long-range spatial influence is generated by each flash, but only triggers a motion signal when at least two flashes are presented. It also needs to explain why the long-range influence of a single flash is not perceptually visible. For example, why are not waves of motion-carrying signals observed to propagate outward from a single flash? It further needs to explain the fact that the motion signal generated by an apparent motion display is perceived to speed up in order to interpolate between flashes that are presented at a larger spatial separation, but at the same time interval; or between flashes that are separated by the same distance but by a shorter time lag (Kolers, 1972). How does a motion signal know how to calibrate its speed to match the variable distance or ISI between a pair of successive flashes? It cannot begin to do so until after the first flash has ended and the second flash has begun. Yet the first flash, by itself, provides no observable evidence that it can generate a motion signal capable of traversing perceptual space. The present article provides an answer to these perplexing questions using just a few processes, each of which is surprisingly simple. The model as a whole illustrates how interactions among simple processes can give rise to emergent properties of considerable subtlety.

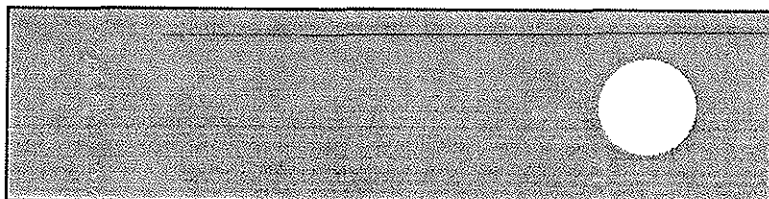
Indeed, the very large psychological literature that exists on the topic of apparent motion—and the more general category of motion perception—indicates a complex interdependency between such stimulus variables as contrast, size, duration, color, and figural organization in determining the perceived motion. In addition, it is clear that the neural networks which compute motion do not exist in isolation from those which are concerned with the extraction of other information from the visual stimulus. Instead, various types of visual analysis are multiplexed by the nervous system. Therefore, the more general problem in understanding motion perception is to discover not just how the brain computes motions, but rather how these computations are embedded in a process of generating a 3-D representation of moving objects.

Due to the difficulty of such a task, the construction of a satisfactory neural model of motion perception is a challenging theoretical problem and probably will remain so for some time. In the present paper we first summarize results from diverse experimental studies which serve to illustrate the complex interrelationship between some of the stimulus factors that

FRAME 1



FRAME 2



**Figure 5.** Two-flash apparent motion display. In the first frame, a single spot is presented, followed after an interstimulus interval (ISI) by the presentation of a spot at a different location in the second frame. When the ISI is small, the two flashes appear to be simultaneous. At longer ISIs, continuous motion from the position of the first spot to that of the second is observed. At still longer ISIs, the spots are perceived to turn on and off in succession, with no accompanying perception of movement.

are known to influence the organization of motion percepts. Then we define the MOC Filter and demonstrate its ability to account for these empirical results.

#### 4. Spatiotemporal Parameters for Generating Apparent Motion

It is instructive to begin our analysis of apparent motion by considering the case of the simple two-flash apparent motion display illustrated in Figure 5. A spot of light on a dark background (or vice versa) is displayed for some duration  $SD_1$ , followed first by a blank *interstimulus interval* (hereafter referred to as the ISI), then by the appearance for a duration  $SD_2$  of a second spot of light in a different location. In some experimental paradigms this whole display sequence may be cycled many times, with a second ISI possibly being inserted after the second flash in the cycle. In this case, we would have the continuously cycling pattern  $SD_1 - ISI_1 - SD_2 - ISI_2 : SD_1 - ISI_1 - SD_2 - ISI_2 : SD_1 - \dots$ , etc. In the discussion that follows we will assume for simplicity that  $SD_2 = SD_1$ , and  $ISI_2 = ISI_1$  (if there is an  $ISI_2$ ). Therefore, we will simply use the abbreviations SD and ISI to signify the frame duration and interstimulus interval, respectively, of the display.

In the apparent motion literature, the delay between the two flashes is sometimes alternatively expressed in terms of the onset-to-onset interval, or *stimulus onset asynchrony* (SOA), instead of in terms of the ISI. It has proven difficult to experimentally determine whether it is the SOA or the ISI that is the critical parameter in determining either the probability or the quality of the motion percept generated by an apparent motion display.

For example, given any particular spatial separation of flashes in a two-flash display that effectively induces a motion percept, there will be a restricted range of SOAs over which this percept can be generated. When the SOA of the display is smaller than the minimum SOA required to produce motion, the subject reports that the two spots appear to blink on and off in place. In this case, it may be difficult for the subject to determine the phase relationship between the flashes. When the SOA of the display is greater than the maximum SOA for motion, the subject reports that the lights turn on and off successively. In this latter case, the subject has no problem determining the temporal order of the flashes; however, no motion is seen.

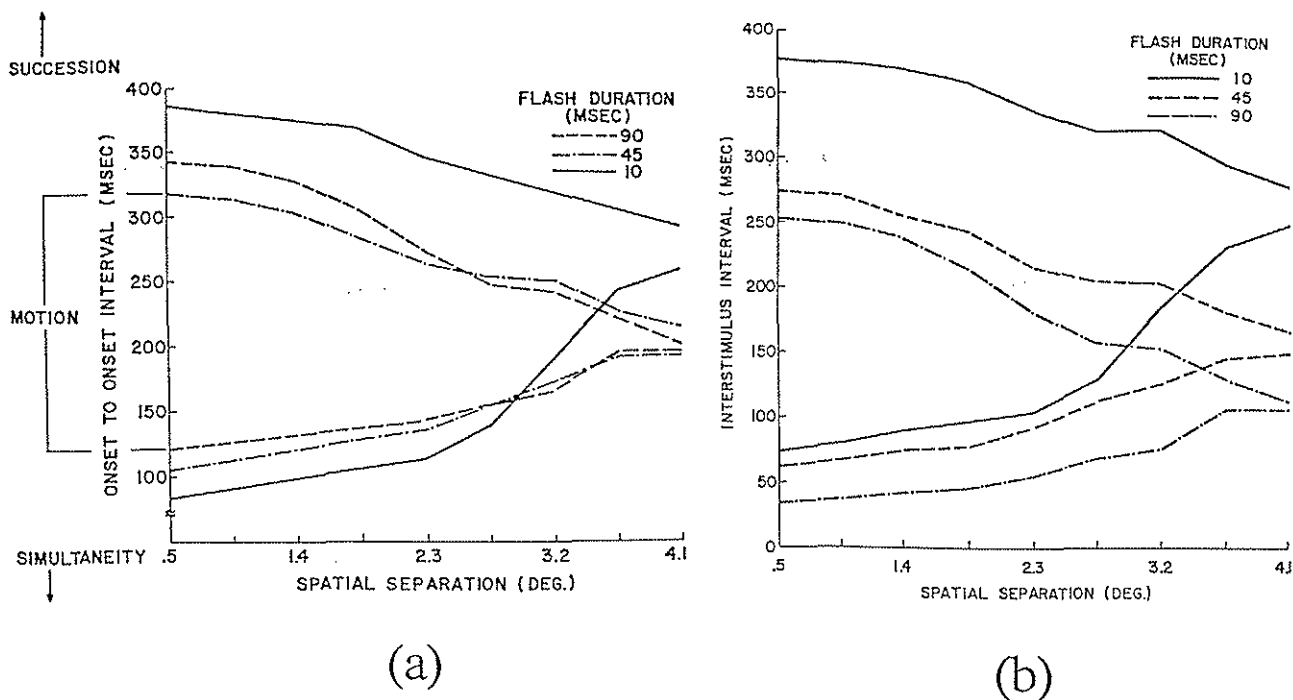
The range of SOAs over which motion is seen constricts as the spatial separation between the flashes is increased, while the SOA corresponding to the midrange remains roughly constant (Burt and Sperling, 1981; Kolers, 1972). At sufficiently large spatial separations, the range collapses to zero and the apparent motion phenomenon disappears altogether.

All of the above facts hold true if one replaces the term 'SOA' with the term 'ISI'. Furthermore, manipulating the stimulus duration also influences the probability that the subject will report motion, all other stimulus parameters being held equal. Since  $SOA = SD + ISI$ , the parameters are not mutually independent and for this reason the separate effects of the three variables have not been isolated. In Figure 6 are shown some classic apparent motion threshold data collected by Neuhaus (1930). In Figure 6a the empirical motion thresholds are expressed in terms of the ISI; and in Figure 6b they are expressed in terms of the SOA. The three sets of upper and lower thresholds in each plot correspond to different values of the parameter SD. Note that in Figure 6a, the longest SD curve has the smallest values, whereas in Figure 6b the shortest SD curve has the smallest value. The experimental difficulty involved in isolating the effects of the three temporal variables underscores the need for a theory which incorporates knowledge obtained from other data bases in order to determine the influence of each variable on the strength of the motion percept. That is, one needs to examine the results of many motion experiments in order to produce a theory which fits all of the facts. An explanation of the results of several of such experiments is provided below.

## 5. Space-Time Separability of the Motion Strength Function

In any apparent motion experiment, the probability of seeing motion along a certain path depends on various parameters of the display, including the luminance of the flashes, the duration of the frame, the distance between the elements seen in motion, and the ISI. Although motion may be seen nearly 100% of the time under optimal stimulus conditions, when the parameters of the display are close to their threshold values for producing a motion percept the probabilistic nature of the apparent motion phenomenon is clear; small changes in the values of relevant parameters then affect the proportion of identical trials on which an apparent motion percept along the path of interest is reported.

Burt and Sperling (1981) performed an important experiment in which competing paths of apparent motion were pitted against one another in such a way that the probability of reporting motion along a given path could be manipulated by varying the ISI of the display. The display used in their experiment is illustrated in Figure 7. The stimulus consisted of several successive frames of two-dimensional dot matrices. The arrows in the figure indicate various paths along which the dots could, in principle, have appeared to move. On each trial of the experiment, the subject was asked to make a forced-choice judgment of the

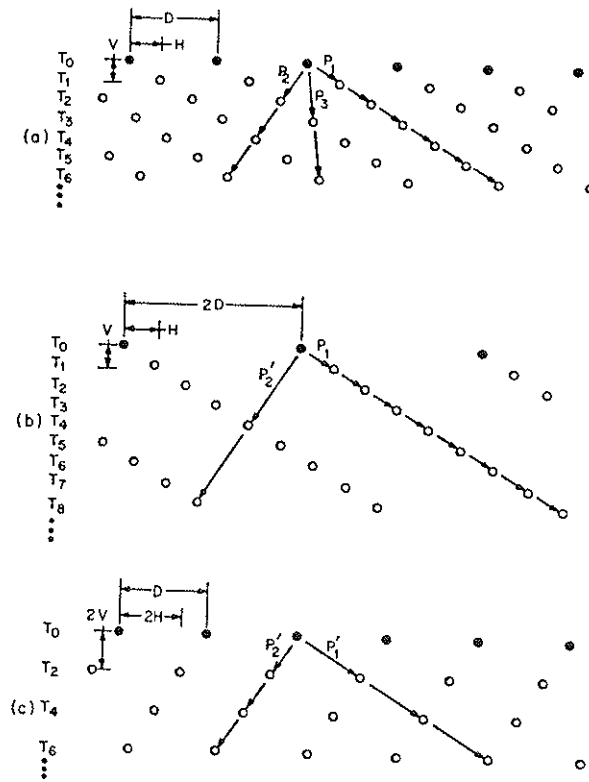


**Figure 6.** Upper and lower thresholds as a function of flash duration for two-flash apparent motion. Adapted from Neuhaus (1930). The lower threshold represents the transition between the percept of simultaneous flashes and continuous movement. The upper threshold represents the transition from perceived movement to perceived succession. (a) Threshold interstimulus intervals (ISIs). (b) Threshold onset-to-onset intervals (SOAs). [Reprinted with permission from Kellers (1972) (Figures 3.2 and 3.1), Pergamon Press.]

path along which the dots moved. The subjects did not find this task difficult; one of the many possible apparent motion paths typically dominated the others. The particular path which statistically dominated depended on the ISI of the display. All other important parameters which might potentially have affected the apparent motion percept—such as the dot luminances or frame durations—were held constant throughout the experiment.

In the vicinity of some particular ISIs, a transition occurred between the dominance of motion along one path and another. Near these transitional ISIs, apparent motion along each of the paths was reported on some percentage of the trials. The probability of reporting motion along a given path was referred to by the authors as the *strength function* for apparent motion along that path. This strength function was found to depend on the spatial separation of the elements along the path, as well as on the ISI of the display.

By moving their subjects nearer to, or farther from, the display, Burt and Sperling were able to determine whether changing the entire scale of the display affected the ISIs at which a transition between dominant apparent motion paths occurred. Significantly, it did not. The authors referred to this finding as *scale invariance* of the motion strength function. Furthermore, on the basis of an ingenious mathematical argument, they concluded that this scale invariance indicated that the underlying strength function for apparent motion to occur between any two elements was *space-time separable*, at least to first order of approximation. Space-time separability of the motion strength function means that the effect of manipulating the spatial parameter—element distance—on the probability of observing



**Figure 7.** Ambiguous motion stimulus of Burt and Sperling (1981). Panel (a) shows a multiple-path motion stimulus  $M_{1,2}$  generated by repeatedly flashing a horizontally oriented row of dots on a CRT screen. Dot spacing within the row is  $D$ . With each new presentation the row is displaced downward a distance  $V$  and to the right a distance  $H$ . Solid circles show the position of dots at time  $T_0$ ; open circles show dot positions at subsequent times  $T_i$ , where  $T_i = T_0 + it$ . Arrows show some possible paths for apparent motion of a dot presented at time  $T_0$ . Path  $P_i$  represents apparent motion to the position of the nearest dot at time  $T_i$ . Generally, all dots of the row appear to move together along the same path. Path dominance is determined by the particular values of  $t, D, V$  and  $H$ . Panel (b) shows stimulus  $M_i$ , which contains a subset of the dots of stimulus  $M_{1,2}$ : Every other dot has been removed. Path  $P_i$  is unchanged, whereas,  $P'_2$  and higher paths are greatly altered. Panel (c) shows stimulus  $M_2$ , which contains another subset of  $M_{1,2}$ : Every other row has been removed. Path  $P'_2$  is unchanged, but the distance between dots along path  $P'_1$  has been doubled relative to  $P_1$  in  $M_{1,2}$ .  $P'_1$  and  $P'_2$  in  $M_1$  and  $M_2$  have the same velocity and direction as  $P_1$  and  $P_2$  in  $M_{1,2}$ ; they differ in dot density along the path. [Figure and caption from Burt and Sperling (1981) (their Figure 1). Reprinted by permission of the American Psychological Association.]

motion between two elements is statistically independent of the effect of manipulating the temporal parameter: ISI. Thus, the total motion strength function is a product of a part which depends only on spatial separation and a part which depends only on the ISI.

It follows that, if we know how the motion strength function depends on spatial separation and also how it depends on ISI, then we know how it depends on any combination of these two factors. On the basis of their data, Burt and Sperling concluded that the motion strength function tended to a finite value when the distance between elements is small, and monotonically decreased to zero as the spatial separation of the elements was increased to large values. They suggested that a Gaussian function might well-describe the dependence of the strength function on spatial proximity in their experiment. As a function of ISI, the strength function in their experiment first increased to a maximum value at about 20 msec, then decreased monotonically to zero at asymptotically large ISIs. They showed that

a function of the form

$$S(I) = I^\beta e^{-\gamma I}, \quad (1)$$

where  $I$  denotes the ISI, provided a good mathematical fit to their data.

The importance of Burt's and Sperling's finding of space-time separability of the apparent motion strength function for the construction of models of the underlying neural process can hardly be overemphasized. It eliminates at least two classes of apparent motion models that otherwise might seem to be likely candidates. We will refer to these as *diffusion* and *traveling wave* models.

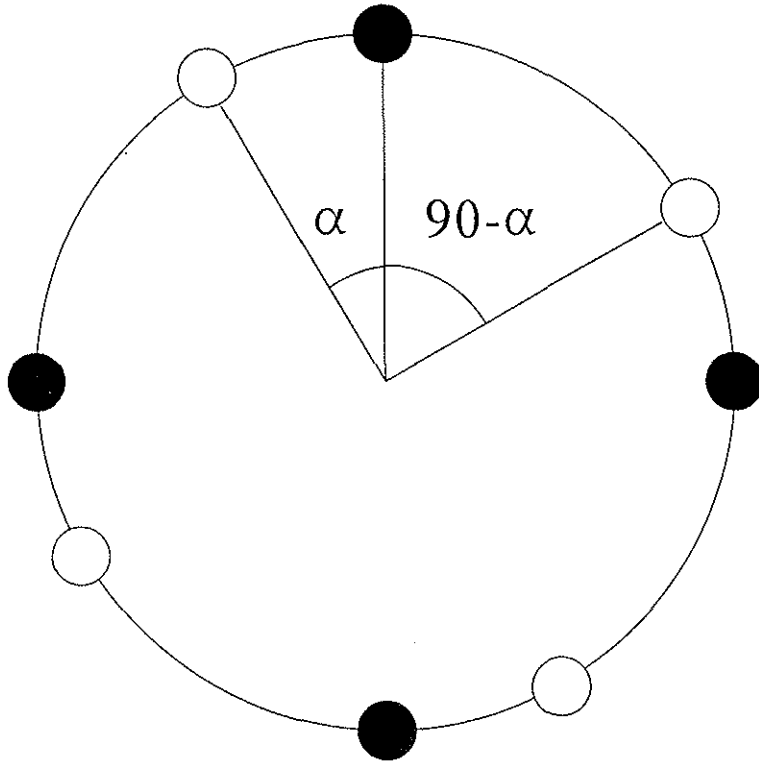
In diffusion models the occurrence of a retinal flash stimulus is assumed to give rise to a spreading neural activation, the peak of which remains centered at the location of the flash throughout time. The spatial spread of the activation, however, is assumed to undergo a progressive broadening over time, until eventually the neural effect of the flash dies out altogether as a result of a decay of the diffusion process. The spatial profile of the motion strength function at any moment in time in a diffusion model corresponds to the profile of the spreading neural activation. If, however, the total motion strength function is space-time separable, the only variation in the spatial motion strength profile over time that is allowed is a waxing and waning of the entire profile by a factor which is independent of spatial position. Clearly, the spatiotemporal profile of a diffusion process is not space-time separable.

The traveling wave model similarly fails the test of space-time separability. In this type of model, a retinal flash is assumed to give rise to a wavelike neural disturbance which propagates away from the location of the flash like ripples from a stone tossed in a pond. Traveling wave models have been discussed informally by workers in the field, but to our knowledge none has ever been proposed as a formal model of apparent motion. At first glance, this type of model might seem to be a likely candidate for explaining the fact that the minimum ISI for apparent motion is an increasing function of spatial separation (as in Figure 6). This is Korte's famous "Third Law" of apparent motion (Korte, 1915), which is discussed in Section 23. One could assume, for example, that no motion percept can occur unless a disturbance due to the first flash has propagated to the location of the second flash by the time of its occurrence. However, traveling wave models do not produce a space-time separable motion function.

In this paper, we shall consider a third idea: that the upper and lower thresholds for apparent motion represents a slice through a 2-dimensional (SOA or ISI by spatial separation) motion strength surface: the slice corresponding to lines of fixed probability for perceiving motion.

## 6. The Shape of the Spatial Component of the Motion Strength Function and its Dependence on Element Size

Independent evidence concerning the nature of the spatial component of the motion strength function was obtained in a recent study by Shechter, Hochstein, and Hillman (1988). In this experiment, subjects were presented with a two-frame apparent motion display in which motion could be observed along either of two competing apparent motion paths. The stimulus is illustrated in Figure 8. Frame 1 of the display consisted of four disks placed at 90 deg intervals along the circumference of a ring. In Frame 2, the stimulus was identical except that all four elements were displaced along the circumference of the ring by equal



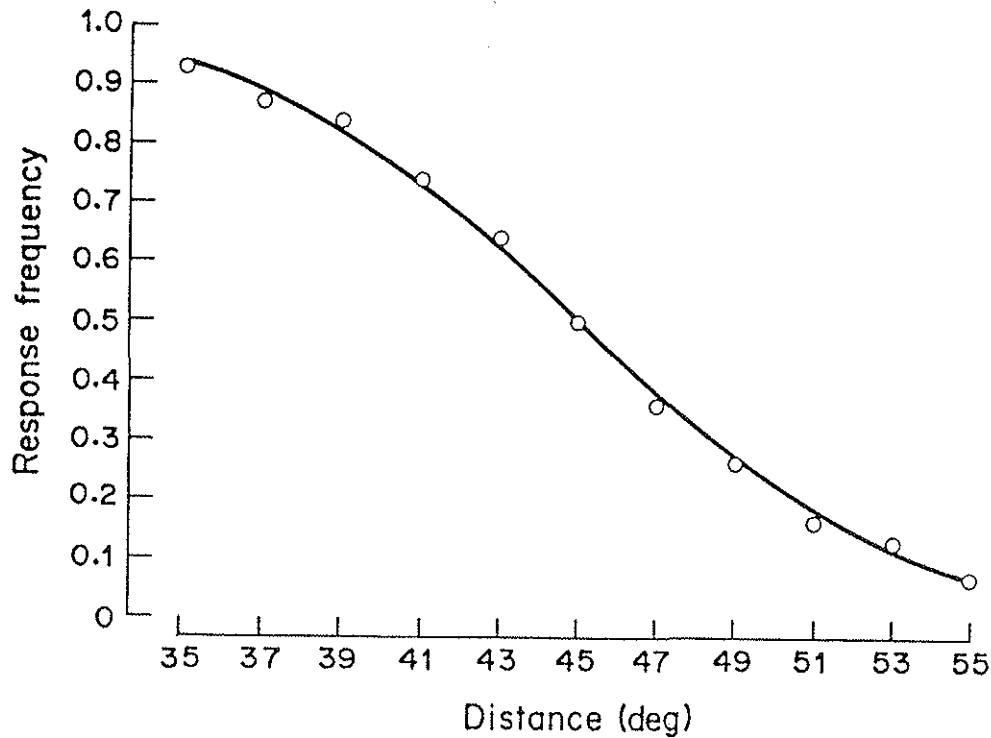
**Figure 8.** Ambiguous apparent motion display used by Shechter, Hochstein, and Hillman (1988). In Frame 1 a circular ring with four disks (represented by the dark disks in the figure) placed at 90 deg intervals around its circumference is presented. This is followed by a blank ISI, then by a second frame in which a discretely rotated version of the Frame 1 ring-and-disks stimuli is presented (represented by open disks). In both frames the actual stimulus is light on a dark background. Depending on the angle of rotation, the observer reports either a counterclockwise stimulus rotation of  $\alpha$  deg, or a clockwise rotation of  $90-\alpha$  deg, each with some probability.

distances, so that the entire stimulus in Frame 2 corresponded to a discrete angular rotation of the Frame 1 stimulus. Because of the symmetry properties of the display, a clockwise stimulus rotation of  $\alpha$  degrees is equivalent to a counterclockwise stimulus rotation of  $90-\alpha$  degrees. For a fixed clockwise rotation, subjects in the experiment always perceived either a clockwise rotation of  $\alpha$  degrees, or else a counterclockwise rotation of  $90-\alpha$  degrees, each with some probability. The task was a 2-alternative forced-choice (2AFC) between the two competing motion percepts.

In their data analysis, Shechter *et al.* converted the probability of a perceived clockwise motion into a Z-score and plotted this Z-score against the degree of clockwise stimulus rotation. They found that the relationship was a linearly decreasing one, indicating that the probability of perceived clockwise motion falls off with increased stimulus rotation according to a cumulative Gaussian function. Their data, and the Gaussian psychometric function which best fits it, are shown in Figure 9.

It remains to be seen whether this finding that the spatial motion strength function decreases with rotation angle according to the formula

$$S(\alpha) = \int_{-\infty}^{(c-\alpha)/\sigma} \frac{1}{\sqrt{2\pi}} e^{-0.5z^2} dz, \quad (2)$$



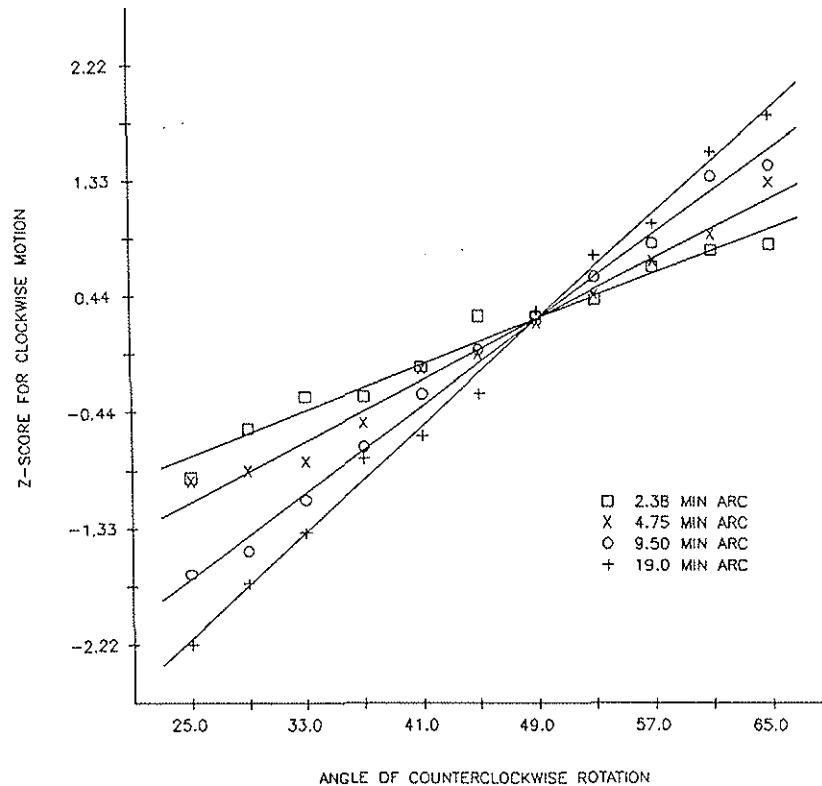
**Figure 9.** Averaged frequencies of reporting motion of the stimulus shown in Fig. 5 in either the clockwise or counterclockwise direction as a function of the angle of rotation in the reported direction. From Shechter, Hochstein, and Hillman (1988) [their Fig. 3]. Lumped results from 20 subjects and three experimental sessions. Solid line is the cumulative Gaussian psychometric function whose parameters best fit the data. [Figure reprinted with permission of Pergamon Press.]

where  $c$  and  $\sigma$  are constants, can be reconciled with the proposal of Burt and Sperling that the motion strength as a function of separation  $W$  is given by the formula

$$S(W) = \phi e^{-W^2/\theta} \quad (3)$$

where  $\phi$  and  $\theta$  are some constants. The present model is based, in part, on assuming the form of  $S(W)$  as in (3); however, the precise shape of  $S(W)$  is not as important as the fact that it is a monotonically decreasing function of separation (Burt and Sperling, 1981).

An interesting variation on the experiment of Shechter *et al.* which extends and clarifies Burt and Sperling's finding of scale invariance was recently performed by Rudd and Bressan (1991; in press). Rudd and Bressan repeated the experiment of Shechter *et al.* with an additional independent variable: disk element size. Again the subjects were asked to make a 2AFC judgment regarding the stimulus rotation direction. The proportion of clockwise judgments at each discrete angular rotation, and for each of four element sizes, was converted to a Z-score. For each of the element sizes, a plot of the Z-score *versus* the angle of counter-clockwise rotation curve was well-fit by a linear regression curve (Figure 10). Furthermore, the slopes of the least-squares linear regression models were found to increase with element size, indicating an interaction between rotation angle and element size. Rudd and Bressan further showed that the dependence of the slopes of these functions on the element diameter was approximately linear. This relationship is not due to an inappropriate choice of the interframe element separation measure, because it was also observed in a control study in



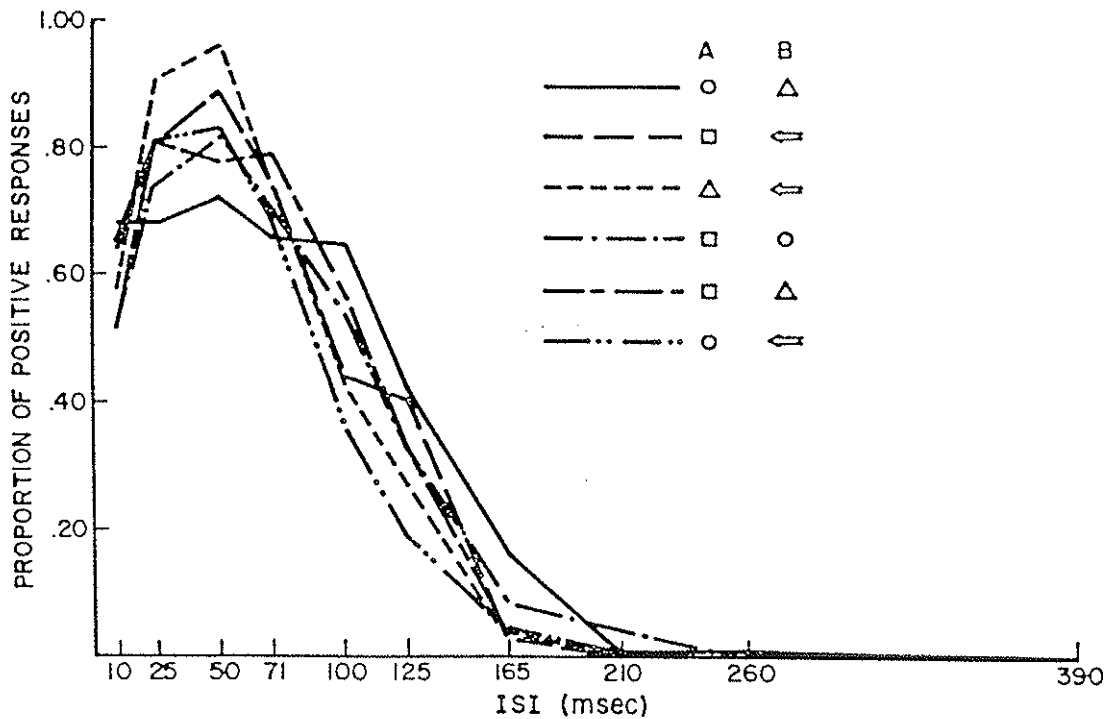
**Figure 10.** Z-score for reporting clockwise motion as a function of the counterclockwise rotation of the stimulus shown in Fig. 5 for four disk element sizes. Slopes of Z-score versus rotation angle curves increase with disk size. [From Rudd and Bressan (in press).]

which the spatial separation of the elements undergoing motion was defined in terms of the inner edge distance between the elements.

The scale invariance property of the motion strength function reported by Burt and Sperling can be deduced from the results of Rudd and Bressan. Because the rate of Z-score fall-off was found by the latter experimenters to be linearly related to the diameter of the elements, the Z-score decrement that would result from moving a subject further away from the display—and thus decreasing the visual angle of the moving elements—would be exactly compensated for by the corresponding reduction of the spatial separation between the elements. The scale invariance observed by Burt and Sperling would result. The results of Rudd and Bressan demonstrate that scale invariance results from the dependence of the motion strength function on both spatial separation and element scale as measured in degrees of visual angle. The effect does not appear to depend on a simultaneous scaling of the ring stimulus and disk elements, nor on the perceived three-dimensional size constancy of the experimental stimulus as the subject is moved with respect to the display.

## 7. The Temporal Motion Strength Function and its Dependence on Flash Duration

Burt and Sperling observed that the temporal motion strength function measured in their experiment was a function with a single maximum at an ISI of about 20 msec and which tended to zero at large values of the ISI. They proposed equation (1) to model this function. This same characteristic shape of the temporal motion strength function has been discovered in a wide variety of experiments. For example, consider the results of an experiment by Kolers



**Figure 11.** The proportion of trials on which motion was seen between dissimilar shapes in a two-flash display. Separate curves correspond to six stimulus pairs. [Data from Kolers and Pomerantz (1971). Figure reprinted from Kolers (1972) with permission of Pergamon Press.]

and Pomerantz (1971) which are plotted in Figure 11. In this experiment, subjects were presented with a continuously cycling display, each cycle of which consisted of two frames separated by an ISI of variable length. The visual objects presented in the two frames were not simple patches of light; instead they consisted of the pictorial figures shown in the diagram. These two figures were presented to different spatial locations in the two frames. Subjects were asked to report whether or not they observed motion between the successively presented figures on each trial. The functions graphed in Figure 11 represent the proportion of trials on which motion was observed for each of the six figure pairs tested as a function of the ISI. Thus, they describe the temporal motion strength functions corresponding to each of the figure pairs. Note that, in each case, the general shape of the motion strength function is similar to the shape deduced by Burt and Sperling from their data. This is quite remarkable given the many superficial dissimilarities between the two apparent motion paradigms: dots *versus* figures; forced choice of (motion/no motion) *versus* direction of motion; two-frames *versus* multiple frames, etc.

Although the shape of the temporal motion strength function is insensitive to these variations in experimental design, it is clear from the results of several experiments that it depends critically on at least one variable: the frame duration. In Figure 12 are plotted the results of an experiment performed by Kolers (1964) which support this claim. This experiment was a simple two-flash apparent motion experiment in which the stimuli were small luminous patches. Each of the curves in the figure corresponds to the motion strength function associated with a particular value of the frame duration variable. Note that, for brief flash durations, the shapes of the functions are similar to those found by Burt and

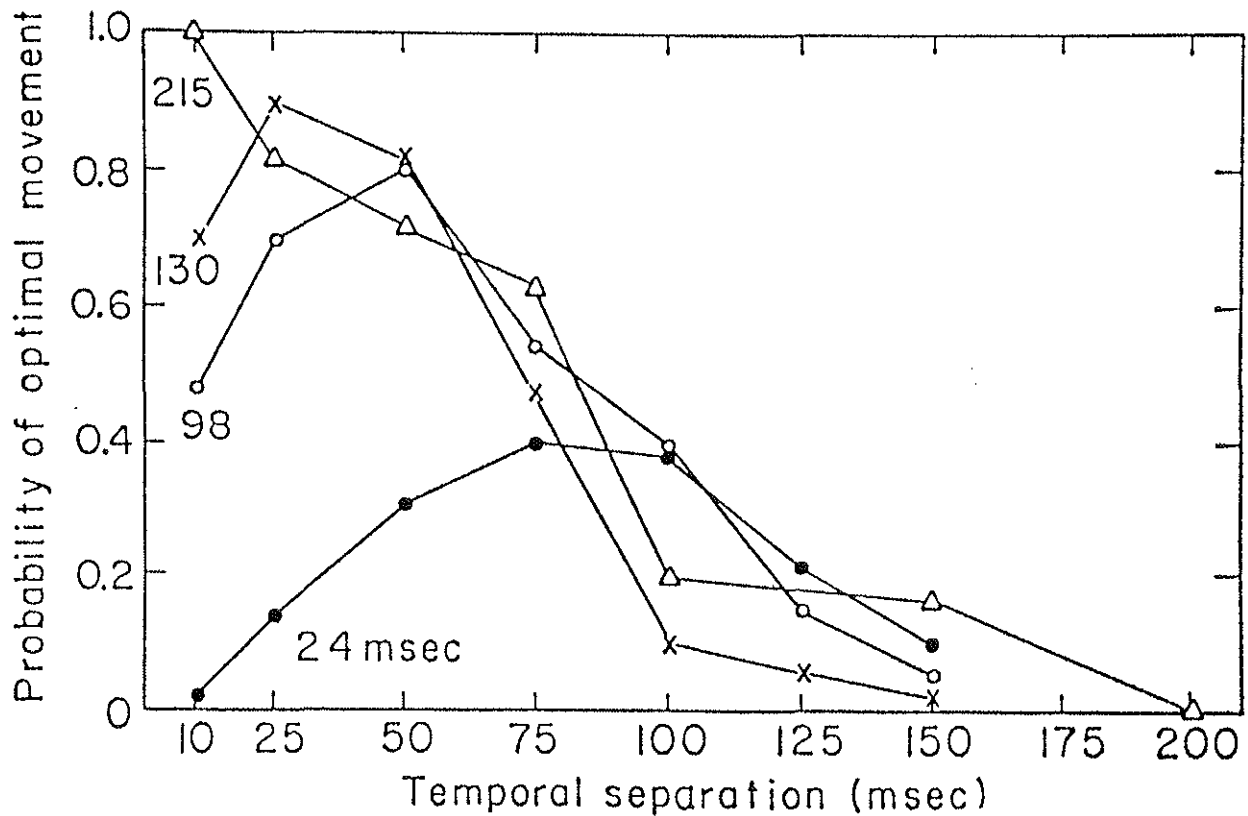
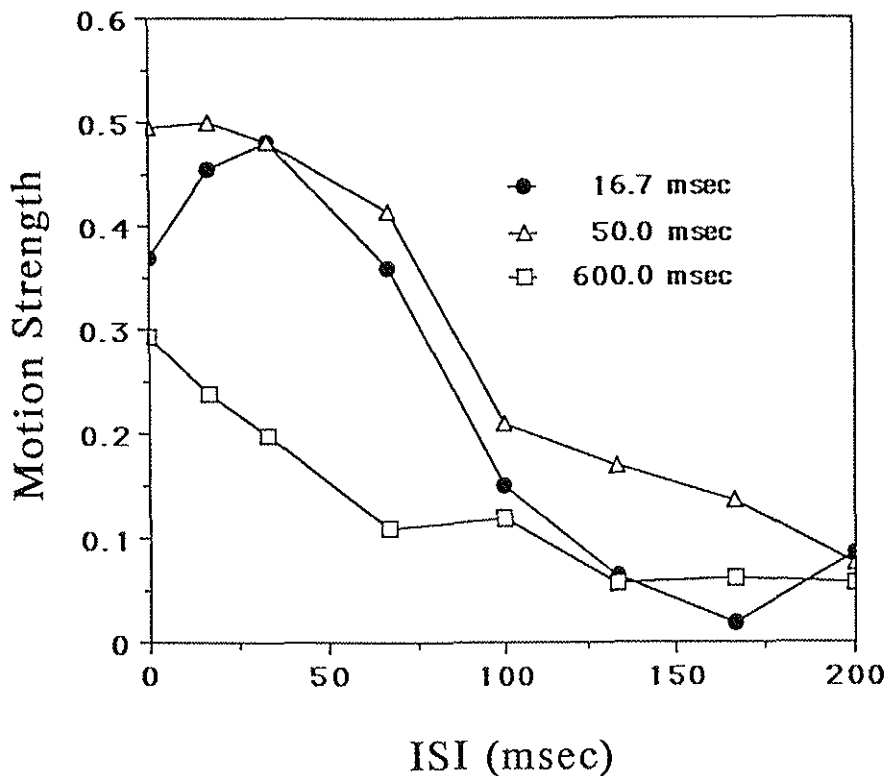


Figure 12. The likelihood of seeing motion between two flashes whose durations and interstimulus intervals were varied. Duration is shown for the separate curves in msec. [Figure adapted from Kolers (1964), reprinted with permission of Scientific American, Inc. Caption from Kolers (1972), used with permission of Pergamon Press.]

Sperling, and by Kolers and Pomerantz. However, as the frame duration is increased, the curves gradually assume the shape of a monotonically decaying function of ISI.

A recent experiment by Bressan and Rudd (in preparation) indicates that this result generalizes to a forced-choice direction-of-motion judgment task using stimuli of the type employed by Shechter *et al.* In their experiment, Bressan and Rudd fixed the angle of rotation of the ring-and-disks stimulus at either a 33 deg counterclockwise rotation (equivalent to a 67 deg clockwise rotation), or a 67 deg counterclockwise (33 deg clockwise) rotation. The two conditions were counterbalanced in order to control for any bias on the part of the subject to report motion in a particular direction. An ISI of variable length was inserted between the two frames, during which the ring appeared without the disk elements. The experiment was carried out with three different durations of Frame 1. The duration of Frame 2 was fixed at a value of 600 msec throughout the experiment.

In Figure 13 are plotted the motion strength functions for a single observer corresponding to Frame 1 durations of 16.7, 50, and 600 msec. On the *y*-axis is plotted the proportion of trials on which the observer reported motion along a path between the most proximal elements. On the *x*-axis is the ISI. When the duration of Frame 1 was brief (16.7 msec), the shape of the motion strength function first rose, then fell, with increasing ISI; a single peak occurred at an ISI of about 33 msec. Thus, the motion strength functions for brief Frame 1 durations were similar in shape to those found by Burt and Sperling, and by Kolers and Pomerantz. As the duration of Frame 1 was increased, the shape of the motion strength function assumed a monotonically decaying profile, with the maximum probability



**Figure 13.** Proportion of trials on which motion was seen in the direction of the nearest disks for the stimulus illustrated in Fig. 5 as a function of ISI. Separate curves correspond to Frame 1 durations of 16.7, 50, and 600 msec. Subject PB. Disk size: 9.50 min arc; disk luminance: 18.36  $cd/m^2$ ; background luminance: 0.03  $cd/m^2$ . [From Bressan and Rudd (in preparation).]

of reporting motion in the direction of the nearest element occurring at an ISI of zero. The authors believe that the relatively low level of motion strength in the 600 msec Frame 1 trials is explained by the fact that these data were gathered in much larger blocks of trials, leading to greater motion adaptation in this experimental condition. They are presently performing a replication of this study in which the conditions are counterbalanced to control for this adaptation.

The overall pattern of results in the Bressan and Rudd experiment is similar to that observed in the data of Kolers (1964), again despite many differences in the nature of the displays and tasks. The fact that a similar data pattern emerges from the two experimental paradigms suggests that it is the duration of Frame 1 rather than that of Frame 2 that controls the shape of the temporal motion strength function. To our knowledge, the effects of manipulating the duration of Frame 2 alone have not been experimentally determined.

## 8. The Insensitivity of the Motion Correspondence Process to Figural Identity

Early researchers (Wertheimer, 1912; Higginson, 1926; Orlansky, 1940; Kolers, 1972) noted the ease with which the apparent motion correspondence process bridged the gap between two figures with different identities. For example, a display consisting of small spot of light presented in Frame 1, followed by the presentation in a different spatial location of a picture of a human face in Frame 2, can generate a perception of the spot moving to the location of the face and being transformed into it in the process. Thus, it is clear that motion perception is not based on a form-dependent matching process.

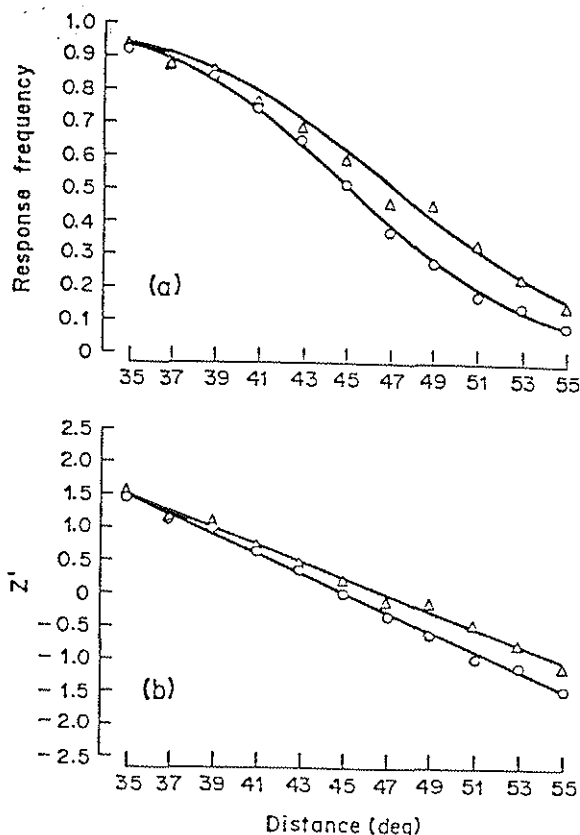
Depending on the shapes displayed in each frame and the timing of the display, the interframe shape disparity may be resolved by the motion process in a variety of ways. With brief frame durations and ISIs, the disparity may be resolved by objectless phi motion. When the frame durations and ISIs are somewhat longer, the disparity may be resolved by a continuous deformation of object shape (van der Waals and Roelofs, 1930; 1931), or even by a transformation in depth (Neuhaus, 1930).

More recent experiments (Kolers and Pomerantz, 1971; Navon, 1976; Burt and Sperling, 1981) have confirmed the relative unimportance of figural identity in determining motion correspondence. The effects of figural identity on both the spatial and temporal components of the motion strength function have also been independently studied. The data in Figure 16, from an experiment by Kolers and Pomerantz, indicate that the shape of the temporal motion strength function is little-affected by the figural identity of the objects undergoing apparent motion. In this experiment, pairs of frames—each containing a simple geometrical shape such as an arrow, circle, triangle, or square—were presented in succession. The effects of varying the ISI and the degree of shape similarity on the probability of perceived motion were simultaneously investigated. The experimenters found that degree of shape similarity accounted for only about 1–3% of the total statistical effect.

Shechter, Hochstein, and Hillman (1988) also investigated the effect of varying figural identity on the spatial motion strength function in the context of a modified ring-and-disks paradigm. To do this, they devised a display consisting of two filled triangles placed at 180° intervals along the ring circumference and two disks placed at the midpoints along the arcs between the triangles. In the second frame, the entire stimulus was rotated 45 deg in either the clockwise or counterclockwise direction. Although the experimenters found the effect of the figural identity manipulation to be statistically significant for determining the direction of the reported motion, the magnitude of the effect was not very large; subjects judged the motion to be in the direction of the figure with the same identity about 58% of the time, or 8% more often than would be expected by chance.

The fact that the bias introduced by figural identity was larger in this study than that found by Kolers and Pomerantz may be partly due to the fact that direction of perceived motion in the ring-and-disk paradigm is potentially determined by the summed effects of four elements per frame, rather than the single flashes used by Kolers and Pomerantz. In addition, the 8% bias found by Shechter *et al.* (1988) was observed in an experimental condition in which no other cues for motion correspondence were operating. In other conditions of their experiment, they measured the combined effects of figural identity and spatial proximity as correspondence cues. The results, shown in Figure 14, indicated that the magnitude of the effect of the figural identity cue was reduced when a strong spatial proximity cue was simultaneously present.

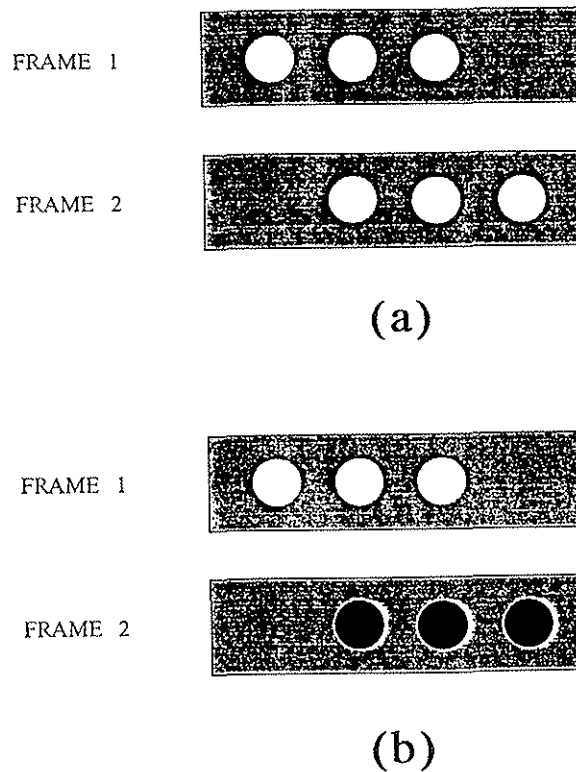
After reviewing the literature concerning the relationship between figural identity and apparent motion, Kolers (1972) concluded that the data would be best explained on the basis of the assumption that there are two parallel subsystems in the human visual system for the computation of motion and the maintenance of figural identity. This view is supported by physiological findings made since the time of Kolers' review, which indicate the existence of parallel visual pathways for motion perception and static form perception. Both the psychophysical and physiological results are clarified by our theoretical results concerning the design of parallel Static BCS and Motion BCS architectures.



**Figure 14.** Combined effects of figural identity and spatial proximity as motion correspondence cues for the stimulus shown in Fig. 5. (a) Upper curve: Frequency of reporting motion in direction that maintains figural identity as a function of the angle of rotation. For one direction of rotation the elements in the stimulus (represented by the disks in Fig. 5) maintained their shape across frames, while for the other direction they changed from disks to filled triangles and vice versa. Lower curve: Frequency of reporting motion in either the clockwise or counterclockwise direction as a function of the angle of reported rotation for filled disks which retained their shapes across frames. (b) Data of (a) converted to Z-scores. [From Shechter, Hochstein, and Hillman (1988), (their Fig. 7). Reprinted with permission of Pergamon Press.]

## 9. Group and Element Apparent Motion: Ternus Displays

Not just the existence of a motion percept, but also its figural organization, can depend on subtle aspects of the display such as the ISI. This fact is nicely illustrated by an ingenious apparent motion display (Figure 15a) originally devised by Josef Ternus (1926/1950). In Frame 1 of the Ternus display, three black elements are placed in a horizontal row on a white background (or the contrast may be reversed). After an ISI, in Frame 2 all three elements are shifted to the right by the distance of the interelement spacing; so that the positions of the two rightwardmost elements in Frame 1 overlap those of the two leftwardmost elements in Frame 2. The entire sequence, Frame 1-ISI-Frame 2, may then be repeated for several cycles. Depending on the ISI of the display, the observer in the Ternus experiment will see either of two bistable motion percepts (Ternus, 1926/1950; Pantle and Picciano, 1976; Petersik and Pantle, 1979; Pantle and Petersik, 1980; Breitmeyer and Ritter, 1986a, 1986b). When the ISI is long, but not so long that the apparent motion gives way to the perception of succession, there is a tendency to see the line of elements move back and forth as a group. This percept is called *group motion*. When the ISI is short, but not so short that the apparent motion gives way to the perception of flickering in place, there is a tendency to see the leftwardmost element in Frame 1 'jumping' to the location of the rightwardmost element in Frame 2, while



**Figure 15.** The Ternus display. (a) Three spots are presented in each frame in such a way that the two leftwardmost spots in Frame 2 occupy the same positions as the two rightwardmost spots in Frame 1. The two frames are repeatedly cycled with ISIs inserted between them. At very short ISIs, all dots appear to flicker in place. At longer ISIs the dots at shared positions appear to remain stationary, while apparent motion occurs between the leftwardmost spot in Frame 1 and the rightwardmost spot in Frame 2 (“element motion”). At still longer ISIs, the three dots appear to move from Frame 1 to Frame 2 and back as a group (“group motion”). (b) When the dots in successive frames have opposite contrast with respect to the frame, only group motion occurs at the ISIs where element motion occurred in (a).

the two central elements remain in place. This percept is called *element motion*.

A number of stimulus variables besides the ISI have been shown to influence the type of motion that is observed in the Ternus display. For example, Petersik and Pantle (1979) found that the percentage of group responses increased with increasing frame duration and interframe interval luminance, as well as with increasing ISI; while it generally decreased with stimulus contrast (but there was some crossover). They also found that the ISI at which the transition from element to group motion occurred was an increasing function of dark adaptation. The transitional ISI has also been shown to decrease with increasing element size, element contrast, frame duration, and viewing eccentricity (Breitmeyer and Ritter, 1986a, 1986b).

These effects are consistent with the explanation that is offered below for group motion and element motion. A particularly demanding Ternus percept was discovered by Pantle and Picciano (1976), which is also explained below. These authors reversed the relative contrast of the three dots on background in the two successive frames (Figure 15b). Then group motion was perceived even at the short ISIs which generated element motion when relative contrast was not reversed between the two frames.

## 10. Motion versus Visual Persistence

It was first suggested by Braddick and Adlard (1978; Braddick, 1980) that the perception of group motion in the Ternus paradigm might be inhibited by the visual persistence of the central elements of the display. This is consistent with the fact that even a small interframe perturbation of the spatial positions of those elements may induce the group motion percept, even when the ISI of the stimulus is zero.

In the case wherein the element positions are not perturbed and the ISI equals zero, one would not expect a motion percept to be generated by the element positions common to both frames. When the ISI is small but nonzero, persistence of the activity of the neural mechanisms responsible for the detection of these central elements would be indistinguishable from the actual persistence of the elements as far as any motion detector ‘looking’ at the outputs of these neural mechanisms is concerned. Therefore, such a persistence could account for the fact that the critical ISI at which the transition from element to group motion occurs is some positive value rather than zero. The situation is the analog of the mechanism by which neural persistence accounts for critical flicker fusion frequency in static form perception, and may in fact be related to that phenomenon.

The idea that visual persistence inhibits group motion has recently been advocated by Breitmeyer and Ritter (1986a, 1986b) who have demonstrated experimentally that a number of factors which are known to be positively correlated with measures of iconic persistence are also positively associated with the value of the critical ISI at which the transition from element to group motion occurs in the Ternus paradigm. Specifically, they have shown that the transitional ISI is a decreasing function of viewing eccentricity, element size, and frame duration. These three variables are all negatively correlated with measures of the visual iconic persistence (Bowen, Pola, and Matin, 1974; Bowling and Lovegrove, 1980; Breitmeyer and Halpern, 1978; Breitmeyer, Levi, and Harwerth, 1981; Di Lollo, 1977; Di Lollo and Hogben, 1985; Meyer and Maguire, 1977; Mezzrich, 1984). None of these studies indicate, however how visual persistence should be modeled, how it may inhibit group motion, or why the illusory percept in element motion does not “collide” with the stationary dots and be thereby terminated.

With these basic facts and ideas about apparent motion in mind, we now proceed to describe the MOC Filter model of motion computation, and to demonstrate that it is capable of accounting for these and other more subtle effects.

## 11. Design of a MOC Filter

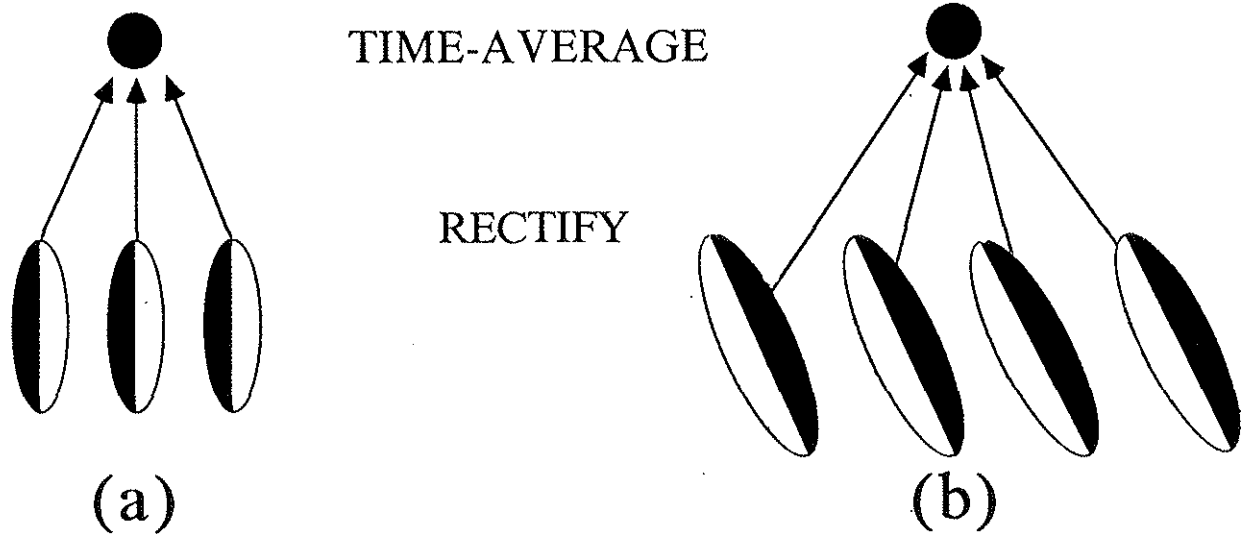
The equations for a one-dimensional MOC Filter were described in Grossberg and Rudd (1989c) and for a two-dimensional MOC Filter in Grossberg and Mingolla (1990a, 1990b). Its five processing levels are described qualitatively below for the two-dimensional case. The simplified equations used for our one-dimensional computer simulations are also provided.

### Level 1: Preprocess Input Pattern

The image is preprocessed before activating the filter. For example, it is passed through a shunting on-center off-surround net to compensate for variable illumination, or to “discount the illuminant” (Grossberg and Todorović, 1988).

In the 1-D theory,  $I_i$  denotes the input at position  $i$ .

### Level 2: Sustained Cell Short-Range Filter



**Figure 16.** The sustained cell short-range filter. Inputs are spatiotemporally filtered by sustained cells with individual oriented receptive fields, and temporal filtering characteristics which are determined by the dynamics of a shunting membrane equation. The output of each sustained cell is rectified and thresholded. The outputs of a spatially aligned array of cells with like orientation, direction-of-contrast, and direction-of-motion are pooled. The breadth of the spatial pooling scales with the size of the simple cell receptive fields, as in (a) and (b).

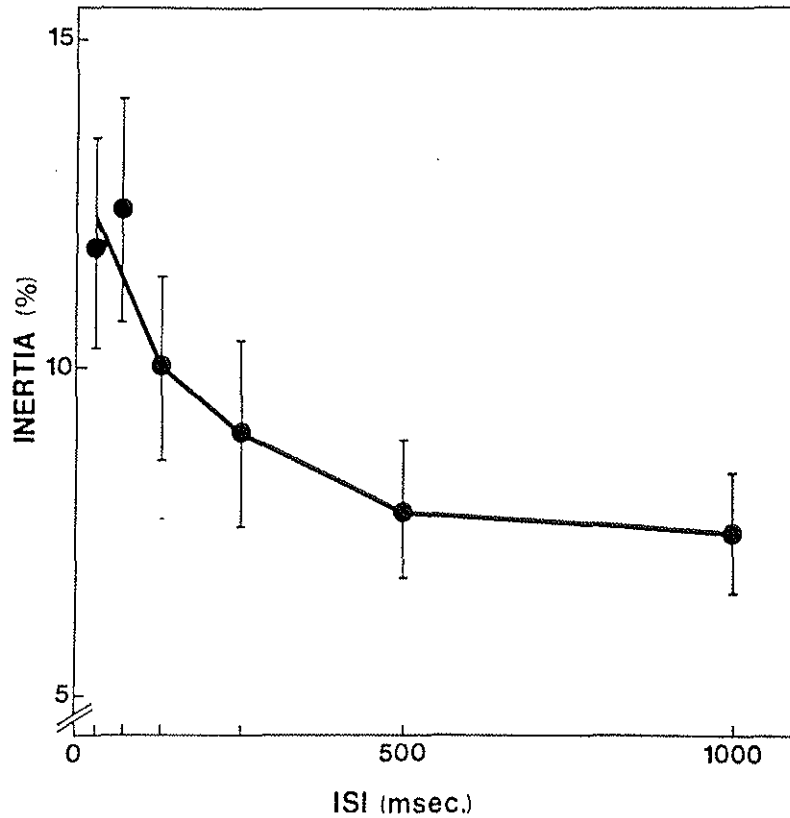
Four operations occur here, as illustrated in Figure 16.

(1) **Space-Average:** Inputs are processed by individual oriented receptive fields, or simple cells, as in Figure 3, which add excitatory and inhibitory contributions from two halves of the receptive field.

(2) **Rectify:** The output signal from a simple cell grows with its activity above a signal threshold. Thus the output is half-wave rectified.

(3) **Short-Range Spatial Filter:** A spatially aligned array of simple cells with like orientation and direction-of-contrast pool their output signals to activate the next cell level. As shown in Figure 16b, the direction of spatial pooling is not necessarily perpendicular to the oriented axis of the simple cell receptive field. The target cells are pooled in a movement *direction* that is not necessarily perpendicular to the simple cell's preferred *orientation*. This spatial pooling plays the role of the short-range motion limit  $D_{\max}$  (Braddick, 1974). The breadth of spatial pooling scales with the size of the simple cell receptive fields (Figures 16a and 16b). Thus " $D_{\max}$ " depends on the spatial frequency content of the image (Anderson and Burr, 1987; Burr, Ross, and Morrone, 1986; Nakayama and Silverman, 1984, 1985; Petersik, Pufahl, and Krasnoff, 1983), and is not a universal constant.

(4) **Time-Average:** The target cell time-averages the inputs that it receives from its short-range spatial filter. This operation has properties akin to the "visual inertia" during



**Figure 17.** ‘Visual inertia’ in apparent motion measured by Anstis and Ramachandran (1987). Ambiguous apparent motion was biased by priming dots, and the degree of bias (inertia) was measured as a function of the interval (ISI) between the priming dot and test. The bias induced by the priming dots was about 12% at short ISIs, and fell monotonically to about 7% for ISIs exceeding 500 msec. (Anstis’ and Ramachandran’s Fig. 6 used by permission of Pergamon Press.)

apparent motion that was reported by Anstis and Ramachandran (1987); see Figure 17.

In the present paper we will be concerned only with simulations involving one spatial dimension; thus, only horizontal motions are considered. It therefore suffices to consider two types of such cells that filter the input pattern  $I_i$ , one of which responds to a light-dark luminance contrast (designated by  $L$ , for left) and the other of which responds to a dark-light luminance contrast (designated by  $R$ , for right). Output pathways from like cells converge (as in Figure 13) to generate inputs  $J_{iL}$  and  $J_{iR}$  at each position  $i$ . The activity  $x_{ik}$  of the  $i$ th target cell at Level 2 obeys an activation equation

$$\frac{d}{dt}x_{ik} = -Ax_{ik} + (1 - Bx_{ik})J_{ik}, \quad (4)$$

where  $k = L, R$ , which performs a time-average of the input  $J_{ik}$ . In (4), constant  $-A$  is the passive decay rate and  $B^{-1}$  is the maximum activity of  $x_{ik}$ .

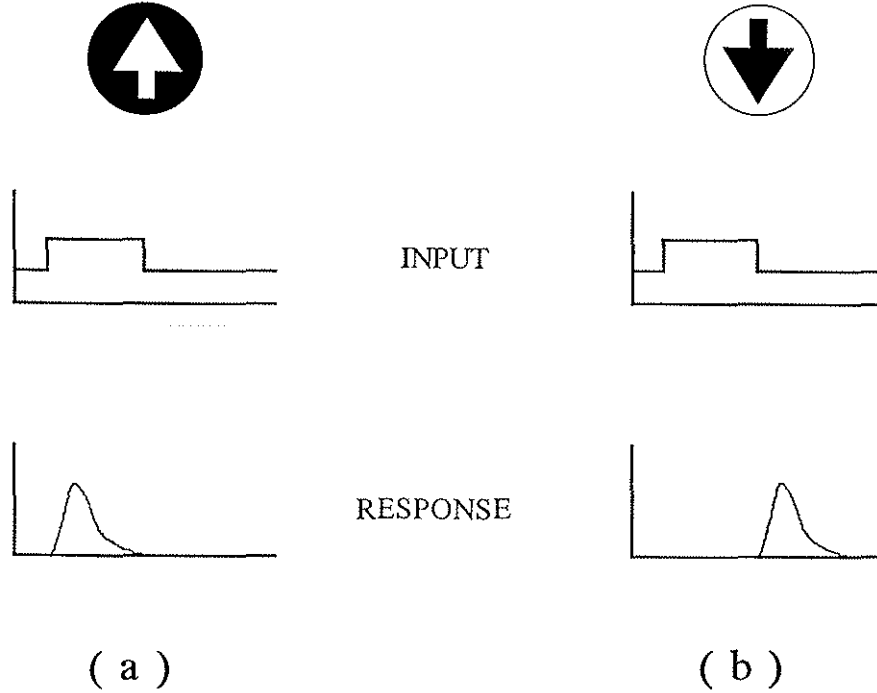
### Level 3: Transient Cell Filter:

In parallel with the sustained cell filter, a transient cell filter reacts to input increments (on-cells) or decrements (off-cells) with positive outputs (Figure 18). These filters use four operations too:

(1) **Space-Average:** This is accomplished by a receptive field that sums inputs over its entire range, unlike the receptive field of a sustained cell. This receptive field is assumed to be unoriented, or circularly symmetric, for simplicity.

## Transient On Cell

## Transient Off Cell



**Figure 18.** Responses over time of transient on- and off-cells. (a) On-cell responses are formed from the positive-rectified and thresholded time derivative of a spatiotemporally filtered image. The spatial filter has an unoriented on-center off-surround receptive field. The temporal filter is based on the dynamics of a shunting membrane equation which time-averages the spatially filtered input. The on-cell thus produces a time-averaged response to an increment in the input. (b) Off-cells are formed from the negative-rectified and thresholded time derivative of a spatiotemporal filter. The off-cell thus produces a time-averaged response to a decrement in the input.

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

(3) **Transient Detector:** The on-cells are activated when the time-average increases (Figure 18a). The off-cells are activated when the time-average decreases (Figure 18b).

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

Here we model the activities of the transient cells in a simple way as the rectified time derivatives of an unoriented space-time average  $x_i$  of the input pattern  $I_i$ . The time derivative is given by the activation equation

$$\frac{d}{dt}x_i = -Cx_i + (D - Ex_i)\sum_j I_j F_{ji}, \quad (5)$$

where  $F_{ji}$  is the unoriented spatial kernel that represents a transient cell receptive field.

Positive and negative half-wave rectifications of the time derivative are performed independently by defining

$$y_i^+ = \max\left(\frac{d}{dt}x_i - \Gamma, 0\right), \quad (6)$$

and

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

where  $\Gamma$  and  $\Omega$  are constant thresholds. The activity  $y_i^+$  models the response of a transient on-cell; and the activity  $y_i^-$  models the response of a transient off-cell.

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

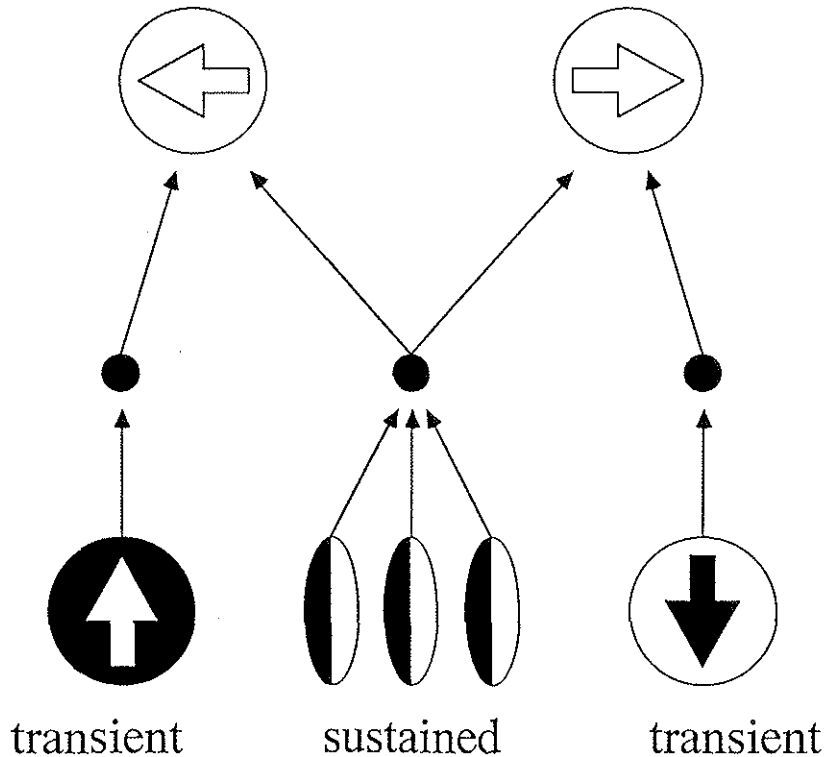
Maximal activation of a Level 2 sustained cell filter is caused by image contrasts moving in either of two directions that differ by  $180^\circ$ . Multiplicative gating of each Level 2 sustained cell output with a Level 3 transient cell on-cell or off-cell removes this ambiguity (Figure 19). For example, consider a sustained cell output from vertically oriented dark-light simple cell receptive fields that are joined together in the horizontal direction by the short-range spatial filter (Figure 16a). Such a sustained cell output is maximized by a dark-light image contrast moving to the right or to the left. Multiplying this Level 2 output with a Level 3 transient on-cell output generates a Level 4 cell that responds maximally to motion to the left. Multiplying it with a Level 3 off-cell output generates a Level 4 cell that responds maximally to motion to the right.

Multiplying a sustained cell with a transient cell is the main operation of the Marr and Ullman (1981) motion detector. Despite this point of similarity, Grossberg and Rudd (1989c) described six basic differences between the MOC Filter and the Marr-Ullman model. For example, none of the operations such as short-range spatial filtering, time-averaging, and rectification occurs in the Marr-Ullman model. In addition, the rationale of the MOC Filter—to design a filter that is sensitive to direction-of-motion and insensitive to direction-of-contrast—is not part of the Marr-Ullman model. This step requires long-range spatial filtering and competitive sharpening, described below, that are also not part of the Marr-Ullman model. This difference is fundamental. The Marr-Ullman model is a product of the “independent modules” perspective. The MOC Filter’s insensitivity to direction-of-contrast can only be formulated within the framework of BCS/FCS complementarity (Figure 1): One cannot understand how a boundary filter’s output can be insensitive to direction-of-contrast unless there is a complementary “seeing” system that is sensitive to direction-of-contrast.

In the 1-D MOC Filter there are two types of sustained cells (corresponding to the two antisymmetric directions-of-contrast), and also two type of transient cells (the on-cells and the off-cells). Consequently, there are four types of gated responses that can be computed. Two of these produce cells that are sensitive to local rightward motion: the  $(L, +)$  cells that respond to  $x_{iL}y_i^+$ , and the  $(R, -)$  cells that respond to  $x_{iR}y_i^-$ . The other two produce cells which 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 a sensitivity to the direction-of-contrast of their inputs from the Level 2 sustained cells from which they are constructed.

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

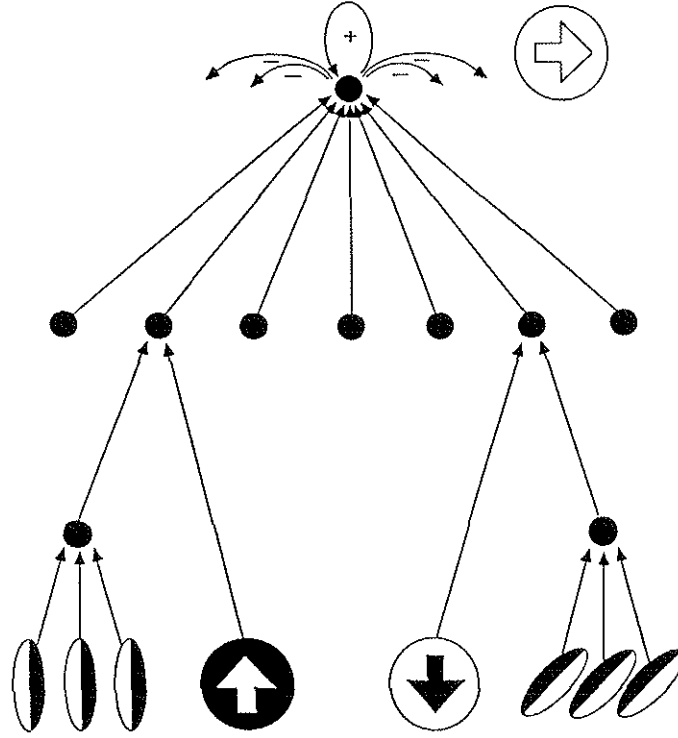
#### Level 5: Long-Range Spatial Filter and Competition



**Figure 19.** Transient cell gating of sustained cell activities to produce directionally-sensitive responses. The short-range filter, which is constructed from like-oriented simple cells, responds ambiguously to a contrast pattern (dark-light in the illustration) moving either to the right or to the left. This ambiguity of motion direction is eliminated by gating the short-range filter response with either a transient on-cell response (to produce a left motion signal) or a transient off-cell response (right motion signal).

Outputs from Level 4 cells sensitive to the same direction-of-motion but opposite directions-of-contrast activate individual Level 5 cells via a long-range spatial filter that has a Gaussian profile across space (Figure 20). This long-range filter groups together Level 4 cell outputs that are derived from Level 3 short-range filters with the same directional preference but different simple cell orientations. Thus the long-range filter provides the extra degree of freedom that enables Level 5 cells to function as “direction” cells, rather than “orientation” cells. It has been shown that cells in MT can also respond to a range of orientations that are not perpendicular to their preferred direction-of-motion (Albright, 1984; Albright, Desimone and Gross, 1984; Maunsell and van Essen, 1983; Newsome, Gizzi, and Movshon, 1983).

The long-range spatial filter broadcasts each Level 4 signal over a wide spatial range in Level 5. Competitive, or lateral inhibitory, interactions within Level 5 contrast-enhance this input pattern to generate spatially sharp Level 5 responses. A winner-take-all competitive network (Grossberg, 1973, 1982) can transform even a very broad input pattern into a focal activation at the position that receives the maximal input. The winner-take-all assumption is a limiting case of how competition can restore positional localization. More generally, we suggest that this competitive process partially contrast-enhances its input pattern to generate a motion signal whose breadth across space increases with the breadth of its inducing pattern. A contrast-enhancing competitive interaction has also been modeled at the complex cell level of the SOC Filter (Grossberg, 1987b; Grossberg and Marshall, 1989). The Level 5 cells of the MOC Filter are, in other respects too, computationally homologous to the SOC Filter



**Figure 20.** Combination of like direction-of-motion activities across space via a long-range Gaussian filter. Local direction-sensitive responses of opposite direction-of-contrast, over a range of orientations, are gated by transient cells of opposite types to produce like direction-of-motion signals. These local signals are combined by a long-range Gaussian spatial kernel to produce a spatially broad pattern of activity across the Level 5 network. This broad pattern is then contrast-enhanced by a competitive, or lateral inhibitory, interaction. The contrast-enhancement restores positional information.

complex cells.

In the 1-D theory, we define the transformation from Level 4 to Level 5 by letting

$$r_i = x_{iLY_i^+} + x_{iRY_i^-}, \quad (8)$$

and

$$l_i = x_{iLY_i^-} + x_{iRY_i^+}, \quad (9)$$

be the total response of the local right motion and left motion detectors, respectively, at position  $i$  of Level 4. Signal  $r_i$  increases if either a light-dark or a dark-light contrast pattern moves to the right. Signal  $l_i$  increases if either a light-dark or a dark-light contrast pattern moves to the left.

These local motion signals are assumed to be filtered independently by a long-range operator with a Gaussian kernel

$$G_{ji} = H \exp [-(j-i)^2/2K^2], \quad (10)$$

which defines the input fields of the Level 5 cells. Thus, there exist two types of direction sensitive cells at each position  $i$  of Level 5. The activity at  $i$  of the right-motion sensitive cell is given by

$$R_i = \sum_j r_j G_{ji}, \quad (11)$$

and the corresponding activity of the left-motion sensitive cell is given by

$$L_i = \sum_j l_j G_{ji}. \quad (12)$$

The Gaussian kernel generates a spatially distributed input to Level 5 in response to even a focal input to Level 1. The next operation spatially sharpens the response at Level 5 to these distributed inputs. This contrast-enhancing competitive, or lateral inhibitory, interaction within Level 5 generates the activities which encode a local measure of motion information. In the simplest case, the competition is tuned to select that population whose input is maximal, as in

$$x_i^{(R)} = \begin{cases} 1 & \text{if } R_i > R_j, j \neq i \\ 0 & \text{otherwise,} \end{cases} \quad (13)$$

and

$$x_i^{(L)} = \begin{cases} 1 & \text{if } L_i > L_j, j \neq i \\ 0 & \text{otherwise.} \end{cases} \quad (14)$$

In the simulations reported in this paper we have made the above assumption for simplicity. The functions  $x_i^{(R)}$  and  $x_i^{(L)}$  change through time in a manner that idealizes parametric properties of many apparent motion phenomena. More generally, we suggest that the competitive process idealized by (13) and (14) performs a partial contrast enhancement of its input pattern and thereby generates a motion signal whose breadth across space increases with the breadth of its inducing pattern.

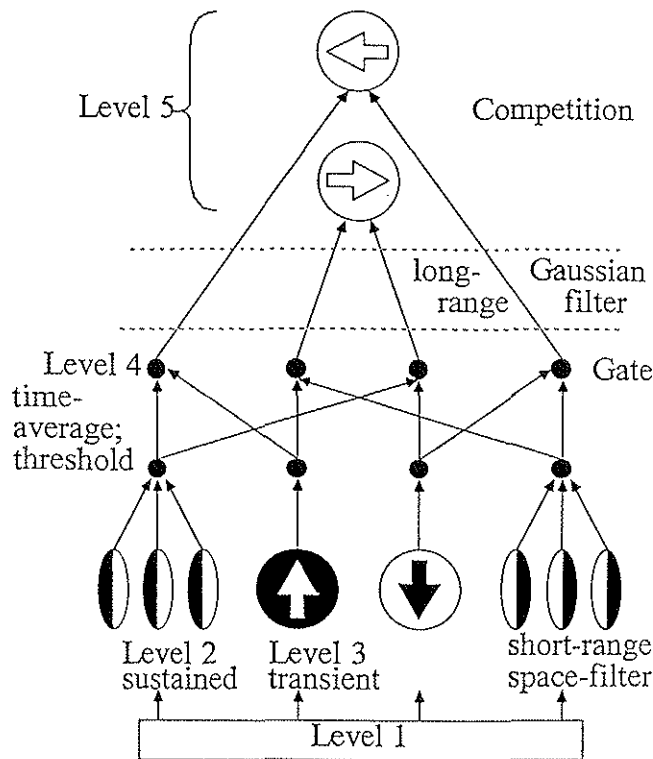
The total MOC Filter design is summarized in Figure 21.

## 12. Gamma Motion: The Apparent Expansion of a Spot at Onset and its Contraction at Offset

When either a light spot on a dark background or a dark spot on a light background is turned on, it appears to expand. When the spot is turned off, it appears to contract. This phenomenon is called *gamma motion* (Kenkel, 1913; Bartley, 1936, 1941; Kolers, 1972). The explanation of gamma motion is a challenge for any model of apparent motion, because it is a case in which nothing actually moves, yet movement is seen. More specifically, it indicates that transient activity in the stimulus is sufficient to generate a motion percept; and that the directionality of this motion percept is dependent on the direction-of-contrast of the local contrast signal, since opposite edges of the stimulus appear to move in opposite directions.

Gamma motion is an exception to the rule that apparent motion fills in missing knowledge regarding an ecological event in order to compensate for a poor spatiotemporal sampling of the stimulus (see, for example, Watson and Ahumada, 1983; Watson, Ahumada, and Farrell, 1983, 1986), which otherwise holds in many cases. In this sense, gamma motion seems “more illusory” than some other forms of apparent motion. Probably for this reason, both Bartley (1941) and Kolers (1972) were inclined to believe that the mechanisms responsible for gamma motion were different from the mechanisms which produce other types of illusory motion. We instead explain gamma motion using the same neural network model that we apply to explain a broad set of apparent, as well as real, motion phenomenon.

The manner in which gamma motion is generated by the MOC Filter is illustrated in Figure 22. In the figure, the stimulus input consists of a 1-D light patch superimposed on a dark background. At Level 2 of the model, an *L*-type sustained cell (sensitive to a light-dark

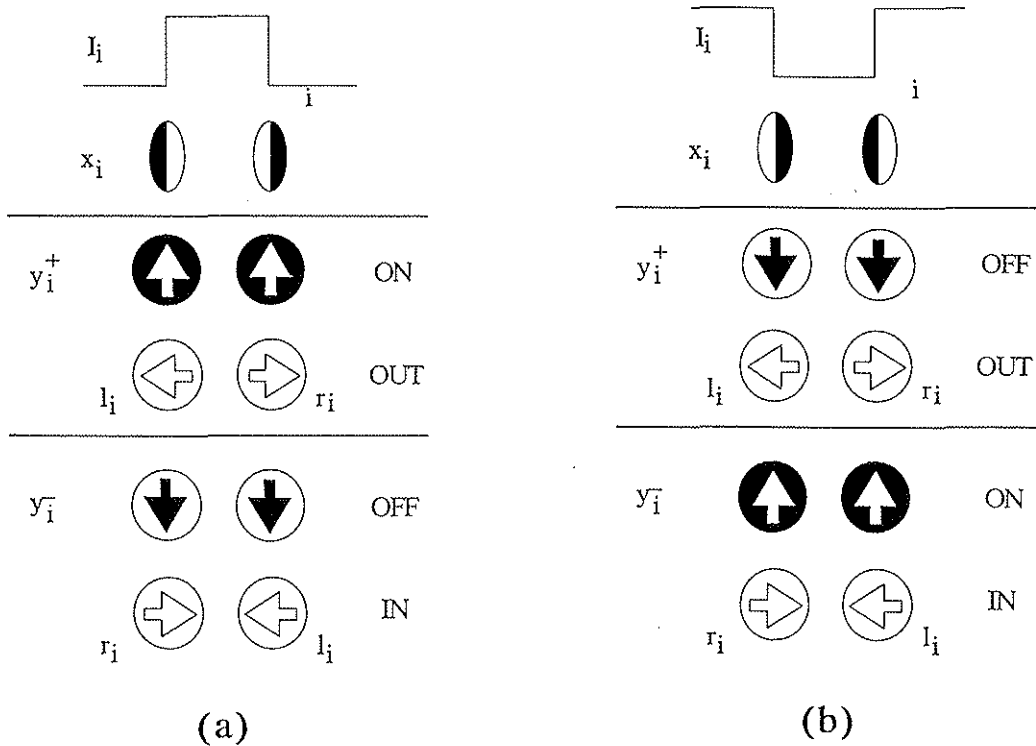


**Figure 21.** The MOC Filter. The input pattern (Level 1) is spatially and temporally filtered in parallel by both sustained response cells with oriented receptive fields that are sensitive to direction-of-contrast (Level 2), and transient response cells with unoriented receptive fields that are sensitive to the direction of contrast change in the cell input (Level 3). Level 4 cells combine sustained and transient cell signals multiplicatively and are thus rendered sensitive to both direction-of-motion and direction-of-contrast. Level 5 cells sum across space and across two types of Level 4 cells to become sensitive to direction-of-motion but insensitive to direction-of-contrast.

contrast pattern in the stimulus) responds at the right edge of the input profile, whereas an *R*-type sustained cell (sensitive to dark-light contrast) responds at the left edge. At the onset of the stimulus, the time derivative of the unoriented cell responses at Level 3 of the model [equation (6)] goes positive, creating responses in the on-cells located at both edges of the luminance patch.

These on-cells gate the responses of the sustained cells located at their same positions to create the activities of the local left and right motion sensitive cells at Level 4 of the model. The on-cell activity at the location of the right edge of the stimulus combines with an *L*-type sustained cell response at that edge to produce a rightward motion signal; while the on-cell activity at the left stimulus edge combines with the activity of an *R*-type sustained cell to produce a leftward motion signal (Figure 22a). Because the on-cells are thresholded, as in (6), these local direction-of-motion signals will be active only as long as the on-cell activities are superthreshold.

The cells at Level 5 of the model receive activations from either local right motion, or local left motion, cells at Level 4, but not both. The effect of the Gaussian smoothing and subsequent sharpening at Level 5 is trivial, because at Level 4 there is activity at only one spatial location in each of the left and right motion channels at any instant. Thus, the output of Level 5 looks like the output of Level 4. The Level 5 output signals an apparent



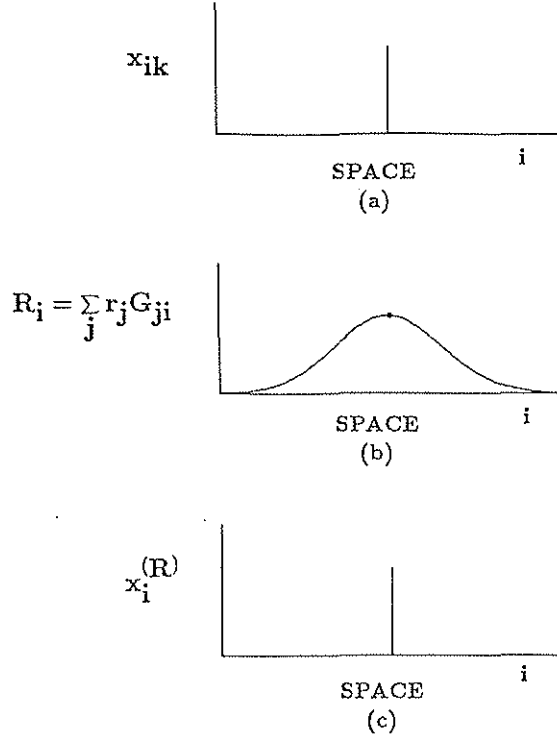
**Figure 22.** Gamma motion. The onset of either (a) a light flash on a dark background or (b) a dark flash on a light background produces an illusion of apparent expansion; an apparent contraction occurs in both cases at stimulus offset. The MOC Filter produces these responses in its output as a result of combining sustained cell responses at the stimulus edges with on- and off-transient activities to create local motion signals.

expansion of the luminance pattern at onset. When the pattern is shut off, the transient on-cell activations at both stimulus edges are replaced by off-cell activations. The result is an apparent contraction of the stimulus (Figure 22a).

The onset of a dark object on a light background reverses the locations of the input edges at which the  $L$ -type and  $R$ -type sustained cells are activated (Figure 22b). In addition, the onset of such an object activates off-cells rather than on-cells, as in equation (7). The reversal of sustained cells combines with the switch from on-cells to off-cells to again make the object appear to expand at onset, as did its counterpart of opposite contrast. The reader may verify that the offset of a dark object on a light background leads to the percept of contraction. In this way, the basic psychophysical observations concerning gamma motion (Bartley, 1936) are successfully mimicked by the model.

### 13. Continuous Motion Paths Generated by Stationary Flashes

In this section we show how a continuous motion signal can be generated between the locations of discrete stroboscopic flashes (Grossberg and Rudd, 1989c), in order to keep our discussion self-contained. In our simulations, an approximation to a continuous motion signal is generated by the MOC Filter whenever more than one of the functions  $x_i^{(R)}, x_{i+1}^{(R)}, x_{i+2}^{(R)}, \dots, x_{i+n}^{(R)}$  are activated sequentially through time, or alternatively, the functions  $x_i^{(L)}, x_{i-1}^{(L)}, x_{i-2}^{(L)}, \dots, x_{i-n}^{(L)}$  are sequentially activated. The goodness of the approximation depends only on the resolution of the simulation, which can be chosen to be arbitrarily fine.



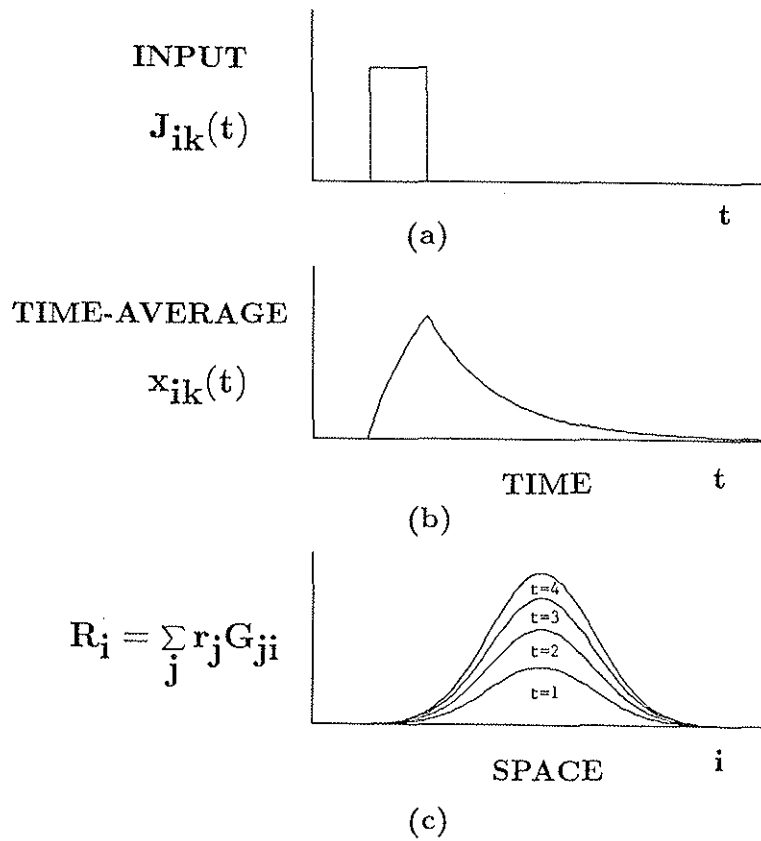
**Figure 23.** Spatial responses at various levels of the MOC Filter to a point input. (a) Sustained activity of a Level 2 cell. (b) Total input pattern to Level 5 after convolution with a Gaussian kernel. (c) Contrast-enhanced output of Level 5 centered at the location of the input maximum.

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. This broad activation pattern (Figure 23b) is generated by the long-range Gaussian filter (10) with kernel  $G_{ij}$  in response to a spatially localized input feature that activates the Level 2 and Level 3 filters at position  $i$  (Figure 23a). The sharp localization of the activities  $x_i^{(R)}$  and  $x_i^{(L)}$  is a result of the contrast-enhancing competitive interaction at Level 5 of the model (Figure 23c). In response to a pair of successive flashes, a continuous motion signal can be generated as the output of Level 5 whenever there is a sufficient overlap between both the spatial and temporal components of the network responses to the separate flashes; that is, when the left or right motion-signaling Gaussian activations generated by the two flashes overlap sufficiently across space, and the corresponding temporal motion signal profiles overlap sufficiently in time.

To understand why this is so, suppose that two successive flashes occur at positions  $i = 0$  and  $i = W$ . Also suppose that the activity  $r_0(t)$  in (8) generated by the first flash is decaying at the same time that the activity  $r_W(t)$  generated by the second flash is growing. If the spatial patterns  $r_0 G_{0i}$  and  $r_W G_{Wi}$  overlap sufficiently, then the total input

$$R_i = r_0 G_{0i} + r_W G_{Wi} \quad (15)$$

to the  $i$ th cell in Level 5 can change in such a way that its maximum value  $x_i^{(R)}(t)$  in (13) occurs sequentially at the positions  $i = 0, i = 1, i = 2, \dots, i = W$ . The result is a percept of continuous motion between the positions of the first and second flashes.

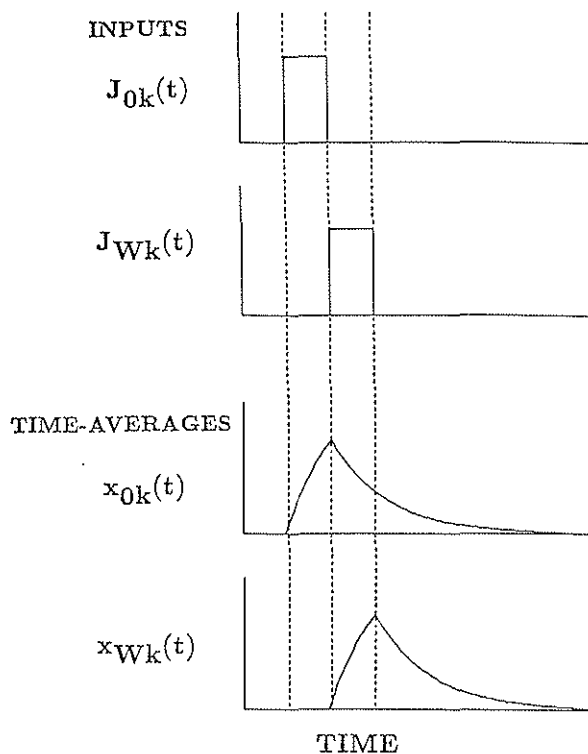


**Figure 24.** Temporal response of the MOC Filter to a point input. (a) The input is presented at a brief duration at location 1. (b) Sustained cell activity at 1 gradually builds after the input onset, then decays after offset. (c) Growth of the input pattern to Level 5 with transient cell activity held constant. The activity pattern retains a Gaussian shape centered at the location of the input, that waxes and wanes through time without spreading across space.

This property of the MOC Filter is illustrated in Figures 24-28. In Figure 24a is shown the temporal profile of the input to the sustained cell centered at position  $i$  that is generated by, say, the first flash. Both  $L$ -type and  $R$ -type sustained cells may contribute to the generation of a rightward motion signal, so we employ the generic subscript  $k = L, R$  here. In Figure 24b is plotted the activity  $x_{0k}(t)$  of the sustained cell at position 0 in response to the input. The characteristics of  $x_{0k}(t)$  which are of interest here are the gradual rise of the response after the onset of the input, and the exponential decay of activity after the signal is turned off. Thus activation persists after the input terminates. Anstis and Ramachandran (1987) have experimentally measured such persistence, and have called it visual inertia (Figure 17).

Assume for simplicity that the transient activity which gates this sustained response is always 'on' and is fixed at a value of 1. This assumption is adopted here to simplify the discussion of how the waxing and waning of sustained cell responses control the motion percept. Given this assumption, the temporal response of the local right motion signal  $r_0(t)$  induced by the input  $J_{0k}(t)$  is equal to  $x_{0k}(t)$ . After the Gaussian convolution, the total input to Level 5 induced by the flash will have the profile of a Gaussian function centered at position 0. The height of this profile changes through time at all spatial positions in proportion to the temporal profile of the local right motion signal generated by the flash.

The growth of the Level 5 input due to the flash is illustrated in Figure 24c. After the offset of the flash, the level of activation decays and the pattern of growth shown in the figure is reversed. The important thing to note about the change of activation over time is

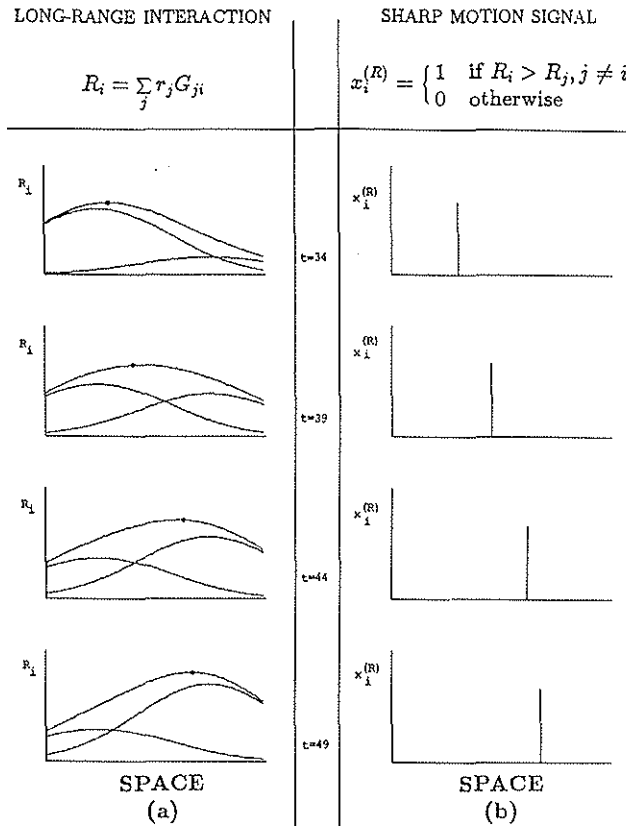


**Figure 25.** Temporal response of the sustained cells at Level 2 to two brief successive point inputs at locations 0 and  $W$ . For an appropriately timed display, the decaying response at position 0 overlaps in time the rising response at position  $W$ .

that the position of the maximum activity across space does not move. Nor does the spatial scale of the reaction spread through time, as would occur if activity diffused across the net. However, if a second flash occurs in the vicinity of this first flash before the activity due to the first flash has fully decayed, the position of the maximum of the total activation can move continuously through time from the first flash position to that of the second flash.

The effect of a temporal overlap of the responses to a pair of inputs is illustrated in Figure 25. In this simulation, the offset of a flash at position 0 is immediately followed by the onset of a flash at position  $W$ . In the upper portion of the figure are shown the successive inputs to the sustained cells at 0 and  $W$ . The time-averaged outputs of these cells are shown in the lower portion of the figure. In this example, the activity of the cell at  $W$  is growing during the same period that the activity of the cell at 0 is decaying.

If the flashes occur sufficiently close to one another in space relative to the width of the Gaussian kernel  $G_{ij}$ , then a traveling wave of activation occurs in the total input (11) to Level 5, as is illustrated in Figure 26a. Going down the page, the frames in the figure represent the spatial configuration of  $R_i$ , the total input to Level 5 at position  $i$ , at successive times following the offset of the first flash. In each frame, three patterns of activity are shown: the total input  $R_i$ , and the two Gaussian components of  $R_i$  which are generated by the individual flashes. The component of  $R_i$  caused by the first flash is largest at the moment of the offset of the first flash, and decays thereafter. While this first component is decaying, the component due to the second flash grows until the second flash is turned off. The sum of the two components changes in such a way that its maximum  $x_i^{(R)}$  across space travels



**Figure 26.** Simulated MOC Filter response to a two-flash display. Successive rows correspond to increasing times following the Frame 1 offset. (a) The two lower curves in each row depict the total input to Level 5 due to each of the two flashes. The input due to the left flash decreases while the input due to the right flash increases. The summed input due to both flashes is a traveling wave whose maximum value across space moves continuously between the two flash locations. (b) Position over time of the contrast-enhanced Level 5 response. Spatial axis is 128 units. Flashes are both of width 12, with left edges at positions 25 and 89. Frame 1 offset time = 32;  $ISI = 0$ ;  $A = .05$ ,  $B = 0$ ,  $K = 42$ . Transient cell activities held constant at the value 1.

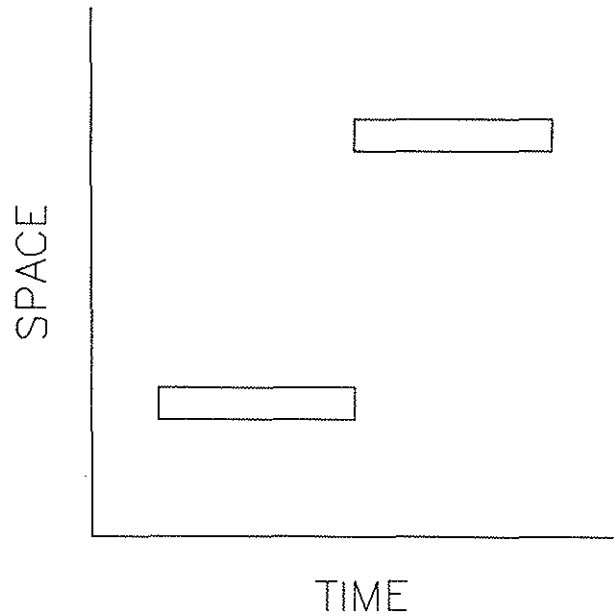
continuously from the position of the first flash to that of the second (Figure 26b).

In summary, the time- and space-averaged responses to individual flashes do not change their positions of maximal activation through time (Figure 24c). In this sense, ‘nothing moves’. When a series of properly timed and spaced flashes is presented, however, the sum of the temporally and spatially averaged responses that they generate can produce a continuously moving peak of activity between the positions of the stroboscopic flashes.

#### 14. Relationship between Flash Spatial Separation and Spatial Scale: Motion Speed-Up and Partial Motions

In a classic apparent motion study with two-flash displays, Neuhaus (1930) showed that there is a restricted range of interflash spatial separations over which apparent motion can be induced. He found that this range extended from close stimulus separations to about 4 degrees of visual angle. In other studies, apparent motion has been reported over separations of up to about 7 degrees (Anderson and Burr, 1987). The range of distances over which apparent motion operates is known to depend on properties of the flashes, such as their durations (Neuhaus, 1930) and spatial scales (Anderson and Burr, 1987).

The MOC Filter also generates a continuous motion path within a restricted range of



**Figure 27.** Space-time diagram of a two-flash apparent motion display. The input is a  $32 \times 32$  matrix of luminance values. Rectangular outlines indicate the spatiotemporal boundaries of a Frame 1 flash of width 3 centered at position 3 and presented from times 4 through 16; and a Frame 2 flash of width 3, centered at 24 and presented from times 16 through 28.

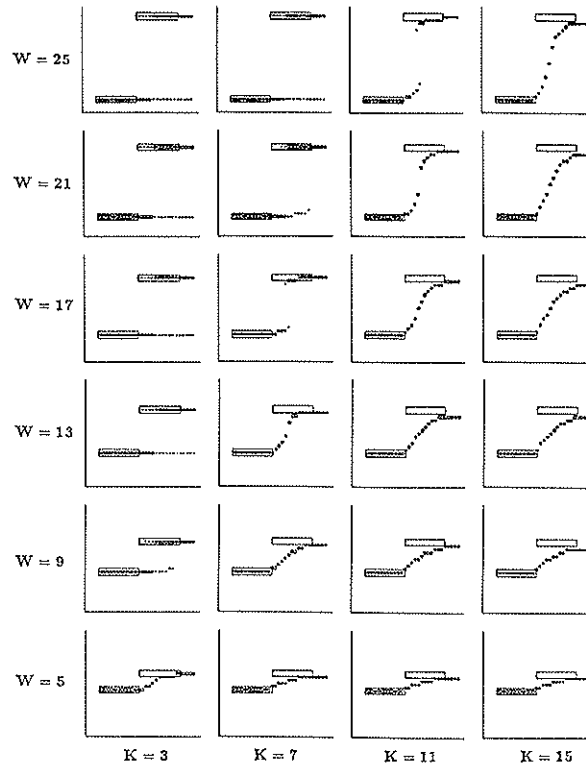
flash separations. This range depends on the size of the Gaussian receptive fields of the Level 5 cells. Grossberg and Rudd (1989c) proved mathematically that a continuous motion path is generated only when the distance between the flashes is less than or equal to twice the value of the spatial width parameter  $K$  of the Gaussian kernel  $G_{ij}$  in (10). In other words,

$$W_{\max} = 2K, \quad (16)$$

where  $W_{\max}$  denotes the upper limit of the range of spatial separations between flashes that can produce continuous motion.

Computer simulations of this and other apparent motion phenomenon are displayed using the scheme illustrated in Figure 27. The rectangular outlines in the figure represent the spatiotemporal boundaries of the stimulus; here, the flashes in a two-flash display. In displaying the results of our computer simulations, we superimpose the paths of the Level 5 outputs  $x_i^{(R)}$  or  $x_i^{(L)}$  on this diagram.

The simulation results displayed in Figure 28 demonstrate the existence of a maximum spatial separation for producing continuous motion; in particular, the  $W_{\max} = 2K$  rule. Going down the columns of the figure, the spatial separation of the flashes is reduced; across rows, the width of the Gaussian filter is increased. In the space-time diagrams displayed in the lower right-hand portion of the figure, the size of the Gaussian filters is large enough to produce a spatial overlap in the network responses to the closely spaced flashes, and a continuously moving wave of activity results from the first flash to the second flash. Note



**Figure 28.** Paths of the MOC Filter output as a function of flash separation  $W$  and Gaussian filter width  $K$ . The rectangular outlines in each panel indicate the spatiotemporal flash boundaries. Large circles indicate locations of the global maximum of the right-motion signal pattern  $R_i$  at 32 time steps. Small circles indicate locations of other local maxima of  $R_i$ . A continuous motion signal path is generated when  $W < 2K$ . Flashes are all 3 pixels wide, with temporal coordinates as in Fig. 24. Parameters  $A = .12$ ;  $B = 0$ . Transient cell activities = 1 throughout.

that the motion wave speeds up as the spatial separation of the flashes increases while the ISI is held constant. In the upper left-hand portion of the figure, the spatial separation of the flashes is too large to be spanned by the Gaussian filters, so the network activations produced by the separate flashes do not combine to create a single moving wave. Instead, their maxima across space remain distinct, rising and falling in place over time as in Figure 24c, corresponding to the percept of blinking in place. Near the value of the critical spatial separation, a partial motion occurs that is similar to reports in the literature of partial motion percepts which occur when stimulus parameters are near the threshold for producing apparent motion (Wertheimer, 1912; Kolers, 1972, p.9).

Because the production of a continuous motion signal from discrete flashes does not depend on the gating of the sustained cells by transient signals, we fixed the value of the transient signals at 1 throughout this simulation. This causes the paths of the left motion and right motion outputs of the model to become degenerate. Then, when a single flash is turned on or off, a single maximum of activity is produced at the location of the center of the flash. The full model would produce a pair of either outward- or inward-directed gamma motions, depending on whether the flash had just been turned on or off. Subsequent sections demonstrate how gating by transient cells may modify properties of the traveling wave in a manner that conforms to challenging data that have thusfar received no other explanation.

## 15. Equal Time Multiple Scale Motions

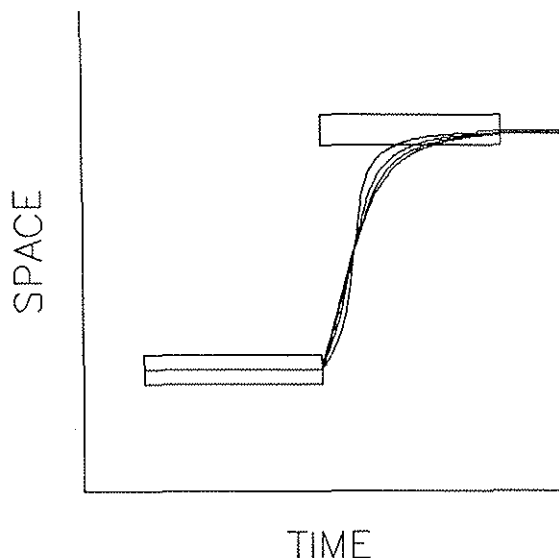
Do MOC Filters which possess different size Gaussian receptive fields at Level 5 produce similar apparent motion paths in response to the same stimulus? The problem of integrating motion signals from multiple scales is one which the brain has solved. For example, single cell recordings from cortical area MT, which is known to be involved in the processing of motion signals (Albright, Desimone, and Gross, 1984; Maunsell and Van Essen, 1983; Newsome, Gizzi, and Movshon, 1983; Zeki, 1974a, 1974b), indicate the existence of direction sensitive cells with large receptive fields of varying size. How are the signals from these motion channels of differing scale integrated into a global motion signal?

As we have shown, in response to two-flash displays with widely separated flashes, MOC Filters at some scales may signal continuous motion while others signal only gamma motion. Thus the bank of parallel multiple scale filters, considered as a group, can signal continuous motion as well as the apparent expansion and contraction of the individual flashes. These are not mutually exclusive percepts, and the fact that different MOC Filter scales carry different motion information is analogous to the fact that spatial frequency channels at different SOC Filter scales carry different information about stimulus form.

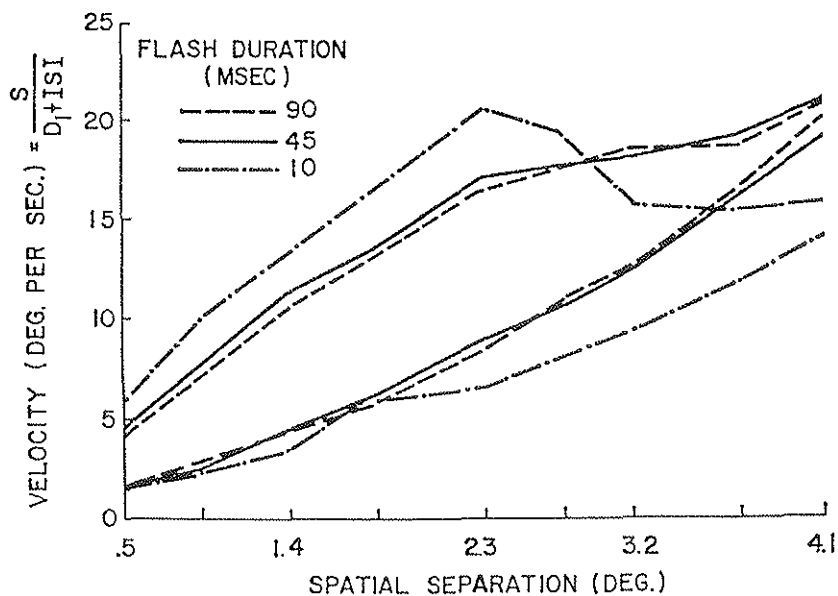
The apparent motion paths generated by four scales of MOC Filters which signal continuous motion in response to the identical two-flash display are shown in Figure 29. The apparent motion paths are indicated by smooth curves. The paths generated at different scales are almost identical. Because the deviations are negligible all along the path, they can be synthesized into a consistent multi-scale motion signal at a later processing stage. Remarkably, these motion paths intersect the point which lies halfway between the flashes at the same time (Grossberg, 1977; Grossberg and Rudd, 1989c). This *Equal Half-Time Property* also applies to the situation in which the spatial separation of the flashes is manipulated while the scale of the MOC Filter is held constant. These properties suggest an explanation of the classical empirical observation (Figure 30) that "large variations in distance are accommodated within a near-constant amount of time" (Kolers, 1972, p. 25). Grossberg and Rudd (1989c) describe the mathematical proofs from Grossberg (1977) of the Equal Half-Time Property and the Acceleration-Deceleration of apparent motion pathways.

None of the motion paths in Figure 29 represents a motion signal of constant velocity, which would be indicated by a straight line on the diagram. Instead, the motion path computed at each filter scale accelerates away from the location of the first flash, followed by a deceleration towards the location of the second flash. The acceleration-deceleration property sheds light on percepts of partial motion. As illustrated in Figure 28, given any spatial separation  $W$  of two flashes, the minimum Gaussian scale size  $K$  that can support a continuous motion percept satisfies  $K = \frac{W}{2}$ . For fixed  $W$ , as  $K$  approaches this critical value from above, the slope of the motion path increasingly steepens, indicating a high velocity signal. It is observed experimentally that near-threshold stimulus conditions are associated with the perception of an accelerating motion away from the first flash, followed by a disconnected deceleration into the location of the second flash (Kolers, 1972, p.9). We suggest that such partial motions may result when high velocity movements of the activity peak exceed the spatiotemporal processing limitations of the neural mechanisms at Level 5 or beyond.

## 16. 'The Less You See It, The Faster It Moves'

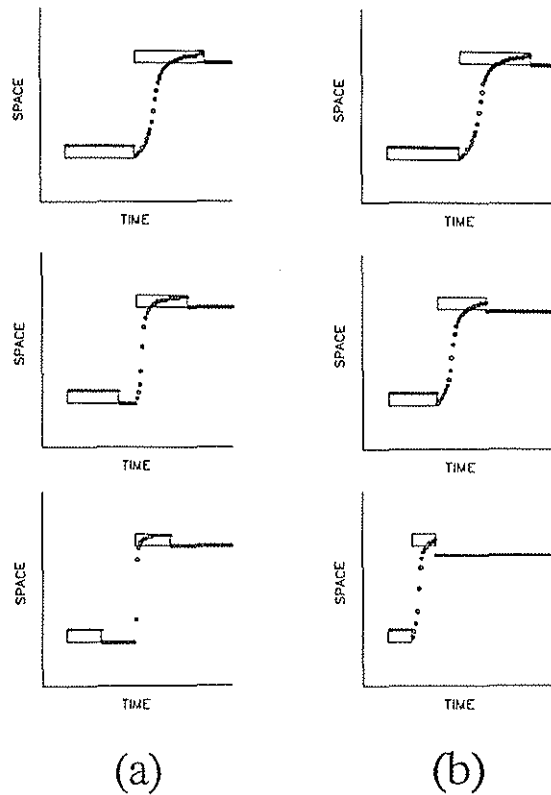


**Figure 29.** Motion paths generated by MOC Filters with different Gaussian filter kernel widths  $K$ . Intersection of the paths occurs at a point halfway between the two flash locations (Equal Half-Time Property). The stimulus was a  $320 \times 320$  luminance matrix. Flash widths equal 21, centered at 80 and 240. Durations equal 121; onsets at time 40 and 160. Parameters  $A = .03$  (equivalent to  $A = .3$  for a  $32 \times 32$  simulation);  $K = 90$  (steepest sigmoid), 110, 130, 150 (shallowest sigmoid). As  $A \rightarrow \infty$ , paths corresponding to different  $K$  converge.



**Figure 30.** Calculated velocity of the spots which generated the motion threshold data shown in Fig. 6. The velocity of the apparent motion increases with flash separation so that large variations in distance are accommodated within a near-constant amount of time. [Figure reprinted from Kolers (1972). Used with permission of Pergamon Press.]

Giaschi and Anstis (1989) measured the apparent speed of motion produced by a continuously cycling two-flash display as a function of the on-time of the flashes. They found that shorter flash durations were associated with higher judged motion velocities. The MOC Filter also produces this effect. In the series of simulations illustrated in Figure 31, the response



**Figure 31.** Simulation of the finding of Giaschi and Anstis (1989) that the apparent velocity of the motion signal produced by a two-flash display increases with decreasing flash duration. (a) With the SOA held constant, a decrease in the duty cycle produces a higher velocity signal (steeper slope on the space-time diagram). (b) With ISI held constant at zero, a decrease in the cycling rate produces a higher velocity signal. In these simulations, flash duration is the only temporal variable that is consistently associated with the observed velocity changes. Input matrix: 128x128. Flash 1 parameters: edges=29, 37; in column 1, on from time 17 through (down column) times 40, 51, and 63; in column 2, on from time 17 through (down column) times 32, 48, and 63. Flash 2 parameters: edges 92, 100; in column 1, on from time 64 through times 87, 98, and 110; in column 2, on from times 33, 49, 64 through 48, 80, and 110. Model parameters:  $A, C = .04, E = 0; D = 1; K = 40$ .

of the model to a single cycle of such a display is shown as a function of flash duration. In Figure 31a, the duty cycle of the display is manipulated by reducing the on-times of the flashes in steps relative to the length of the SOA. As the flash duration is reduced, the slopes of the paths of the moving wave maxima become progressively steeper on the space-time diagram, indicating an increase in the velocity of the MOC Filter output.

Giaschi and Anstis performed several control experiments to verify that it was the reduction of the flash duration, rather than an increase in the ISI *per se*, that produced the empirical apparent velocity increase. We also checked to make sure that this was the case for the MOC Filter results. In order to demonstrate this, we ran a second simulation series in which the ISI was held constant at zero while the flash duration—and thus the SOA of the display—was manipulated. The results are shown in Figure 31b. In the simulation results, the MOC Filter again generates a higher velocity signal when the flash duration is short. Here, a faster moving wave is associated with a shorter SOA and fixed ISI, while in the previous simulation it was associated with a longer ISI and a fixed SOA. In both cases higher velocity signals are associated with shorter flash on-times. Thus, it is the flash duration rather than the ISI or SOA that determines the speed of the motion signal produced by the

MOC Filter, as it was the flash duration that determined the speed of the motion percept reported by the observers in the Giaschi and Anstis experiment.

## 17. Split Apparent Motion

Results such as motion speed-up with decreasing ISI and with shorter flash duration show that the early stages of biological motion processing cannot be velocity detectors *per se*, but rather are sensitive to subtle combinations of stimulus intensity, duration, and spatial relationships. The phenomenon of *split motion* (DeSilva, 1926) shows, in addition, that motion processing does not necessarily select a globally preferred direction of motion. Under the appropriate experimental conditions, apparent motion can be observed to occur simultaneously along competing pathways (Kolers, 1972). Split motion can be observed when a single flash presented in Frame 1 is followed by a pair of flashes in Frame 2. If the two flashes which are presented in the second frame are alike in all respects (e.g., size, luminance, orientation, shape), and are equidistant from the location of the Frame 1 flash, then the first flash will appear to 'split' and move simultaneously to both of the Frame 2 flash positions.

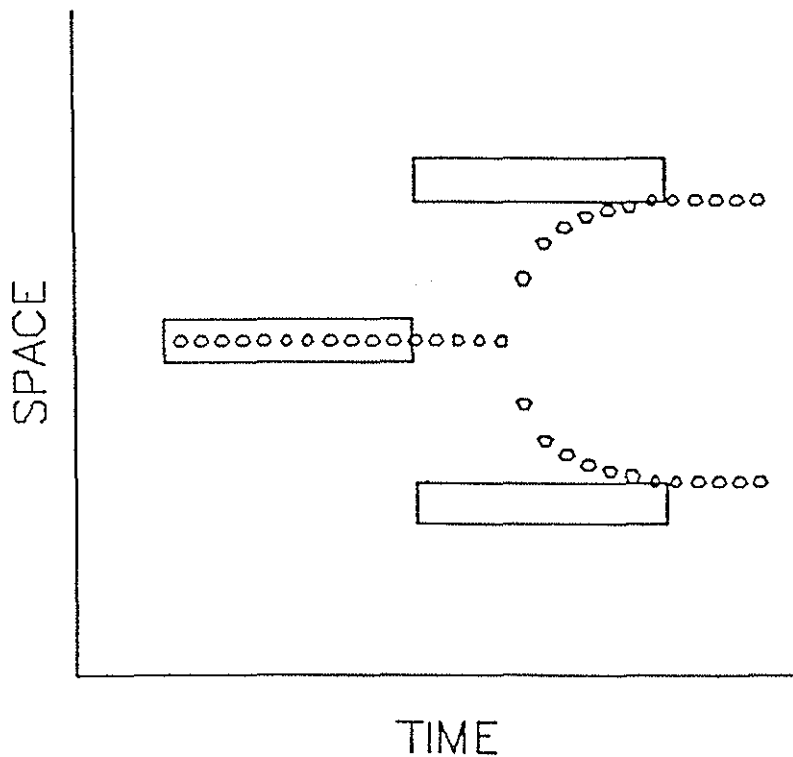
A MOC Filter simulation of split motion is illustrated in Figure 32. In this simulation, the value of the transient cell activities are again fixed at 1. A single maximum of activity is located at the position of the Frame 1 flash at the end of that frame. After the onset of Frame 2, this single maximum divides into two separate maxima, which then follow separate paths to each of the two Frame 2 flash locations.

Split motion will occur in the MOC Filter model whenever the conditions which lead to an apparent motion signal act to create motion paths of equal strength, but in opposing directions. Split motion is a challenging phenomenon for the motion theorist because it eliminates any candidate motion mechanisms which compute a unique direction-of-motion signal. In the MOC Filter, it is a matter simply of motion waves traveling in different, noncompeting directions.

## 18. Ternus Display: Group Motion

Ternus motion percepts probe more deeply the existence and ordering of MOC Filter processing levels. Indeed, prior to our explanation of these percepts, many scientists believed that the switch from the group motion percept to the element motion percept must somehow depend upon a prior stage of object segmentation, or even cognitive processing. Our theory explains these key properties of Ternus motion as manifestations of early motion filtering. On the other hand, *all* levels of the MOC Filter are needed to explain the full range of Ternus data. This fact supports the hypothesis that no fewer than the MOC Filter levels can exist *in vivo*. In Figure 30 the basic stimulus in the Ternus experiment is illustrated using our space-time diagram convention. Recall from Section 9 that group motion (all three elements move as a whole) is typically observed when the ISI of the Ternus display is brief (as in Figure 33a); and element motion (one element jumps across the other two which remain perceptually stationary) is typically observed when the ISI is longer (Figure 33b).

In the series of computer simulations illustrated in Figure 34, we investigated the dependence of the group motion percept on the Gaussian filter scale  $K$ . We also held the transient activities at the value 1 in order to demonstrate that, in the absence of transient cell gating, the model generates percepts of simultaneity, group motion, and succession as the ISI is increased, but no element motion.

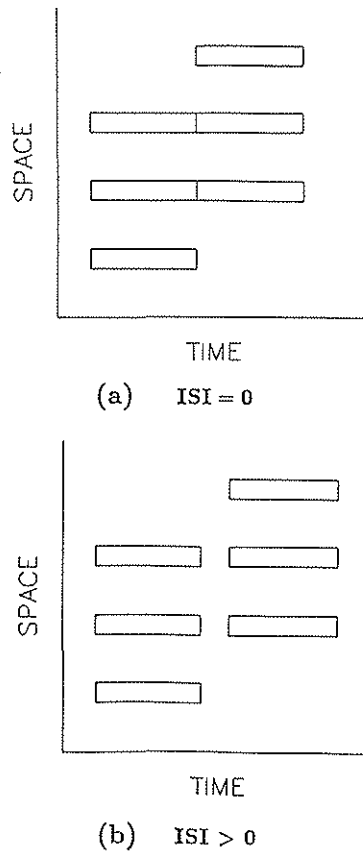


**Figure 32.** Simulation of split apparent motion. In Frame 1, a single flash is presented, followed in Frame 2 by a pair of flashes which are equidistant from the first flash. With the value of the transient signals fixed at 1, a single maximum of activity across space is observed at the center of the Frame 1 flash for the duration of that frame. After the onset of Frame 2, this maximum splits into two separate local maxima which follow separate paths to each of the Frame 2 flashes. Input matrix:  $128 \times 128$ . Frame 1: edges at 60, 68; on from time 17 through time 63. Frame 2: edges at 29, 37, 91, 99; on from time 64 through time 110. Parameters  $A = .04$ ;  $B = 0$ ;  $K = 22$ . Transient activities = 1 throughout.

In Figure 34, no continuous motion percept is generated when the filter scale  $K$  is too small to span the distance between the flashes, as expected from our analysis of two-flash displays. When  $K$  exceeds a critical value, a single continuous motion path is observed. This path begins at the center of the group of elements presented in Frame 1 and travels to the center of the group of elements presented in Frame 2. We identify this apparent motion path with the percept of group motion. The explanation of this phenomenon and its identification with group motion is illustrated in Figure 35.

In Figure 35a, three simultaneous point flashes are represented. Figure 35b represents their individual and total inputs to Level 5. This figure shows that any Gaussian filter scale which is sufficient to span the interframe distance between the furthest elements in the Ternus display will also produce sufficient overlap in the network activations induced by the individual elements to form a single unimodal pattern of input to Level 5. After the sharpening of the input at Level 5, this broad pattern of activation generates a continuous motion path, centered within the three flashes, which moves in the manner perceived during group motion.

In Figure 35c, the sharpened output from Level 5 is represented as a winner-take-all position centered in the middle of the three flashes. More generally, we suggest that the relatively flat shape of the total input to Level 5 in Figure 35b would generate a blob of



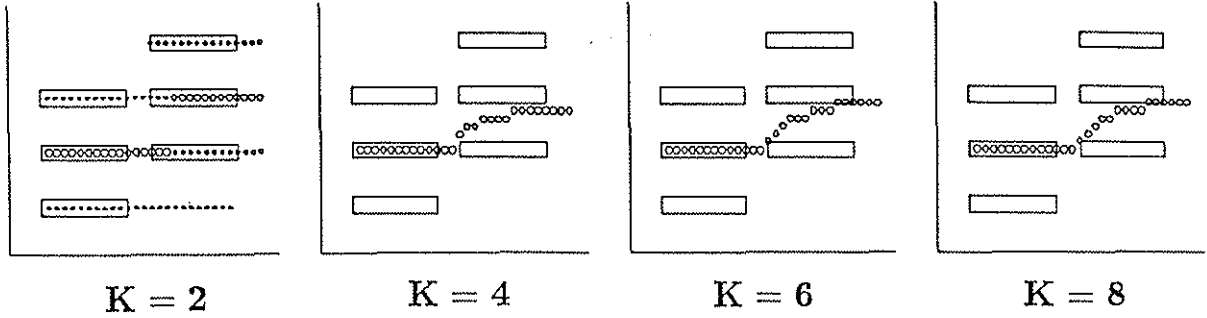
**Figure 33.** Space-time diagrams of two Ternus displays. Rectangular outlines indicate flash boundaries. (a)  $ISI = 0$ . (b)  $ISI > 0$ .

motion centered in the middle of the three flashes.

### 19. Ternus Display: Element Motion and Transient Cell Gating of Sustained Cells

Production of element motion by the MOC Filter depends on the gating of the sustained cells by the transient cell activities which vary in time according to equations (6) and (7). To see this intuitively, consider the separate cases of Ternus displays in which the ISI either is zero (Figure 33a), or some positive constant (Figure 33b). When the ISI is zero, the input at the locations of the two central elements in the display is unchanged during the transition from Frame 1 to Frame 2. Thus, no off-cell activity will be generated by the offset of the first frame at these locations; nor will any transient on-cell activity be generated by the onset of the second frame. Because of the transient cell gating of the sustained cells, occurrence of a transient cell response is required for the production of local motion signals at these locations. Thus, the lack of transient responses during the interframe transition has the effect of gating off the contributions that the sustained cells at these locations would have made to the total motion signal had the transient cells been active.

When the ISI of the Ternus display is zero, two of the three components of total network activity, which together can generate a group motion signal (as in Figure 34), are gated off during the interframe transition. Only the leftwardmost element in Frame 1 and the rightwardmost element in Frame 2 can then contribute to the apparent motion signal, since these are the only positions at which transient signals occur during the interframe transition. With the potential contributions of the central elements gated off, the MOC Filtered image



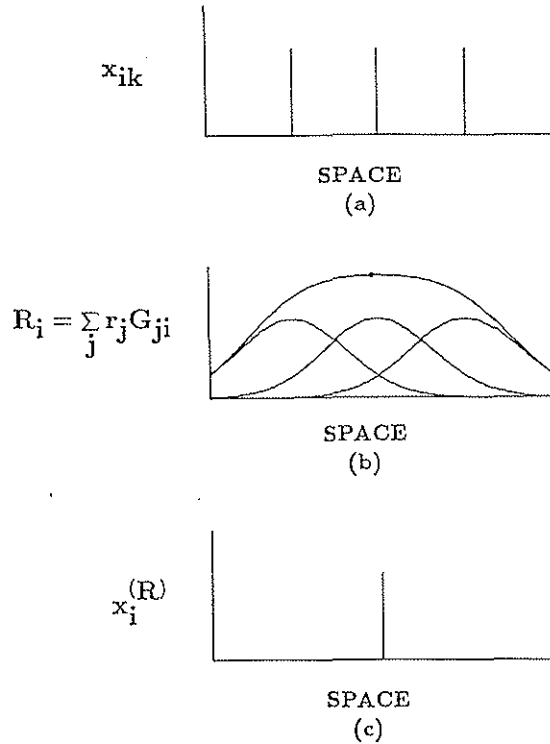
**Figure 34.** Generation of group motion by a Ternus display. Space-time paths of global maxima (large circles) and other local maxima (small circles) corresponding to Gaussian filters of four widths  $K$ . In order to generate apparent motion,  $K$  must be large enough so that responses to the individual flashes will combine to produce a single moving global peak. For such a  $K$ , group motion always occurs if the transient cell activities are fixed at a positive constant value. Input matrix:  $32 \times 32$ . Width of each flash: 3 pixels. Frame 1 flash centers at locations 6, 13, 20. Frame 2 flash centers at locations 13, 20, 27. Frame 1 on from time 4 through time 15. In (a) Frame 2 on from time 17 through time 28; in (b) from time 20 through time 31. Parameters  $A = .12$ ;  $B = 0$ . Transient activities = 1 throughout.

of the Ternus display looks like that of a two-flash motion display. Hence element motion is generated.

On the other hand, when the ISI of the display is sufficiently large for the activities of the transient cells at the central element positions to have time to build, the second and third flash positions will once again contribute to the motion percept, as in Figure 35b, so group motion will occur, as in Figure 34.

The influence of transient cell gating on the temporal profiles of sustained cell responses is illustrated in Figure 36. In response to a brief input  $I_i$ , the time-averaged response  $x_{ik}$  of a sustained cell located at  $i$  first gradually rises, then decays. The transient activity is modeled for simplicity as the time derivative of a similar time-averaged activation. The transient on-cell response  $y_i^+$  is the half-wave rectified positive part of  $dx_i/dt$ ; and the transient off-cell response  $y_i^-$  is the half-wave rectified negative part of  $dx_i/dt$ .

The two Level 4 activities illustrated,  $x_{iLY_i^+}$  and  $x_{iRY_i^-}$ , both signal right motion [as in equation (8)]. If the activity  $x_{iRY_i^-}$  generated by the offset of a flash decays while the activity  $x_{iLY_i^+}$  generated by a later flash is rising, a continuous right motion wave between the two flash locations will be produced, provided that the spatial parameters of the display are also appropriate.



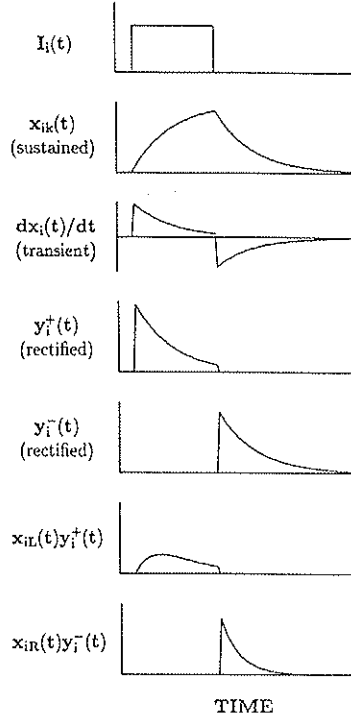
**Figure 35.** Spatial summation of network activations underlying group motion response the Ternus display. (a) Three individual sustained response activations across space generate in (b) Gaussian profiles as input to Level 5 that sum to create a unimodal total input whose maximum value in (c) is centered at the middle of the display elements. 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 frame of the display.

## 20. Simulating the Transition between Element and Group Motion: Same Direction-of-Contrast

The MOC Filter simulations of the Ternus effect utilize all of the mathematical features defined in Section 11. The elements of the Ternus display simulated here are light on a dark background. Because  $L$ -type sustained cells respond to a light-dark pattern and  $R$ -type sustained cells to a dark-light pattern,  $x_{iL}$  and  $x_{iR}$  activations occur at the right and left edges, respectively, of the elements. These responses were gated with the activities  $y_i^+$  and  $y_i^-$  of transient on-cells and off-cells, respectively, to derive the local right and left motion signals  $r_i$  and  $l_i$ . The local motion signals were separately convolved with a Gaussian kernel, as in (11) and (12), to form the separate motion path outputs  $x_i^{(R)}(t)$  and  $x_i^{(L)}(t)$ , as in (13) and (14).

Figure 37 illustrates the main effect of the transient cells on transforming sustained cell reactions into motion signal functions as the ISI is varied, at the two positions which receive flashes during both Frame 1 and Frame 2 of the Ternus display (Figure 15a). For concreteness, we illustrate here only the two functions which contribute to a right motion signal.

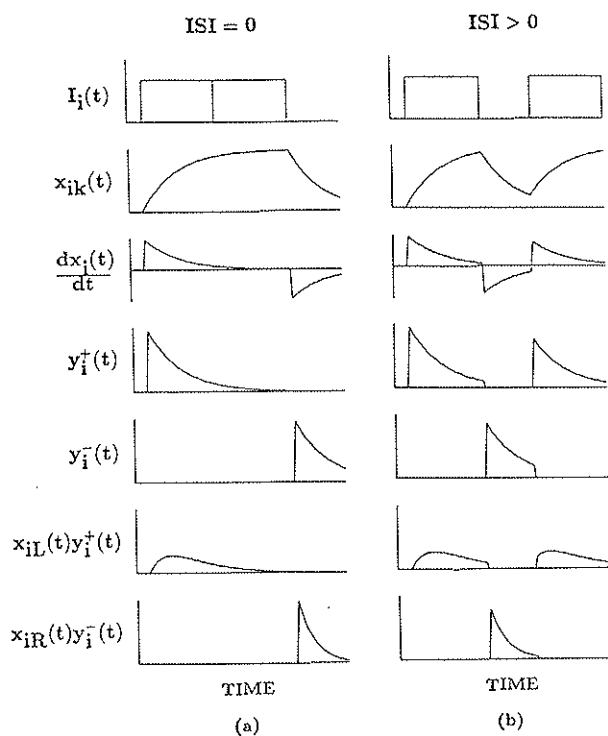
In response to a Ternus display with  $ISI = 0$ , the sustained cell activations  $x_{ik}$  do not have a chance to decay between offset of Flash 1 and onset of Flash 2. In Figure 37a, the



**Figure 36.** Multiplicative gating of Level 2 sustained activities by Level 3 transient activities to generate a direction-of-motion sensitive response at Level 4. Presentation of an input  $I_i$  produces sustained responses  $x_{ik}$  ( $k = L, R$ ) and a transient response  $dx_i/dt$ . The activity  $x_{iL}$  is gated by the rectified on-cell response  $y_i^+$  to generate an  $x_{iL}y_i^+$  response that is sensitive to direction-of-motion and direction-of-contrast; and by the rectified off-cell response  $y_i^-$  to generate an  $x_{iR}y_i^-$  response that signals the same direction-of-motion (rightward). Time axis is 128 units. Parameters  $A, C, D = .12$ ;  $B, E = 0$ ;  $\Sigma_j I_j F_{ji} = J_{iL}$  or  $J_{iR}$ , whichever is nonzero, and  $= 0$  otherwise.

sustained cells respond directly to the inputs  $I_i$ . If, as occurs in vivo, intermediate cellular stages time-averaged the inputs  $I_i$  before generating outputs to the sustained cells, then the same property would also hold at small, but positive, ISIs. Because the sustained cells do not decay significantly during the interflash interval, the on-transient cells  $y_i^+$  and the off-transient cells  $y_i^-$  are inactive during the interflash interval at the Ternus positions that receive two flashes. Thus a motion signal is generated only at the onset of the first flash and the offset of the second flash. Since, as in Figures 25 and 26, a motion signal is generated by interaction of the *off*-response to the first flash with the *on*-response to the second flash, no motion signal is generated at the two positions that receive two flashes. Only the first and fourth positions generate an off-response to the first flash and an on-response to the second flash, respectively. Hence only the combinations of inputs to the long-range Gaussian filter from these positions generate a traveling wave at Level 5. This wave has the properties of element motion.

In contrast, suppose that the ISI is chosen sufficiently large that the transient detectors can respond both to the offset of Frame 1 and to the onset of Frame 2, as in Figure 37b. Then, at each of the three flash positions in Frame 1, a transient off-response is generated at Level 2 when Frame 1 shuts off. All three positions can therefore generate a sustained-transient motion off-response at Level 4. These three Level 4 responses input to the long-



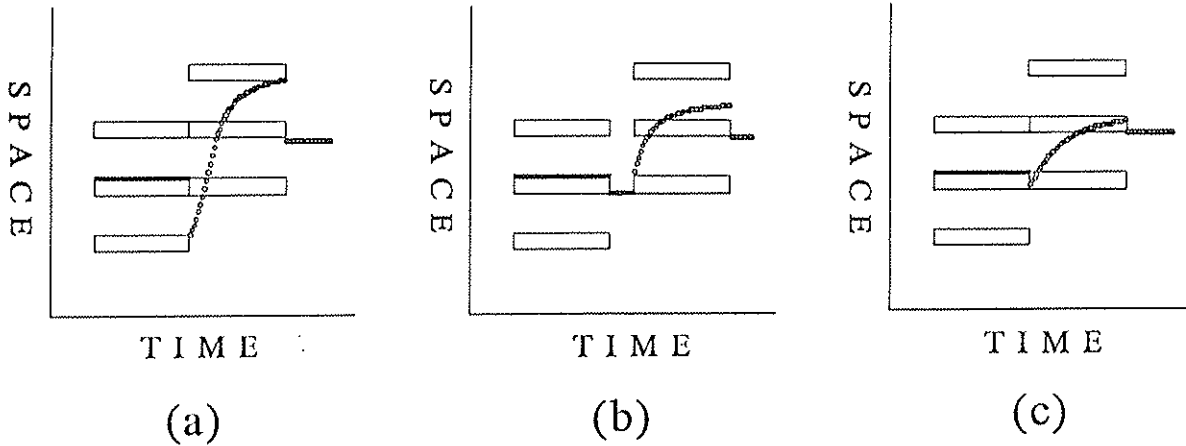
**Figure 37.** Mechanisms for generating group and element motion in the Ternus display. (a) Element motion when the ISI is small: At the locations of Ternus display elements 2 and 3, no transient responses  $dx_i/dt$  are generated at the offset of Frame 1 or the onset of Frame 2. Thus, no contribution to the overall motion signal is made by these elements. Element motion results. (b) Group motion when the ISI is sufficiently large: Gated sustained-transient signals develop at all display locations, including those of Elements 2 and 3; thus all locations contribute to the overall unimodal motion signal, as in Figure 32b. Group motion results. Time axis is 128 steps. Flash durations = 40. In (a)  $ISI = 0$ . In (b)  $ISI = 40$ . Parameters  $A, C, D = .09$ ;  $B, E = 0$ ;  $\Sigma_j I_j F_{ji} = J_{iL}$  or  $J_{iR}$ , whichever is nonzero, and = 0 otherwise.

range Gaussian filter to generate a unimodal total off-input to Level 5 that is centered at the middle flash of Frame 1, as in Figure 35b. The same is true for the total on-response to the onset of Flash 2; except now the total on-input to Level 5 is centered at the middle flash of Frame 2. These off-responses and on-responses combine via the long-range Gaussian filter to generate a traveling wave with properties of group motion at Level 5. The motion paths computed by the model in response to Ternus displays with  $ISI = 0$  and  $ISI > 0$  are displayed in Figure 38a and 38b, respectively.

In summary, the paths of the sharpened Level 5 signals  $x_i^{(R)}$  depend on the ISI in such a way as to mimic the ISI-dependence of the paths of element and group motion in response to the Ternus display illustrated in Figure 15a.

## 21. A Crucial Test: Simulating Group Motion at Short ISIs with Reverse Contrast Stimuli

Pantle and Picciano (1976) showed that group motion occurs even at short ISIs if the relative contrast of stimulus-to-background is reversed between the two successive frames (Figure 15b). This phenomenon is simulated in Figure 38c. Its explanation uses essentially all the processing levels of the MOC Filter, as well as their ordering. The Pantle-Picciano



**Figure 38.** Simulated group and element motion responses to three Ternus displays. (a) Element motion when  $ISI = 0$  and the flashes in the two frames are of the same direction-of-contrast. (b) Group motion when  $ISI > 0$  and the flashes in the two frames are of the same direction-of-contrast. (c) Element motion when  $ISI = 0$  and the flash contrast is reversed between frames. Input matrix  $128 \times 128$ . Element widths 9 pixels. Frame 1 center locations: 12, 48, 84. Frame 2 center locations: 48, 84, 120. Frame durations 56. Frame 1 onset time = 2. ISIs = 0 and 14. Parameters  $A, C, D = .05$ ;  $B, E = 0$ ;  $K = 60$ .  $\sum_j I_j F_{ji} = J_{iL}$  or  $J_{iR}$ , whichever is nonzero, and = 0 otherwise.

effect is thus a strong test of the hypothesis that no fewer set of levels can be used to explain motion data at this level of subtlety.

The main property leading to an explanation is, however, simple. Suppose that the three stimulus dots in Frame 1 are more luminous than their background, whereas the three dots in Frame 2 are less luminous than their background. Onset of Frame 1 then activates transient on-cells, as before. However, onset of Frame 2 activates transient off-cells, no matter how small the ISI is chosen, due to the reversal of contrast between Frame 1 and 2. The offset of Frame 1 can therefore activate off-cells, just as in the case of a large ISI without contrast-reversed frames. All three locations in Frame 1 will therefore influence the formation of a unimodal right-motion signal, centered at the middle of Frame 1 and decaying through time.

The effects of Frame 2 onset require further consideration, because the three stimulus elements generate luminance decrements. Here we use the same properties were used to explain gamma motion in response to an input decrement in Figure 22b. In particular, an input decrement activates sustained cells that are sensitive to the opposite direction-of-contrast from the sustained cells activated by an input increment; compare Figure 19b with 19a. However, these sustained cells are gated by transient off-cells, rather than by transient on-cells. The net effect, as in our explanation of gamma motion, is to generate the same combination of local right motion signals and left motion signals, given either contrast

polarity. In summary, the reverse-contrast Frame 2 enables transient off-cells to be activated despite the short ISI, but these off-cells combine with sustained cells that generate the same *directions* of local motion signals that are generated without a reversal of contrast in the large ISI case.

Further argument is needed to guarantee that a group motion signal is generated at Level 5. The key point is that Frame 1 and Frame 2 activate sustained cells that are sensitive to opposite directions-of-contrast but the same direction-of-motion. The long-range Gaussian filter pools inputs from Level 4 that are sensitive to opposite directions-of-contrast in order to create output signals from Level 5 that are insensitive to direction-of-contrast. Thus the local right-motion signals activated by Frame 1 and Frame 2 input to the same right-motion detection filter at Level 5, thereby generating a group motion signal.

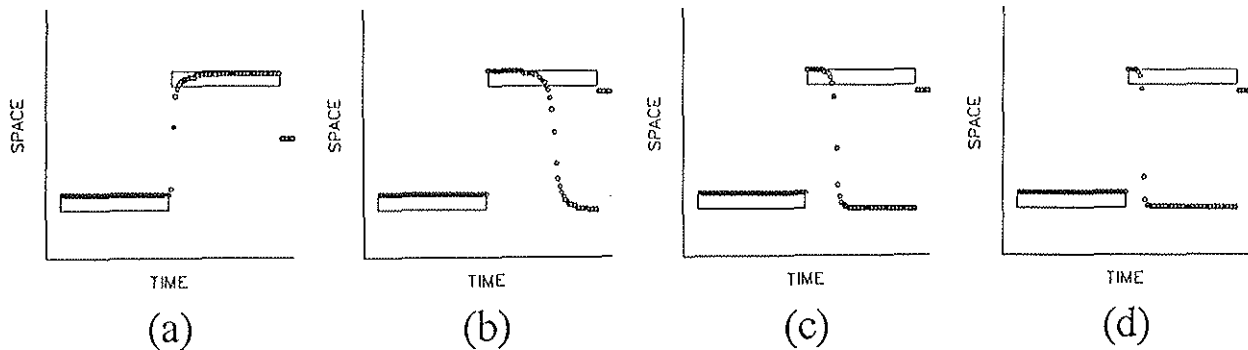
In summary, our explanation of reverse-contrast group motion at zero ISI uses all properties of the MOC Filter: sustained cell rectification and time-averaging, transient cell activation and rectification, sustained-transient gating, combining all sustained-transient cells sensitive to the same direction-of-motion but opposite directions-of-contrast via the long-range filter, and competitive sharpening of the motion output signal.

## 22. Delta Motion: Motion from the Second Flash to the First Flash

Another experimental probe of MOC Filter dynamics is the phenomenon of delta, or reverse, motion. ‘Delta motion’ is a motion that appears to travel from the second flash towards the first flash. This percept tends to occur when the luminance or contrast of the second flash is large compared to that of the first flash (Korte, 1915; Kolers, 1972, p.17). Delta motion depends upon the fact that the transient and sustained activities in the model are based on shunting equations whose averaging rate speeds up when input intensity increases.

Because the responses of these front-end filters speed up with increasing input intensity, the peak response of the MOC Filter to a low intensity flash can lag behind the response to a subsequent high intensity flash. If the intensity of the first flash in a two-flash display is made small enough relative to that of the second flash, the response of the transient and sustained mechanisms to the first flash may become sufficiently delayed compared to the response to the second flash that the phase lag due to the difference in neural response rates will become greater than the opposing phase lag due to the ISI. When this happens, the peak in the neural activity generated by the first flash will follow the peak activity due to the second flash and a reversed motion wave will result.

A model simulation of delta motion is illustrated in Figure 39. When the intensities of the two flashes are equal (Figure 39a), a motion wave is generated which moves from the location of the first flash to that of the second, as in the earlier simulations. When the intensity of the second flash is larger than that of the first (Figure 39b), an initial wave motion from the first flash to the second still occurs; however, this wave now has a much higher velocity than the forward wave which occurs when the flashes are of equal intensity. In fact, this movement occurs so rapidly that no trace of it appears in the output illustrated because of our discrete time approximation to the continuous curve. The spatial width of the long-range Gaussian filter is sufficient to generate motion under these conditions, but it may be undetectable due to its high speed, as in the empirical phenomenon of partial motion percepts. This initial rapid movement away from the first flash is followed by a slower reverse motion wave directed back towards the first flash. The velocity of this reverse motion wave



**Figure 39.** Simulation of delta motion. (a) When the intensities of the two flashes are equal, a forward motion signal is produced by the MOC Filter. (b)–(d) When the intensity of the second flash is sufficiently greater than that of the first, the direction of the motion signal is reversed. The velocity of the reverse motion wave increases with the intensity of the second flash. Input matrix  $128 \times 128$ .  $J_{ik} = \sum_j I_j F_{ji} = I_i$  at the appropriate stimulus edges, and  $= 0$  otherwise. Within Flash 1:  $I_i = 10$ . Within Flash 2 [(going across panels (a)–(d)]  $I_i = 10, 200, 500, 1000$ . Flash widths  $= 9$ ; durations  $= 57$ . Flash 1: left edge 24, on at time 8. Flash 2: left edge 88, on at time 65. Parameters  $A, B, C, D, E = .001$ .

increases monotonically with the intensity of the second flash, as illustrated in Figure 39b-d.

The phenomenon of delta motion illustrates the importance of the nonlinear temporal filtering properties of the transient and sustained mechanisms in the model, which result from their shunting inhibitory dynamics. Nonlinear responses will be generated by these filters when nonzero values are assigned to the parameters  $B$  and  $E$  in (4) and (5). In our previous simulations, these parameters were both set to zero. This corresponds to the special case in which the Level 2 and 3 filters behave like linear RC circuits. The full model with nonlinear filtering mechanisms continues to produce appropriate forward motion in response to displays in which the two flashes are of equal intensity, in addition to producing reverse motion under the appropriate conditions. A modified MOC Filter with linear front end mechanisms could produce forward motion, but not delta motion.

### 23. Apparent Motion Thresholds: The Joint Effects of Spatial Separation, Flash Duration, and Interstimulus Interval

In his classic study of apparent motion thresholds for two-flash displays, Neuhaus (1930) examined the individual and joint effects of manipulating the spatial separation and the flash durations on the range of ISIs over which a good motion percept was obtained. His data were briefly discussed in Sections 4 and 14, and graphed in Figure 3. The main effect of the

spatial separation variable has already been discussed. Here we briefly summarize the facts that are relevant to the present discussion.

At any fixed spatial separation, there is a restricted range of ISIs for which motion will be observed. This range progressively narrows with increasing flash separation. At ISIs which are briefer than the minimum ISI for motion, ‘simultaneity’ or ‘blinking in place’ is observed. At ISIs which are longer than the maximum ISI for motion, successive flashes are observed without an accompanying experience of continuous motion.

The lower motion threshold has received a great deal of attention in the literature. The fact that the minimum ISI for perceived motion increases with increasing spatial separation (Figure 6a) is sometimes referred to as Korte’s “Third Law” of apparent motion, after the student of Kurt Koffka’s who first drew attention to the phenomenon (Korte, 1915; Boring, 1950; Kolers, 1972). A similar dependence of the minimum SOA for perceived motion on increasing spatial separation also obtains (Figure 6b). Interestingly, whereas the minimum ISI decreases with flash duration (Figure 6a), the minimum SOA increases with flash duration (Figure 6b).

We show here that these properties follow in the model from the assumption that, *at threshold*, the signal generated by Flash 2 but evaluated at the location of Flash 1 is a fixed fraction of the signal generated by Flash 1 at its own location. In other words, at threshold, the ability of the second flash to generate a motion signal, starting at the location of the first flash, depends upon the size of the second flash’s signal relative to the first flash’s activation at the location of the first flash. This is a type of Weber Law for the motion threshold. Weber Law sensitivity is a consequence of designing Level 5 and/or subsequent levels using shunting on-center off-surround networks (Grossberg, 1983, 1987c). Assuming a right motion signal for concreteness, this threshold condition can be expressed mathematically by the following equation:

$$\frac{r_w(t)e^{-W^2/2K^2}}{r_0(t)} = \varepsilon, \quad (17)$$

where 0 and  $W$  are the locations of the first and second flashes, respectively; and  $r_0(t)$  and  $r_w(t)e^{-W^2/2K^2}$  are the magnitudes over time of the right motion signals which are generated by the first and second flashes at position 0. Parameter  $\varepsilon$  is the threshold Weber ratio for motion.

Mathematical expressions for the threshold ISI and SOA based on (17) are derived in the Appendix. The expressions are

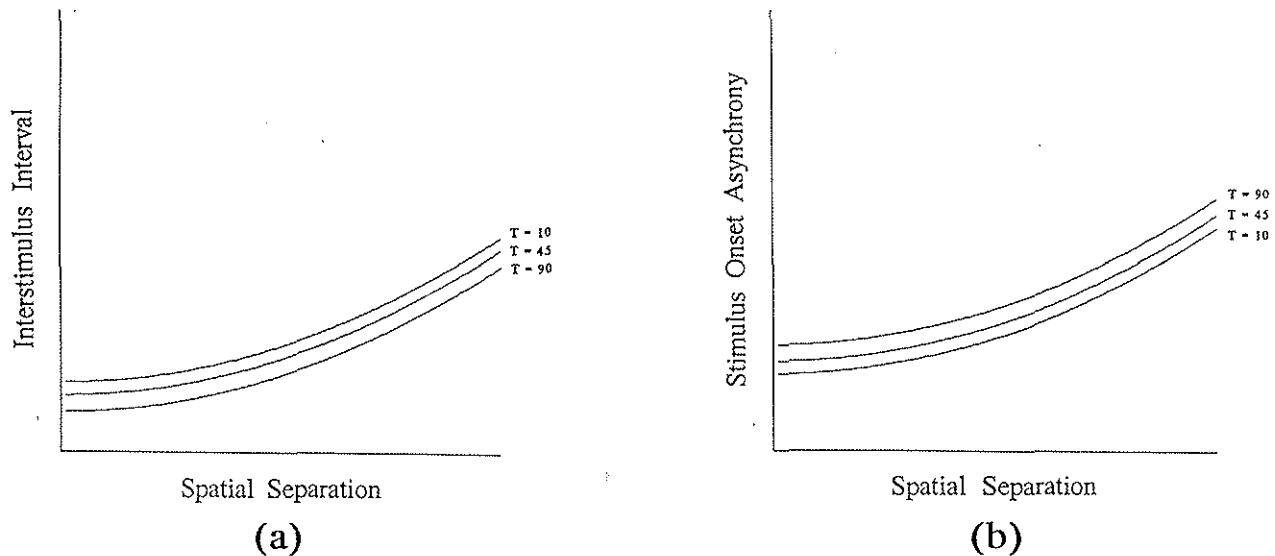
$$ISI = \frac{1}{2A} \left[ \ln(\varepsilon) - AT + \ln(1 - e^{-AT}) + \frac{W^2}{2K^2} \right], \quad (18)$$

and

$$SOA = \frac{1}{2A} \left[ \ln(\varepsilon) + AT + \ln(1 - e^{-AT}) + \frac{W^2}{2K^2} \right], \quad (19)$$

where  $T$  is the flash duration and  $W$  is the spatial separation. The behavior of ISI and SOA as  $T$  and  $W$  are varied is illustrated in Figure 40.

These theoretical lower threshold ISIs and SOAs for motion depend on both spatial separation and flash duration in a manner which is consistent with the data plotted in Figure



**Figure 40.** Theoretical lower apparent motion thresholds based on equation (17) in the text. (a) Minimum ISI for motion as a function of flash separation for flash durations of 10, 45 and 90 msec. Curves are solutions to equation (18) in the text, where  $A = 1 \text{ msec}^{-1}$ ;  $\epsilon = .1$ ;  $K = .354$ . (b) Minimum SOA for motion as a function of flash separation for the same flash durations and parameters as in (a).

6. Both temporal thresholds increase with flash separation. At all spatial separations, ISI is a decreasing function of flash duration, while SOA is an increasing function of flash duration within a broad parameter range.

As demonstrated in the Appendix, the correct theoretical dependence of the SOA on flash duration follows from the gating of the sustained cells by the transient cells. If the transient activities are held constant at the value 1 throughout time, the resulting expression for the threshold SOA does not depend on the flash duration. Thus, the fact that (18) and (19) behave appropriately as the flash duration is manipulated provides further support for the idea that motion signals are derived from gated sustained cell and transient cell activities.

Our proposed explanation of the effect of flash duration on the upper and lower apparent motion thresholds is more involved than this, however. Figure 3a shows that a decrease in flash duration increases the upper threshold ISI at which an apparent motion percept is replaced by a percept of temporally discrete stimuli. This property is paradoxical because it suggests that a decrease in flash duration, at a fixed ISI, makes it easier to generate an off-response to the first flash that overlaps the on-response to the second flash. In other words, a *shorter* flash duration implies a *longer* duration of network activation. Moreover, this longer duration, or "visual inertia" (Figures 17 and 24), persists after the flash terminates. Thus, the network can "remember" the duration of a previous flash as an activation whose persistence varies inversely with flash duration.

The form of the desired family of activation curves is suggested by the data reported

by Kolers (1964) for the probability of seeing motion as a function of flash duration and ISI (Figure 12). Note that the effects of shorter duration flashes increase more slowly *and* decrease more slowly in Figure 12, thereby persisting longer, within the range of durations studied in Figure 6. These properties cannot be explained using only the MOC Filter equations defined in Section 11. In particular, after a flash shuts off, the rate of decay of  $x_{ik}$  in equation (4) equals  $A$  and of  $x_i$  in (5) equals  $C$ . Both decay rates are independent of the duration of the previous flash. We are therefore led to ask: Is there a principled extension of the model that can explain these paradoxical data properties?

## 24. Shunting Cascades and Habituating Transmitter Gates

Mechanisms capable of modeling these data are, in fact, already part of the total Motion BCS model of which the MOC Filter equations in Section 11 form a part. These mechanisms are the following ones:

### (1) Shunting Cascade

In equations (4) and (5), sustained cell and transient cell outputs are caused by a single stage of shunting activation. We suggest that a cascade of two or more successive stages of shunting activation give rise to the sustained cell and transient cell output signals. For simplicity, we assume here that exactly two stages of shunting activation occur.

For example, as summarized in Section 11, Level 1 of the complete MOC Filter includes a shunting on-center off-surround network that discounts the illuminant (Grossberg and Todorović, 1988). This shunting stage was not included in the previous simulations, since it was not rate-limiting in explaining their targeted data. It could play the role of the other shunting stage that we need. Shunting cascades have also been used to model the earliest stages of photoreceptor transduction (Carpenter and Grossberg, 1981; reprinted in Grossberg, 1987c). For present purposes, all we need is one extra stage, wherever it might occur.

A shunting cascade provides a persistent short term memory of flash duration. Equations (20) and (21) clarify how this happens:

$$\frac{d}{dt}u = -au + (b - u)J, \quad (20)$$

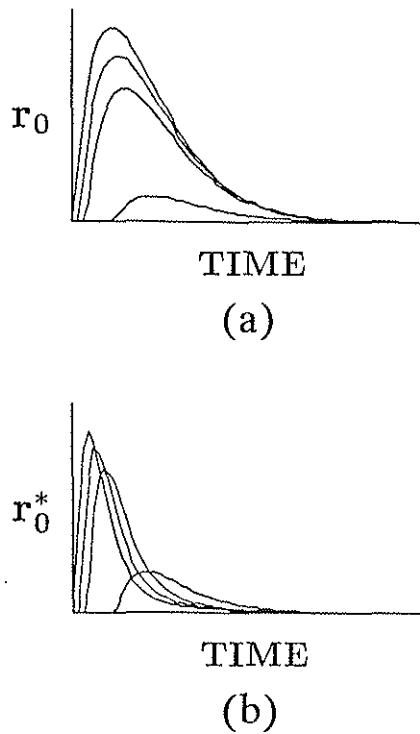
$$\frac{d}{dt}v = -cv + (d - v)u. \quad (21)$$

In (20), input  $J$  activates potential  $u$ . In (21), the output  $u$  from the first stage activates potential  $v$  of the second stage. A briefer input  $J$  in (20) causes a smaller activation  $u$ . A smaller activation  $u$  causes  $v$  to grow more slowly. This can be seen by rewriting (21) as

$$\frac{d}{dt}v = -(c + u)v + du. \quad (22)$$

In (22),  $v$  grows at a rate  $(c + u)$  that varies with the size of  $u$ . Thus, at the moment when  $J$  shuts off, both  $u$  and  $v$  are smaller if the duration of  $J$  is chosen briefer.

After  $J$  shuts off, activity  $u$  decays at a constant rate  $\alpha$ , as in equation (4). Throughout the decay period, a variable  $u$  that started out smaller remains smaller, because it decays at



**Figure 41.** Theoretical motion strength functions generated by flash offset as a function of time after offset. Functions corresponding to four different flash durations are illustrated. Strength functions with shallower peaks represent motion signals produced by briefer flashes. (a) Motion strength functions generated by a MOC Filter with a two-stage shunting cascade and no transmitter habituation. (b) Motion strength functions generated by the same MOC Filter with transmitter habituation. The curves in (b) exhibit many of the properties of the empirical motion strength functions graphed in Fig. 9. See text for details. Flash durations: 4, 16, 22, 36 (roughly in same proportion as those which generated the empirical functions in Fig. 9). Parameters  $A, C, D = .1$ ;  $B, E = .01$ ;  $\Sigma_j I_{ji} F_{ji} = J_{iL}$  or  $J_{iR}$ , whichever is nonzero, and  $= 0$  otherwise; in (b)  $T = 1$ ;  $U = 2$ ;  $V = 200$ .

a constant rate. On the other hand, by (22), variable  $v$  decays slower if  $u$  is smaller. Thus a *briefer* flash causes a *smaller* activation that decays *slower*, as in Figure 12.

Despite this useful qualitative property, the exact form of the family of curves generated as a function of flash duration does not conform to the data in Figure 12. A simulated family of  $v$  curves as a function of input duration is shown in Figure 41a, where an activation generated by a longer flash remains larger than one generated by a shorter flash throughout its period of growth and decay. The curves due to longer flashes do not *cross-over* the curves due to shorter flashes, as they do in the data curves of Figure 12. Somehow, a larger activation needs to cause a faster *relative* rate of decay than is provided by a shunting cascade. This extra degree of freedom needs, moreover, to operate at a stage subsequent to that of the shunting cascade.

## (2) Habituating Transmitter Gate

A mechanism that is formally competent to generate the cross-over property also occurs within the full Motion BCS model. It is a process whereby a neurotransmitter is released at a processing stage subsequent to a shunting cascade. In response to a signal  $y$ , the transmitter process  $z$  is released at a rate proportional to the product  $yz$ ; thus, transmitter

is released by mass action. The release of transmitter inactivates, or habituates, the amount of available transmitter at a rate  $Vyz$  that is proportional to amount released. Transmitter also accumulates to a target level  $u$  at a constant rate  $T$ . In all, the habituating transmitter gate obeys an equation of the form

$$\frac{dz}{dt} = T(u - z) - Vyz \quad (23)$$

(Grossberg, 1969, 1972b, 1982, 1987a). Equation (23) can be rewritten in the form

$$\frac{dz}{dt} = -(T + Vy)z + Tu. \quad (24)$$

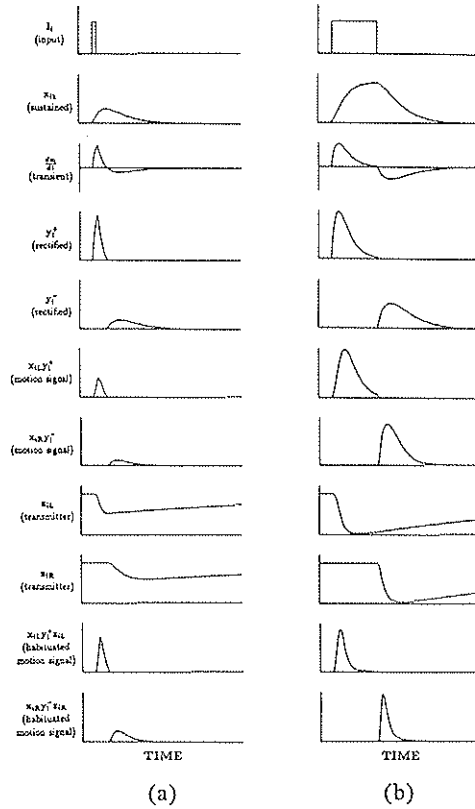
Thus, if signal  $y$  increases, the rate  $(T + Vy)$  of transmitter habituation increases. In addition, the rate  $yz$  of transmitter release first quickly increases with  $y$ , and then decreases more slowly with  $z$ , but at a rate that increases with  $y$ . As a function of input duration, the gated signal  $yz$  generates the family of curves in Figure 41b, which emulates the data in Figure 12.

Habituating transmitter gates have previously been hypothesized to exist in the Motion BCS at a stage subsequent to the sustained-transient cells (Grossberg, 1990c). These gates occur within a network of gated dipole opponent processes. The opponent processes help to reset resonating segmentations in response to moving images. The reset event consists of an antagonistic rebound that prevents massive smearing of motion percepts. The existence and size of a rebound depends upon the relative balance of transmitter habituation within the on-channel and off-channel of a gated dipole (Grossberg, 1972b, 1987a). Details concerning the modeling and formal properties of shunting cascades followed by habituating transmitter gates are now provided.

We assume that the sustained and transient cell activities are each based on a sequence of exactly two filtering operations, each of which are of the type described by equations (4) and (5). That is, the sustained cell activity  $x_{ik}$  in (4) is now assumed to be input to a second shunting stage that is characterized by a membrane equation of the same form, but in general with different parameters. We identify the output of this second stage with the sustained cell activity in the extended model. The transient activity is also assumed to be the time-derivative of a similar two-stage shunting model. As in the original model, the positive rectified and thresholded transient activity forms the on-cell response in the modified MOC Filter, while the negative rectified and thresholded transient activity forms the off-cell response.

Figure 42 illustrates the simulated activities at various levels of the network which contribute to the generation of a right motion signal in the extended model. As noted above, the neural responses produced by the extended model exhibit a flash duration dependence that is not observed in the original model. Figure 42a displays simulated MOC Filter responses to a brief flash, and Figure 42b displays responses of a MOC Filter with identical parameters to a longer flash.

In response to a brief flash (Figure 42a), the two-stage shunting cascade smooths its input  $I_i$  in such a way that the maximum sustained cell activity  $x_{ik}$  may peak after the offset of the stimulus. This phenomenon is empirically observed in the records of cortical responses to brief flashes (Duysens, Orban, Cremieux, and Maes, 1985). In response to a



**Figure 42.** Activities generated at various levels of a MOC Filter with a two-stage shunting cascade in response to (a) a brief flash, and (b) a long duration flash. The equations used to generate these simulations and the role of each of these activities in producing a motion wave are described in Section 23 of the text. Parameters  $A, C, D = .1$ ;  $B, E = .01$ ;  $\sum_j I_{ji} P_{ji} = J_{iL}$  or  $J_{iR}$ , whichever is nonzero, and  $= 0$  otherwise;  $T = 1$ ;  $U = 2$ ;  $V = 200$ ;  $\Pi = \Phi = .001$ . Time axis is 128 steps. Stimulus in (a) on at time 11 through time 14; in (b) on at time 11 through time 46.

long flash, on the other hand (Figure 42b), the peak sustained cell response occurs at, or near, the time of the flash offset.

The on-cell and off-cell responses are modeled for simplicity as the time derivative of a process with identical parameters as those of the sustained response, although in general the parameters of the processes which generate the sustained and transient activities are different. For a brief flash, the rectified on-cell activity  $y_i^+$  is terminated by the offset of the flash. For a longer flash, it decays away while the flash is still on. When the sustained activity  $x_{ik}$  peaks after the flash offset, as in Figure 42a, the transient activity  $dx_i/dt$  does not go negative until after some delay following the flash offset. This fact is reflected in the trace of the off-cell activity  $y_i^-$ , which does not begin immediately after the offset of the stimulus, as it did in the one-stage MOC Filter, but rather only after some time lag. When the off-cell activity does set in, it does not instantaneously reach its peak value, as in Figure 36. Instead it gradually rises, then decays away. When the flash duration is long, there is no time lag following the flash offset before  $y_i^-$  begins to build. Some time is, however, required for the off-cell activity to reach its peak value, unlike the case of the one-stage MOC Filter response.

The effects of flash duration on the phase lag and smoothing properties of the sustained and transient cell activities are inherited by the local motion signals which are based on them. Consider the two right motion components,  $x_{iL}y_i^+$  and  $x_{iR}y_i^-$ , which are generated

at edges of opposite direction-of-contrast in response to flash onset and offset, respectively. Note that the motion signal component which is produced at flash offset appears only after some delay in the case of a brief flash (Figure 42a). In addition, the peak of this motion component occurs at a later time than that which is generated by a longer flash (Figure 42b). With an appropriate choice of parameters for the two shunting stages, the profile of off-cell activity and its corresponding local motion signal can both be made, when the flashes are long, to approximate the exponential decay profile that is characteristic of the one-stage MOC Filter.

We assume that the transmitter gating stage acts subsequent to the computation of the local left motion signals  $l_i$  and the right motion signals  $r_i$ . Let  $z_{il}$  be the transmitter gate of  $l_i$  and  $z_{ir}$  be the transmitter gate of  $r_i$ . Then, as in equation (23),

$$\frac{dz_{il}}{dt} = T(U - z_{il}) - V l_i z_{il}, \quad (25)$$

and

$$\frac{dz_{ir}}{dt} = T(U - z_{ir}) - V r_i z_{ir}. \quad (26)$$

The habituated motion signals are modeled as the thresholded product of the unhabituated signals and the amounts of transmitter available for transmitting each of these signals, as in

$$l_i^* = l_i z_{il} - \Phi, \quad (27)$$

and

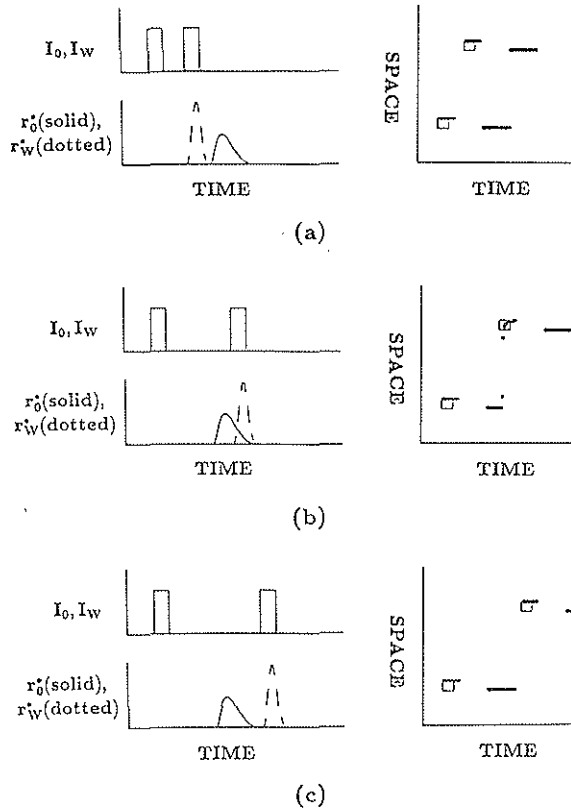
$$r_i^* = r_i z_{ir} - \Pi, \quad (28)$$

in which  $\Phi$  and  $\Pi$  are the left and right local motion thresholds.

Consider a right-motion signal  $r_i$ . By (8),  $r_i$  is the sum of components  $x_{iLY_i^+}$  and  $x_{iRY_i^-}$ . At most one of these components can be positive at each position  $i$  at any time. Thus each  $z_{iR}$  in (26) responds either to an  $L$ -type sustained cell  $x_{iL}$ , or a  $R$ -type sustained cell  $x_{iR}$ , but not both. This fact is summarized by writing  $z_{ir}$  as  $z_{iL}$  when it multiplies  $x_{iLY_i^+}$  and as  $z_{iR}$  when it multiplies  $x_{iRY_i^-}$ . (A more complete notation would be  $z_{irL}$  and  $z_{irR}$ .) The functions  $x_{iLY_i^+} z_{iL}$  and  $x_{iRY_i^-} z_{iR}$  are the habituated motion signals due to flash onset and offset, respectively. When these two signal profiles are generated at different positions, they may combine via the long-range Gaussian spatial filter, and their weighted sum is contrast-enhanced to form the MOC Filter output.

Temporal overlap of the motion signals generated by the Frame 1 offset and Frame 2 onset is needed to produce apparent motion from a two-flash display. When the motion thresholds  $\Phi$  and  $\Pi$  in (27) and (28) are set to positive values, a restricted range of ISIs exists over which continuous motion is generated. This property of the model is illustrated in Figure 43.

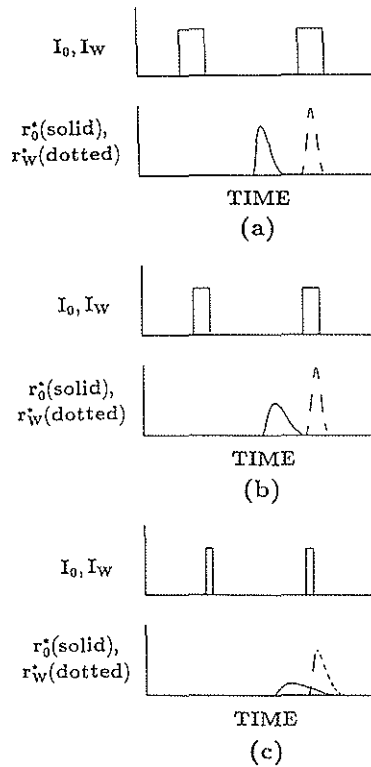
In Figure 43a, the ISI is too short to produce an overlap of the Frame 1 (solid curve) and Frame 2 (dashed curve) motion signal contributions. The signal produced by the onset of Frame 2 decays to a subthreshold level before the signal due to the offset of the first flash rises to a superthreshold level. The corresponding space-time diagram of the MOC Filter output indicates blinking in place, with no continuous motion signal. In Figure 43c,



**Figure 43.** Effect of ISI on the overlap in the motion signals from Frame 1 and Frame 2 flashes required for a continuous motion signal. In (a) the ISI is too short for an overlap to occur. The motion signal  $r_W^*$  induced by the onset of the Frame 2 flash at position  $W$  precedes the motion signal  $r_0^*$  induced by the offset of the Frame 1 flash at position 0. No continuous motion path is observed in the corresponding space-time diagram of the MOC Filter output. In (b) the ISI is appropriate for motion to occur;  $r_W^*$  and  $r_0^*$  overlap and a moving wave is observed in the MOC Filter output. In (c) the ISI is too large for overlap to occur.  $r_0^*$  precedes  $r_W^*$  and no moving wave is observed. Parameters  $A, C, D = .03$ ;  $B, E = .002$ ;  $\Sigma_j I_{ji}; F_{ji} = J_{iL}$  or  $J_{iR}$ , whichever is nonzero, and  $= 0$  otherwise;  $T = 1$ ;  $U = 2$ ,  $V = 5000$ ;  $\Pi = \Phi = .001$ . Input matrix  $128 \times 128$ . Frame 1 flash edge locations: 29, 37; Frame 2 flash edge locations: 92, 100.  $I_0$  on from time 16 through time 25. In (a)  $I_W$  on from time 38 through time 47. In (b)  $I_W$  on from time 64 through time 73. In (c)  $I_W$  on from time 80 through time 89.

on the other hand, the ISI is too long to produce the overlap required for motion; the signal due to the Frame 1 offset dies away before the contribution of the second flash gets started. Again, as shown in the associated space-time diagram, no continuous motion signal is produced. At an intermediate ISI (Figure 43b), the timing of the flashes is right for producing continuous motion. In this case, there is an appropriate overlap between the two motion signal contributions, with the component due to the Frame 1 offset slightly preceding the component due to the Frame 2 onset. The associated space-time diagram of the MOC Filter output exhibits a continuous apparent motion signal.

With this background, it is now possible to simulate the main effect seen in the upper threshold curves of the Neuhaus data in Figure 6a; namely, that the range of ISIs over which motion is seen increases as flash duration decreases. In Figure 44, the ISI of the display is held constant while the frame duration is varied. As the frame duration is decreased in panels (a)–(c), the overlap of the motion signal contributions from the two flashes increases. This is primarily because of an increase in the length of time that the motion signal due to offset of the Frame 1 flash is above threshold, although the motion signal due to the



**Figure 44.** The effect of manipulating flash duration of the range of ISIs over which a continuous motion signal can be produced by a MOC Filter with shunting cascade. In (a)–(c) the durations of the flash inputs at positions 0 and  $W$  are equal and simultaneously increased, while the ISI remains fixed. Continuous motion can occur in the model only when the motion signal  $r_0^*$  due to the offset of the flash at 0 overlaps the motion signal  $r_W^*$  due to the onset of the later flash at  $W$ . Decreasing the flash duration increases the range of ISIs over which such an overlap can occur. In the simulation depicted, the appropriate overlap is seen only when the flash duration is sufficiently short, in (c). Parameters  $A, C, D = .03$ ;  $B, E = .002$ ;  $\Sigma_j I_j; F_j; i = J_{iL}$  or  $J_{iR}$ , whichever is nonzero, and  $= 0$  otherwise;  $T = 1$ ;  $U = 2$ ;  $V = 5000$ ;  $\Pi = \Phi = .001$ . Time axis is 128 steps. In (a)  $I_0$  on from time 23 through time 37;  $I_W$  on from time 88 through time 102. In (b)  $I_0$  on from time 28 through time 37;  $I_W$  on from time 88 through time 97. In (c)  $I_0$  on from time 33 through time 37;  $I_W$  on from time 88 through time 92.

Frame 2 onset is also broadened to a lesser degree. The degree of overlap in these profiles determines the corresponding probabilities of generating motion signals near threshold over time; i.e., the motion strength functions. The joint effects of ISI and flash separation can be accounted for by combining these temporal characteristics with the assumption of a Weber Law for motion detection (Section 23). Alternatively, the desired result follows by assuming that the thresholding of the motion signal occurs after the long-range Gaussian filter. Such a deterministic threshold may be replaced by a statistical threshold in the presence of homogeneous neural noise. Any of these variations on the MOC Filter will produce a pattern of results which is consistent with that observed empirically in the motion threshold data of Neuhaus (Figure 6a).

## 25. Effects on the Motion Strength Function of Flash Duration, ISI, and Figural Identity

These simulation results help to clarify the dependence of motion strength on the ISI and flash durations of the display, as summarized in Figures 12 and 13. If the frame duration is short, the motion strength is relatively weak when the  $ISI = 0$ , rises to a single maximum as

the ISI is increased to an optimal value, and monotonically decreases with further increase in the ISI. As the flash duration is parametrically increased, the motion strength function attains larger peak values at smaller ISIs and decays at faster rates as a function of ISI, thereby generating the cross-over effect that was discussed in Section 24. In the limit of long durations, the motion strength function is maximal at a very small ISI, and thereafter monotonically decreases with ISI (Kolers, 1964; Bressan and Rudd, in preparation).

As shown in Figure 41b, formally similar properties can be reproduced by the modified MOC Filter with a two-stage shunting cascade in the sustained and transient filters, followed by habituating transmitter gating. In order to understand how the simulation in Figure 41b is related to the motion strength functions in Figures 12 and 13, first consider the temporal profiles of the network responses in the original one-stage model. In that model, the component of the motion signal which is generated by the offset of Frame 1 is an exponentially decaying function of the time after offset (as in Figure 36, line 7), and the component of the motion signal generated by the onset of Frame 2 is a unimodal function (Figure 36, line 6).

By appropriate manipulation of the sustained and transient filter parameters in equations (4) and (5) [i.e., making the shunting parameters  $B$  and  $E$  large], the rise times of the filters can be made arbitrarily short in comparison to their decay times. When the rise times of the filters are much faster than the decay times, the motion signal contribution due to the onset of Frame 2 converges on a delta function, while the contribution due to the offset of Frame 1 retains the shape of a gradually decaying exponential function. Since the probability of a motion signal depends on the temporal overlap of these two functions (weighted by a spatial proximity factor), in this extreme limit the brief on-signal will simply act as a probe of the off-signal profile; thus, the motion strength function will approximately be a monotonically-decaying function of the ISI, as is found experimentally with long flash durations.

According to this analysis, the shape of the temporal motion strength functions in Figures 12 and 13 should depend primarily on the duration of the first frame. This is consistent with the fact that the change in the form of the motion strength function which is observed when only the duration of the first frame is manipulated (Figure 13) is similar to that which is observed when the durations of both frames are simultaneously manipulated (Figure 12). To explain the dependence on flash duration, we need to analyse the two-stage MOC Filter.

To start, we consider the two-stage shunting cascade, but ignore the effects of the transmitter gate on motion signal habituation. In the limiting case of a fast second-stage shunt, the behavior of the two-stage model is the same as the one-stage model just discussed, because the second filter then just convolves the first stage output with a delta function. In general, however, the rise and decay constants of the second filter vary over time according to the prevailing value of the output of the first filter, which in turn depends on the parameters of the stimulus, such as its intensity and duration.

In particular, the time constants of both filter stages can be chosen to be so sensitive to the input to the first filter that the envelope of the motion signal produced by the onset of a flash of sufficiently large intensity is well-approximated by a delta function. Under such a choice of parameters, the output of the second filter can be made to decay more quickly when the stimulus duration is long than when it is brief. This result should be intuitively clear if one keeps in mind the fact that the time-averaging of the stimulus which is performed by the first filter (see equation (20)) ensures that a brief flash will have the same effect on the time constants of the second filter (see equation (21)) as a longer but less intense flash. Thus flash

duration and intensity have equivalent effects on the second filter due to the time-averaging action of the first filter.

Since the motion signal due to flash onset is assumed to be well-approximated by a delta function, we need only consider the motion signal generated by the offset of the first flash of a two-flash display in order to see whether the model can generate motion strength functions which exhibit the correct flash duration dependence. We therefore assume that the effect of a threshold measurement carried out at any particular ISI is to probe the strength of the motion signal generated by the offset of the first flash. In Figure 41a, the effects of flash duration on the unhabituated motion signal generated at flash offset are illustrated. The shapes of the simulated functions in Figure 41a have some, but not all, of the properties of empirical motion strength functions (Figures 12 and 13). In general, the peak motion strength generated by a long flash is appropriately greater than that generated by a short flash. Also, the curves peak at longer time lags following flash offset when the flash duration is shorter. This property is consistent with the fact that peak motion strength occurs empirically at longer ISIs when the flash duration is shorter (Figure 12).

On the other hand, the theoretical motion strength functions plotted in Figure 38a fail to account for the crossover of the curves in Figure 12. That is, the theoretical motion strengths generated by flashes of longer duration are greater than those which are generated by briefer flashes at all ISIs; whereas the empirical strength functions corresponding to long flashes decay away more quickly than those which correspond to short flashes. It is for this reason that the upper thresholds for apparent motion measured by Neuhaus (1930) increase with decreasing flash duration in Figure 6a.

The discrepancy in these results is eliminated by the motion signal habituation due to transmitter gating of signals from the shunting cascade. The results of a simulation of the habituated motion signals generated by the offset of the Frame 1 flash are shown in Figure 41b. The flash duration dependence of these habituated motion signals closely mimics that of the data in Figure 12, including the tendency for the motion strength functions generated by long flashes to decay away at shorter ISIs than those generated by short flashes. Further parameter adjustments could improve this fit. However, we have already made a very rough approximation by assuming that the motion signal component due to the onset of the second flash acts as an instantaneous probe of the signal due to the Frame 1 offset. The argument presented here is intended only to provide a qualitative insight into the location and types of mechanisms that seem to govern data properties which have long resisted any explanation.

In Figure 11 are graphed the temporal motion strength functions generated by flashes of different figural shape. The functions generated by the different shapes are essentially identical, as also occurs in the model. This is a result of the fact that the MOC Filter computes motion on the basis of primitive locally filtered images, rather than on the basis of higher-order features or cognitive variables.

In particular, the local motion signals from different orientationally tuned cells at Level 2 of the MOC Filter are pooled by target cells at Level 5 (Figure 20). The long-range Gaussian filter averages across a band of orientations in order to generate cells that are sensitive to direction-of-motion, insensitive to direction-of-contrast, and less sensitive to orientation than individual Level 2 cells. These properties are consistent with experimental reports that, whereas apparent motion is not sensitive to a form matching process, it can be influenced by stimulus orientation (Shechter, Hochstein, and Hillman, 1988). It has also been reported

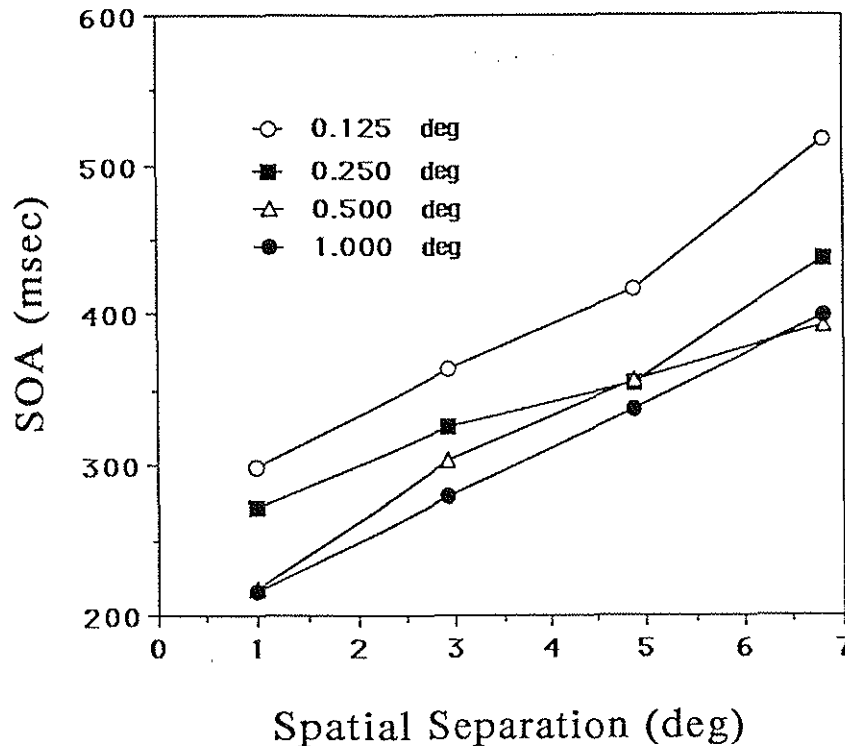


Figure 45. Lower threshold SOAs for two-flash apparent motion as a function of spatial separation (from an unpublished experiment by one of the authors). Flashes were bright squares on a dark background, generated on an Apple II computer. Separate curves represent data generated by flashes of four different sizes. Flash widths are indicated on the graph. Data from three subjects are combined.

that apparent motion occurs only between similar spatial frequencies (Watson, 1986). In our model, this is due to the fact that several copies of the MOC Filter exist, each fed by a different range of receptive field sizes (Figure 29), just as in the analogous theory of SOC Filter design for static form perception (Grossberg, 1987b; Grossberg and Marshall, 1989).

## 26. Ternus Display: The Effects of Flash Duration and Element Size

The same multiple-stage filtering assumption that explains the effect of flash duration on two-flash apparent motion thresholds can also explain the flash duration dependence of the threshold ISI at which the transition from element motion to group motion occurs in the Ternus paradigm, as discussed in Section 10 (Breitmeyer and Ritter, 1986a, 1986b). Although we do not simulate this effect here, it is now easy to see how these two phenomenon are related.

For any fixed flash separation of a two-flash Ternus display, the lower motion threshold will be determined by the persistence of the on-signal generated by the first flash; and the upper motion threshold will be limited by the off-signal generated by this flash. In general, the persistence of both of these signals will depend on flash duration. In particular, the on-signal persistence will be long when the flash duration is short (Figure 42a), which is consistent with the psychophysically measured flash duration dependence of visual iconic persistence (Bowen, Pola, and Matin, 1974; Bowling and Lovegrove, 1980).

The development of an off-cell response is progressively delayed as the duration of the flash is shortened. This results in a delayed onset of the motion signal component due to the

offset of the first flash (as in Figure 42a), and thus to a delayed motion wave. No motion will be observed unless the on-cell signal due to the second flash overlaps with the off-cell signal due to the first flash, which for brief flashes does not form until after some minimum delay. As discussed above, this results in a threshold minimum ISI for apparent motion.

In the context of the Ternus display (Figure 15a), when the flash duration is short, an element does not contribute to the motion wave until the end of a critical minimum period after the Frame 1 offset. This minimum period is the time that it takes for the motion signal generated by the off-transient activity to get started. If the element is turned on again at this same spatial location before this minimum delay is reached, the contribution that this element would otherwise have made to the motion wave will be killed off before it gets started. When this happens, however, any locations at which elements are not reactivated may still contribute to a motion signal, provided that there is sufficient spatial and temporal overlap between the contributions to the motion signal generated by the first and second frames. Thus, element motion can still occur. However, there is a minimum, but nonzero, ISI for the occurrence of group motion. The model hereby relates the flash duration dependence of two-flash motion thresholds with that of the element-to-group motion transitional ISI in the Ternus experiment. This relationship has not, to our knowledge, previously been pointed out in the motion literature.

A similar relationship between two-flash lower motion thresholds and element-to-group transitional ISIs is also found experimentally to depend on element size (Petersik and Pantle, 1979; Breitmeyer and Ritter, 1986a, 1986b). In Figure 45 are shown some data from a previously unpublished study of two-flash motion thresholds which was carried out by one of the present authors (Rudd, unpublished data). In this experiment, subjects viewed a continuously cycled two-flash display generated by an Apple II computer. The independent variables—element size and spatial separation—were varied factorially. The subject's task was to indicate the lower motion threshold motion by adjusting the SOA of the display continuously with a game control device until the frames alternated at the "fastest rate that produced a percept of good motion". The ISI of the display was zero throughout the experiment.

The combined mean threshold SOA values for three subjects is plotted for each of the conditions. Each point on the graph represents a total of between 74 and 82 measurements, depending on the condition. The results indicate that the effect of decreasing the element size is to raise the motion thresholds at all spatial separations by approximately equal amounts. The data indicate either no interaction between the size and separation variables, or else an interaction which is too small to be observed by eye.

One possible explanation of this result is that the responses of the sustained and/or transient filters to the smaller stimuli is slower and/or weaker than the response of these filters to the larger stimuli. This would have the effect in the MOC Filter of delaying the time at which the motion signal contributions from successive flash *onsets* would reach a threshold level (even in the one-stage model). This explanation is consistent with a large body of data in the vision literature which indicate that the visual system responds more sluggishly to stimuli of smaller scale, over a large range of scales (Ferree and Rand, 1929; Teichner and Krebs, 1972; Breitmeyer, 1975; Harwerth and Levi, 1978; Rudd, 1988, in press).

The same hypothesis also accounts for the fact that the ISI at which the transition from element to group motion occurs in the Ternus experiment increases with decreasing element

size. Smaller stimuli generate weaker and/or slower first-stage neural responses which result in a slower decay of the on-signal output of the shunting cascade, thus leading to a greater visual persistence, and a higher element-to-group motion transitional ISI, as reported by Breitmeyer and Ritter (1986a).

## 27. 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 MOC Filter; the breadth of the Gaussian filter from Level 4 to Level 5 of the MOC Filter (Figure 21); the breadth of the endstopping operation from complex cells to hypercomplex cells stage of the CC Loop (Figure 3); 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 MOC Filter. Subsequent spatial interactions within each copy are assumed to be related to receptive field size in a self-similar fashion (Grossberg, 1987b; Grossberg and Marshall, 1989). These relationships among spatial scales enable a variety of difficult motion properties to be explained.

In particular, 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 support the hypothesis that BCS interactions become independent of direction-of-contrast no later than Level 5 of Figure 21, thereby enabling flashes with different directions-of-contrast to interact across space, as in Figure 26, to generate a motion signal.

Such a motion signal within the BCS generates output signals to the FCS that define a boundary web (Section 1) which contains the filling-in of 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 interactions (Grossberg, 1987b), which clarifies how a binary switch between colors can occur at all, and filling-in of the winning opponent color within the moving boundary structure (Cohen and Grossberg, 1984; Grossberg and 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 MOC Filter circuit of in Figure 21 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.

Chubb and Sperling (1988) have also distinguished between short-range and long-range mechanisms that differ in their sensitivity to changes in image contrast. Their experiments suggest that the long-range mechanism behaves as if it performs a full-wave rectification of image data. In the MOC Filter, the long-range Gaussian filter combines pairs of half-wave rectified signals that are sensitive to opposite directions-of-contrast. This multiple-stage operation has the same net effect as full-wave rectification. Although the processing stages of the MOC Filter lead to the net effect of full-wave rectification, the analysis of Ternus motion and reverse-contrast Ternus motion in Sections 20 and 21 shows that all of these processing stages are needed to explain a wide range of motion data.

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 and Bell, 1976; Ramachandran and Anstis, 1983), with increasing field size (Baker and Braddick, 1982; Chang and Julesz, 1983), and with increasing element size (Petersik, Pufahl, and Krasnoff, 1983). These properties are clarified by observations that  $D_{max}$  varies with the spatial frequency content of the image (Burr, Ross, and Morrone, 1986; Nakayama and Silverman, 1984, 1985) and that receptive field size varies with spatial frequency (Anderson and 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 interactions within the copy are related to receptive field size by a property of self-similarity, as indicated in Figure 16.

## 28. Multiplexing of motion direction and motion depth

In the static BCS (Figure 3), it has been shown how cells become binocular at the complex cell level (Grossberg, 1987b; Grossberg and 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 21). This hypothesis 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.

As noted in Section 11, 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, Desimone, and Gross, 1984; Maunsell and van Essen, 1983). The organization of these cells into hypercolumns whose cells vary with respect to direction-of-motion, rather than orientation, has been described in Grossberg (1991) as a manifestation of FM Symmetry.

## 29. Concluding Remarks: Towards a Unified Theory of Biological Vision

The MOC Filter model suggests a unified mechanistic explanation of a large data base concerning short-range and long-range apparent motion, both here and in Grossberg and Rudd (1989c). When supplemented by a Motion CC Loop, or MOCC Loop, as in Grossberg and Rudd (1989c) and Grossberg and Mingolla (1990a, 1990b, 1990c), the Motion BCS model

suggests a solution of various motion segmentation problems, such as the global aperture problem, motion capture, and induced motion.

These seemingly paradoxical perceptual properties may now be explained as manifestations of ecologically useful mechanisms. Such mechanisms generate resonant emergent segmentations whose output signals are sensitive to direction-of-motion but insensitive to direction-of-contrast, and control rapid switching between the complementary perceptual states of resonance and reset. The coherence needed for globally unambiguous perception is achieved by these resonant segmentations, and their rapid reset prevents the resonances from causing massive perceptual smearing (Grossberg, 1991). In this broader theoretical context, various data concerning negative aftereffects, such as the MacKay illusion (1957), the waterfall illusion (Sekuler, 1975), and aftereffects of long-range apparent motion (von Grünau, 1986), may also be explained using mechanisms that were herein used to explain threshold properties such as Korte's Laws. Taken together, these results provide a foundation for building a principled neural theory of motion perception, or, more correctly, a neural theory of *motion form* perception.

This theory may itself be subsumed under a more general theory which reveals the Motion BCS and the Static BCS to be variations on a common architectural design. In this theory, the Static BCS and the Motion BCS are viewed as two parallel subsystems of a larger, symmetric system design called FM Symmetry (Grossberg, 1991). FM Symmetry suggests how many previously intractable facts about static form perception and motion form perception may now be given a unified explanation. In particular, the different geometries of static form perception and motion form perception are clarified. Moreover, interactions between the Static BCS and Motion BCS lead to predictions concerning how the cortical stream  $V1 \rightarrow V2 \rightarrow MT$  may contribute to the perception of moving-form-in-depth, and to an explanation of apparent motion of illusory contours (Ramachandran, Rao, and Vidyasagar, 1973; Ramachandran, 1985) that may be used as a perceptual probe of these neurobiological predictions.

The total theory which unifies the Static BCS and the Motion BCS under the organizing principle of FM Symmetry is called FACADE Theory, in order to connote that the theory's final representations multiplex together properties of Form-And-Color-And-DEpth. FACADE Theory offers a new foundation for a general theory of biological vision. Its suggested resolution of many classical paradoxes in visual perception points towards a wealth of new empirical, theoretical, technological, and even philosophical issues that have only begun to be explored.

## REFERENCES

- Albright, T.D. (1984). Direction and orientation selectivity of neurons in visual area MT of the Macaque. *Journal of Neurophysiology*, **52**, 1106–1130.
- Albright, T.D., Desimone, R., and Gross, C.G. (1984). Columnar organization of directionally sensitive cells in visual area MT of the macaque. *Journal of Neurophysiology*, **51**, 16–31.
- Anderson, S.J. and Burr, D.C. (1987). Receptive field size of human motion detection units. *Vision Research*, **27**, 621–635.
- Anstis, S.M. and Mather, G. (1985). Effects of luminance and contrast on direction of ambiguous apparent motion. *Perception*, **14**, 167–179.
- Baker, C.L., Jr. and Braddick, O.J. (1982). The basis of area and dot number effects in random dot motion perception. *Vision Research*, **22**, 1253–1259.
- Anstis, S.M. and Ramachandran, V.S. (1987). Visual inertia in apparent motion. *Vision Research*, **27**, 755–764.
- Bartley, S.H. (1936). The relation of retinal illumination to the experience of movement. *Journal of Experimental Psychology*, **19**, 475–485.
- Bartley, S.H. (1941). **Vision, a study of its basis**. New York: D. Van Nostrand.
- Boring, E.G. (1950). **A history of experimental psychology**. Englewood Cliffs, NJ: Prentice-Hall.
- Bowen, R.R., Pola, J., and Matin, L. (1974). Visual persistence effects of flash luminance, duration, and energy. *Vision Research*, **14**, 295–303.
- Bowling, A. and Lovegrove, W. (1980). The effects of stimulus duration on the persistence of gratings. *Perception and Psychophysics*, **27**, 574–578.
- 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. and Adlard, A. (1978). Apparent motion and the motion detector. In J.C. Armstrong, J. Krauskopf, and B.R. Wooten (Eds.), **Visual psychophysics and psychology**. New York: Academic Press.
- Breitmeyer, B.G. (1975). Simple reaction time as a measure of the temporal response properties of transient and sustained channels. *Vision Research*, **15**, 1411–1412.
- Breitmeyer, B.G. and Halpern, M. (1978). Visual persistence depends on spatial frequency and retinal locus. Paper presented at the annual meeting of the Psychonomic Society, San Antonio, TX, November, 1978.
- Breitmeyer, B.G. and Levi, D.M., and Harwerth, R.S. (1981). Flicker-masking in spatial vision. *Vision Research*, **21**, 1377–1385.
- Breitmeyer, B.G. and Ritter, A. (1986a). Visual persistence and the effect of eccentric viewing, element size, and frame duration on bistable stroboscopic motion percepts. *Perception and Psychophysics*, **39**, 275–280.
- Breitmeyer, B.G. and Ritter, A. (1986b). The role of visual pattern persistence in bistable stroboscopic motion. *Vision Research*, **26**, 1801–1806.

- Bressan, P. and Rudd, M.E. (in preparation). Effects of ISI and flash duration on direction judgments in an ambiguous motion paradigm.
- Burr, D.C., Ross, J. and Morrone, M.C. (1986). Smooth and sampled motion. *Vision Research*, **26**, 643–652.
- Burt, P. and Sperling, G. (1981). Time, distance, and feature trade-offs in visual apparent motion. *Psychological Review*, **88**, 171–195.
- Carpenter, G.A. and Grossberg, S. (1981). Adaptation and transmitter gating in vertebrate photoreceptors. *Journal of Theoretical Neurobiology*, **1**, 1–42.
- Chang, J.J. and Julesz, B. (1983). Displacement limits for spatial frequency filtered random-dot cinematograms in apparent motion. *Vision Research*, **23**, 1379–1385.
- Chubb, C. and Sperling, G. (1988). Two motion perception mechanisms revealed through distance-driven reversal of apparent motion. *Proceedings of the National Academy of Sciences*, **86**, 2985–2989.
- Cohen, M.A. and Grossberg, S. (1984). Neural dynamics of brightness perception: Features, boundaries, diffusion, and resonance. *Perception and Psychophysics*, **36**, 428–456.
- DeSilva, H.R. (1926). An experimental investigation of the determinants of apparent visual movement. *American Journal of Psychology*, **37**, 469–501.
- DeValois, R.L., Albrecht, D.G., and Thorell, L.G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, **22**, 545–559.
- Di Lollo, V. (1977). Temporal characteristics of iconic memory. *Nature*, **267**, 241–243.
- Di Lollo, V. and Hogben, J.H. (1985). Suppression of visible persistence. *Journal of Experimental Psychology: Human Perception and Performance*, **11**, 304–316.
- Duysens, J., Orban, G.A., Cremieux, J., and Maes, H. (1985). Visual cortical correlates of visible persistence. *Vision Research*, **25**, 171–178.
- Exner, S. (1875). Ueber das Sehen von Bewegungen und die Theorie des zusammengesetzten Auges. *Sitzungsberichte Akademie Wissenschaft Wien*, **72**, 156–190.
- Ferree, C.E. and Rand, G. (1929). Intensity of light and speed of vision: Effect of size of object and difference of coefficient of reflection as between object and background. *Journal of Experimental Psychology*, **12**, 363–391.
- Foster, K.H., Gaska, J.P., Nagler, M., and Pollen, D.A. (1985). spatial and temporal frequency selectivity of neurons in visual cortical areas V1 and V2 of the macaque monkey. *Journal of Physiology*, **365**, 331–363.
- Gengerelli, J.A. (1948). Apparent movement in relation to homonymous and heteronymous stimulation of the cerebral hemispheres. *Journal of Experimental Psychology*, **38**, 592–599.
- Giaschi, D. and Anstis, S. (1989). The less you see it, the faster it moves: Shortening the “on-time” speeds up apparent motion. *Vision Research*, **29**, 335–347.
- 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. (1984). Outline of a theory of brightness, color, and form perception. In E. Degreef and J. van Buggenhaut (Eds.), **Trends in Mathematical Psychology**. North Holland: Amsterdam.
- 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, I: Cognition, learning, reinforcement, and rhythm**. Amsterdam: Elsevier/North-Holland.
- Grossberg, S. (Ed.) (1987c). **The adaptive brain, II: Vision, speech, language, and motor control**. Amsterdam: Elsevier/North-Holland.
- Grossberg, S. (1990a). Self-organizing neural architectures for motion perception, adaptive sensory motor control, and associative mapping. In M. Caudill (Ed.), **Proceedings of the international joint conference on neural networks**, January, **II**, 213–216, Hillsdale, NJ: Erlbaum Associates.
- Grossberg, S. (1990b). Neural FACADES: Visual representations of static and moving form-and-color-and-depth. *Mind and Language*, **5**, 411–456.
- Grossberg, S. (1990c). 3-D vision and figure-ground separation by visual cortex. Submitted for publication.
- Grossberg, S. (1991). Why do parallel cortical systems exist for the perception of static form and moving form? *Perception and Psychophysics*, **49**, 117–141.
- Grossberg, S. and 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. and Mingolla, E. (1985a). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review*, **92**, 173–211.
- Grossberg, S. and Mingolla, E. (1985b). Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations. *Perception and Psychophysics*, **38**, 141–171.
- Grossberg, S. and Mingolla, E. (1987). Neural dynamics of surface perception: Boundary webs, illuminants, and shape-from-shading. *Computer Vision, Graphics, and Image Processing*, **37**, 116–165.
- Grossberg, S. and Mingolla, E. (1990a). Neural dynamics of motion segmentation: Direction fields, apertures, and resonant grouping. In M. Caudill (Ed.), **Proceedings of the international joint conference on neural networks**, **I**, 11–14. Hillsdale, NJ: Erlbaum Associates.
- Grossberg, S. and Mingolla, E. (1990b). Neural dynamics of motion segmentation. In **Proceedings of Vision Interface '90**. Halifax, Nova Scotia, May 14–18, 1990.
- Grossberg, S. and Mingolla, E. (1990c). Neural dynamics of motion Perception: Direction fields, apertures, and resonant grouping. Submitted for publication.
- Grossberg, S., Mingolla, E., and Todorović, D. (1989). A neural network architecture for preattentive vision. *IEEE Transactions on Biomedical Engineering*, **36**, 65–84.

- Grossberg, S. and Rudd, M.E. (1989a). Neural dynamics of visual motion perception: Group and Element apparent motion. *Investigative Ophthalmology Supplement*, **30**, 73.
- Grossberg, S. and Rudd, M.E. (1989b). A neural architecture for visual motion perception: Group and element apparent motion. In **Proceedings of the international joint conference on neural networks**, June 19, 1989, Washington, DC.
- Grossberg, S. and Rudd, M.E. (1989c). A neural architecture for visual motion perception: Group and element apparent motion. *Neural Networks*, **2**, 421–450.
- Grossberg, S. and Rudd, M.E. (1990). Cortical dynamics of visual motion perception: Short- and long-range motion. *Investigative Ophthalmology Supplement*, **31**, 529.
- Grossberg, S. and Todorović, D. (1988). Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena. *Perception and Psychophysics*, **43**, 241–277.
- Harwerth, R.S. and Levi, D.M. (1978). Reaction time as a measure of suprathreshold grating detection. *Vision Research*, **24**, 933–941.
- Heggelund, P. (1981). Receptive field organization of complex cells in cat striate cortex. *Experimental Brain Research*, **42**, 89–98, 90–107.
- Higginson, G.D. (1926). Apparent visual movement and the Gestalt. *Journal of Experimental Psychology*, **9**, 228–252.
- Hogben, J.H. and DiLollo, V. (1985). Suppression of visual persistence in apparent motion. *Perception and Psychophysics*, **38**, 450–460.
- Hubel, D.H. and 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. and Wiesel, T.N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, **195**, 215–243.
- Hubel, D.H. and Wiesel, T.N. (1977). Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London (B)*, **198**, 1–59.
- Kenkel, F. (1913). Untersuchungen über Zusammenhang zwischen Erscheinungsgrosse und Erscheinungsbewegung beim einigen sogenannten optischen Täuschungen. *Zeitschrift für Psychologie*, **61**, 358–449.
- Kolers, P.A. (1964). The illusion of movement. *Scientific American*, **211**, 98–106.
- Kolers, P.A. (1972). **Aspects of motion perception**. Oxford: Pergamon Press.
- Kolers, P.A. and Pomerantz, J.R. (1971). Figural change in apparent motion. *Journal of Experimental Psychology*, **87**, 99–108.
- Kolers, P.A. and von Grünau, M. (1975). Visual construction of color is digital. *Science*, **187**, 757–759.
- Lappin, J.S. and Bell, H.H. (1976). The detection of coherence in moving random-dot patterns. *Vision Research*, **16**, 161–168.
- Korte, A. (1915). Kinematoskopische Untersuchungen. *Zeitschrift für Psychologie*, **72**, 194–296.
- MacKay, D.M. (1957). Moving visual images produced by regular stationary patterns. *Nature*, **180**, 849–850.

- Marr, D. and 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. and van Essen, D.C. (1983). Response properties of single units in middle temporal visual area of the macaque. *Journal of Neurophysiology*, **49**, 1127–1147.
- Meyer, G.E. and Maguire, W.M. (1977). Spatial frequency and the mediation of short-term visual storage. *Science*, **198**, 534–525.
- Mezrich, J.J. (1984). The duration of visual persistence. *Vision Research*, **24**, 631–632.
- Nakayama, K. and Silverman, G.H. (1984). Temporal and spatial characteristics of the upper displacement limit for motion in random dots. *Vision Research*, **24**, 293–299.
- Nakayama, K. and Silverman, G.H. (1985). Detection and discrimination of sinusoidal grating displacements. *Journal of the Optical Society of America A*, **2**, 267–273.
- Navon, D. (1976). Irrelevance of figural identity for resolving ambiguities in apparent motion. *Journal of Experimental Psychology: Human Perception and Performance*, **2**, 130–138.
- Neuhaus, W. (1930). Experimentelle untersuchung der scheinbewegung. *Archiv für die gesamte Psychologie*, **75**, 315–458.
- Newsome, W.T., Gizzi, M.S., and Movshon, J.A. (1983). Spatial and temporal properties of neurons in macaque MT. *Investigative Ophthalmology and Visual Science*, **24**, 106.
- Orlansky, J. (1940). The effect of similarity and difference in form on apparent visual movement. *Archives of Psychology*, **246**.
- Pantle, A.J. and 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. and Picciano, L. (1976). A multistable movement display: Evidence for two separate motion systems in human vision. *Science*, **193**, 500–502.
- Peterhans, E. and von der Heydt, R. (1989). Mechanisms of contour perception in monkey visual cortex. II. Contours bridging gaps. *The Journal of Neuroscience*, **9**, 1749–1763.
- Petersik, J.T. and 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., and Krasnoff, E. (1983). Failure to find an absolute retinal limit of a putative short-range process in apparent motion. *Vision Research*, **23**, 1663–1670.
- Pollen, D.A., Gaska, J.P., and Jacobson, L.D. (1989). Physiological constraints on models of visual cortical function. In R.M.J. Cotterill (Ed.), **Models of brain function**, 115–135. Cambridge: Cambridge University Press.
- Prazdny, K. (1983). Illusory contours are not caused by simultaneous brightness contrast. *Perception and Psychophysics*, **34**, 403–404.
- Ramachandran, V.S. (1985). Apparent motion of subjective surfaces. *Perception*, **14**, 127–134.
- Ramachandran, V.S. and Anstis, S.M. (1983). Displacement thresholds for coherent apparent motion in random dot patterns. *Vision Research*, **23**, 1719–1724.
- Ramachandran, V.S., Rao, V.M., and Vidyasagar, T.R. (1973). Apparent motion with subjective contours. *Vision Research*, **13**, 1399–1401.
- Rudd, M.E. (1988). Quantal fluctuation limitations on reaction time to sinusoidal gratings. *Vision Research*, **28**, 179–186.

- Rudd, M.E. (in press). A variable integration time model of threshold vs intensity curves. *Journal of the Optical Society of America A*.
- Rudd, M.E. and Bressan, P. (1991). Invariance in apparent motion strength due to cancelling of separate 2-D size and proximity effects. *Investigative Ophthalmology Supplement*, **34**, 826.
- Rudd, M.E. and Bressan, P. (in press). Quantitative analyses of apparent motion thresholds: Interactions between spatial separation, size, and luminance. Submitted to *Vision Research*.
- Sekuler, R. (1975). Visual motion perception. In E.C. Carterette and M.P. Friedman (Eds.), **Handbook of perception, Volume V: Seeing**. New York: Academic Press.
- Shapley, R. and Gordon, J. (1985). Nonlinearity in the perception of form. *Perception and Psychophysics*, **37** (1), 84–88.
- Shechter, S., Hochstein, S., and Hillman, P. (1988). Shape similarity and distance disparity as apparent motion correspondence cues. *Vision Research*, **28**, 1013–1021.
- Spigel, I.M. (1968). Problems in the study of visually perceived movement: An introduction. In R.H. Haber (Ed.) **Contemporary theory and research in visual perception**. New York: Holt, Rinehart, and Winston, pp.103–121.
- Squires, P.C. (1931). The influence of hue on apparent visual movement. *American Journal of Psychology*, **43**, 49–64.
- Tanaka, M., Lee, B.B., and Creutzfeldt, O.D. (1983). Spectral tuning and contour representation in area 17 of the awake monkey. In J.D. Mollon and L.T. Sharpe (Eds.) **Colour Vision**. New York: Academic Press.
- Teichner, W.H. and Krebs, M.J. (1972). Laws of simple reaction time. *Psychological Review*, **79**, 344–358.
- Ternus, J. (1926/1950). Experimentelle Untersuchungen über phänomenale Identität. *Psychologische Forschung*, **7**, 81–136. Abstracted and translated in W.D. Ellis (Ed.), **A sourcebook of Gestalt psychology**. New York: Humanities Press, 1950.
- van der Waals, H.G. and Roelofs, C.O. (1930). Optische scheinbewegung. *Zeitschrift für Psychologie und Physiologie des Zinnesorgane*, **114**, 241–288.
- van der Waals, H.G. and Roelofs, C.O. (1931). Optische scheinbewegung. *Zeitschrift für Psychologie und Physiologie des Zinnesorgane*, **115**, 91–190.
- von der Heydt, R., Peterhans, E., and Baumgartner, G. (1984). Illusory contours and cortical neuron responses. *Science*, **224**, 1260–1262.
- von Grünau, M.W. (1986). A motion aftereffect for long-range stroboscopic apparent motion. *Perception and Psychophysics*, **40**, 31–38.
- Watson, A.B. (1986). Apparent motion occurs only between similar spatial frequencies. *Vision Research*, **26**, 1727–1730.
- Watson, A.B. and Ahumada, A.J. (1983). A look at motion in the frequency domain. NASA Technical Memo. 84352 (National Aeronautics and Space Administration, Washington, DC).
- Watson, A.B., Ahumada, A.J., and Farrell, J.E. (1983). The window of visibility: A psychophysical theory of fidelity in time-sampled motion displays. NASA Technical Paper 2211 (National Aeronautics and Space Administration, Washington, DC).
- Watson, A.B., Ahumada, A.J., and Farrell, J.E. (1986). Window of visibility: A psychophysical theory of fidelity in time-sampled visual motion displays. *Journal of the Optical Society of*

*America A*, **3**, 300–307.

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.

---

**TABLE 1**  
**PROPERTIES OF MOC FILTER**

---

<b>INPUT END</b>	<b>OUTPUT END</b>
(1) SENSITIVE TO DIRECTION- OF-CONTRAST	INSENSITIVE TO DIRECTION- OF-CONTRAST
(2) SENSITIVE TO ORIENTATION	SENSITIVE TO DIRECTION- OF-MOTION
(3) SHORT-RANGE INTERACTIONS	LONG-RANGE INTERACTIONS
(4) MONOCULAR	BINOCULAR

---

## APPENDIX: DEPENDENCE OF THRESHOLD ISI AND SOA ON FLASH SEPARATION AND DURATION

It is shown below how the model can be used to explain the lower threshold curves of Neuhaus (1930) which are graphed in Figure 3; namely, both threshold ISI and SOA increase as a function of flash separation  $W$ , although ISI decreases and SOA increases as a function of flash duration  $T$ . First we show that most of these properties follow from the use of sustained cells alone, and then that sustained-transient cell gating is sufficient to obtain them all. In both cases, we let 0 be the position of Flash 1 and  $W$  the position of Flash 2. Signify the right motion signal over time produced by Flash 1 at 0 by  $r_0(t)H$ , and the right motion signal produced by Flash 2 at 0 by  $r_w(t)He^{-W^2/2K^2}$ , as implied by equations (8), (10), and (11). Assume that at the motion threshold, the ratio of these two quantities is a constant, or

$$\frac{r_w(t)e^{-W^2/2K^2}}{r_0(t)} = \varepsilon. \quad (\text{A1})$$

### Transient Cells Always "On"

In order to obtain a rough idea of the implications of (A1) for the dependence of threshold ISI and SOA on the spatial separation  $W$ , flash duration  $T$ , and interstimulus interval  $I$ , we first assume that the transient activities are always 'on' and equal to 1.

Let the onset time of Flash 1 be  $t = 0$ . For simplicity in equation (4), set  $B = 0$ , then

$$r_0(t) = x_{0R}(t) = \begin{cases} \frac{J}{A}(1 - e^{-At}) & \text{if } 0 \leq t \leq T \\ \frac{J}{A}(1 - e^{-AT})e^{-A(t-T)} & \text{if } t \geq T, \end{cases} \quad (\text{A3})$$

and

$$r_w(t) = x_{wL}(t) = \begin{cases} 0 & \text{if } 0 \leq t \leq T + I \\ \frac{J}{A}(1 - e^{-A(t-T-I)}) & \text{if } T + I \leq t \leq 2T + I \\ \frac{J}{A}(1 - e^{-AT})e^{-A(t-2T-I)} & \text{if } t \geq 2T + I. \end{cases} \quad (\text{A4})$$

Inspection of (A3) and (A4) shows that ratio (A1) reaches its maximum at time  $t = 2T + I$ , when Flash 2 shuts off,  $r_w(t)$  is maximized, and  $r_0(t)$  is decaying. Substituting (A3) and (A4) both evaluated at time  $t = 2T + I$  into (A1) yields

$$\frac{(1 - e^{-AT})e^{-W^2/2K^2}}{(1 - e^{-AT})e^{-A(I+T)}} = \varepsilon, \quad (\text{A5})$$

or

$$e^{-W^2/2K^2} = \varepsilon e^{-A(I+T)}. \quad (\text{A6})$$

Taking the logarithm of both sides of (A6) yields

$$-W^2/2K^2 = \ln(\varepsilon) - A(I+T), \quad (\text{A7})$$

or

$$SOA = I + T = \frac{\ln(\varepsilon) + W^2/2K^2}{A} \quad (A8)$$

and

$$ISI = \frac{\ln(\varepsilon) + W^2/2K^2}{A} - T. \quad (A9)$$

According to (A8) and (A9), the threshold SOA and ISI for motion should both be increasing quadratic functions of spatial separation  $W$ . The theoretical ISI decreases with flash duration  $T$ , while the SOA does not depend on duration. All of these facts except the last are consistent with the data of Neuhaus (1930). This latter difficulty can be overcome by introducing the transient cells into the threshold computation.

### Sustained-Transient Gating

Let  $C = A$ ,  $D = 1$ , and  $B = E = 0$  in equations (4) and (5). From equations (8), (A3), and (A4), we obtain

$$\begin{aligned} r_W(t) &= x_{WL}(t)y_W^+(t) \\ &= x_{WL}(t)\left|\frac{d}{dt}x_W(t)\right| \\ &= \frac{J^2}{A}\left(1 - e^{-A(t-T-I)}\right)e^{-A(t-T-I)}, \end{aligned} \quad (A10)$$

and

$$\begin{aligned} r_0(t) &= x_{0R}(t)y_0^-(t) \\ &= x_{0R}(t)\left|\frac{d}{dt}x_0(t)\right| \\ &= \frac{J^2}{A}(1 - e^{-AT})^2e^{-2A(t-T)}, \end{aligned} \quad (A11)$$

during the interval of the second flash. After Flash 2 is shut off, its influence will not further increase. Thus,

$$\frac{r_W(t)}{r_0(t)} = \frac{(1 - e^{-A(t-T-I)})e^{AI}e^{A(t-T)}}{(1 - e^{-AT})^2}, \quad (A12)$$

which is maximized within the time interval  $T + I \leq t \leq 2T + I$  at time  $t = 2T + I$ . At  $t = 2T + I$ , (A1) becomes

$$\frac{e^{2AI}e^{AT}e^{-W^2/2K^2}}{(1 - e^{-AT})} = \varepsilon. \quad (A13)$$

Solving for  $I$  yields

$$ISI = \frac{1}{2A} \left[ \ln(\varepsilon) - AT + \ln(1 - e^{-AT}) + \frac{W^2}{2K^2} \right], \quad (A14)$$

and

$$SOA = I + T = \frac{1}{2A} \left[ \ln(\varepsilon) + AT + \ln(1 - e^{-AT}) + \frac{W^2}{2K^2} \right]. \quad (A15)$$

A comparison of (A15) with (A8), and (A14) with (A9), shows the effect of the transient cells. By (A15), the threshold SOA now *increases* with  $T$ , as in the data of Neuhaus (1930), for all values of  $A$ . The quadratic increase of SOA with distance  $W$  also obtains. By (A14), the threshold ISI increases quadratically with  $W$ , but *decreases* with  $T$  if

$$T > \frac{1}{A} \ln(2). \quad (A16)$$

These curves are plotted in Figure 37.