

A NEURAL MODEL OF FIRST-ORDER AND SECOND-ORDER MOTION PERCEPTION AND MAGNOCELLULAR DYNAMICS

Aijaz A. Baloch¹, Stephen Grossberg², Ennio Mingolla³, and C.A.M. Nogueira⁴
Department of Cognitive and Neural Systems

and
Center for Adaptive Systems
Boston University
677 Beacon Street
Boston, MA 02215

Submitted: December 1998

Journal of the Optical Society of America A, in press
Technical Report CAS/CNS-TR-96-030
Boston, MA: Boston University

Running Head: Neural Model of Second-Order Motion

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

Key words: first-order motion, second-order motion, , display, visual cortex, lateral geniculate nucleus, magnocellular processing, neural network.

1. Supported in part by the Defense Research Projects Agency and the Office of Naval Research (ONR N00014-95-1-0409).

2. Supported in part by the Defense Research Projects Agency and the Office of Naval Research (ONR N00014-95-1-0409) and the Office of Naval Research (ONR N00014-95-1-0657).

3. Supported in part by the Defense Research Projects Agency and the Office of Naval Research (ONR N00014-95-1-0409) and the Office of Naval Research (ONR N00014-94-1-0597).

4. Present address: Morgan Stanley & Co. Inc., 1585 Broadway, New York, NY 10019.

Acknowledgments: The authors wish to thank Robin Amos, Robin Locke and Diana Meyers for their valuable assistance in the preparation of the manuscript.

ABSTRACT

A neural model of motion perception simulates psychophysical data concerning first-order and second-order motion stimuli, including the reversal of perceived motion direction with distance from the stimulus (`, display`), and data about directional judgments as a function of relative spatial phase or spatial and temporal frequency. Many other second-order motion percepts that have been ascribed to a second non-Fourier processing stream can also be explained in the model by interactions between ON and OFF cells within a single, neurobiologically interpreted magnocellular processing stream. Yet other percepts may be traced to interactions between form and motion processing streams, rather than to processing within multiple motion processing streams. The model hereby explains why monkeys with lesions of the parvocellular layers, but not the magnocellular layers, of the lateral geniculate nucleus (LGN) are capable of detecting the correct direction of second-order motion, why most cells in area MT are sensitive to both first-order and second-order motion, and why after APB injection selectively blocks retinal ON bipolar cells, cortical cells are sensitive only to the motion of a moving bright bar's trailing edge. Magnocellular LGN cells show relatively transient responses while parvocellular LGN cells show relatively sustained responses. Correspondingly, the model bases its directional estimates on the outputs of model ON and OFF transient cells that are organized in opponent circuits wherein antagonistic rebounds occur in response to stimulus offset. Center-surround interactions convert these ON and OFF outputs into responses of lightening and darkening cells that are sensitive both to direct inputs and to rebound responses in their receptive field centers and surrounds. The total pattern of activity increments and decrements is used by subsequent processing stages (spatially short-range filters, competitive interactions, spatially long-range filters, and directional grouping cells) to determine the perceived direction of motion.

1 Introduction

1.1 First-Order and Second-Order Motion

Apparent motion percepts generated by displays in which nothing actually moves provide important clues to the neural processes that govern motion perception. Most of the old studies of motion perception could be attributed to what Braddick (1974) would later call the long-range mechanism. Braddick (1974) used random dot kinematograms wherein a rectangular area with horizontal or vertical orientation was displaced from one frame to the next. For appropriate values of spatial displacement and temporal interval between frames (interstimulus interval), subjects observed a clear perception of motion and figure-ground separation. Braddick observed that displacements beyond a quarter of a degree would not provide figure-ground separation. He also observed that an increase in the interstimulus interval decreased perceptual segregation. Braddick then suggested that two different processes govern apparent motion: one short-range and the other long-range. A spatial limit, or maximum displacement threshold D_{max} , was proposed over which the short-range process can be activated. D_{max} was later shown to vary significantly depending on the choice of parameters, such as an increment of the target area (Lappin and Bell, 1976; Baker and Braddick, 1982; Chang and Julesz, 1983a, 1983b; van Doorn and Koenderink, 1984), presentation of the target to more peripheral sites (Baker and Braddick, 1982), and use of multi-frame kinematograms (Nakayama and Silverman, 1984).

More recently, Bischof and Di Lollo (1990), Cavanagh and Mather (1989), Grossberg and Rudd (1989), and Sperling (1989) have argued that the differences in perception obtained for short-range and long-range processes can be more easily attributed to a difference in the stimuli used to test each case, with stimuli classified as first-order or second-order stimuli. A first-order stimulus is a stimulus whose motion can be discriminated by spatially tracking a difference of mean luminance or color over time. In an illustrative second-order motion stimulus, the density of moving dots inside and outside a central square is the same, so there is no difference in luminance between the regions. Second-order motion percepts can discriminate two such areas, even if their mean luminance and color are the same, if they differ in their spatial, temporal, or ocular distribution of mean luminance or color (Cavanagh and Mather, 1989), or when foreground and background vary in their binocular disparity or texture. The motion model developed here correctly detects the perceived direction of motion for a variety of first-order and second-order motion stimuli.

1.2 Fourier and Non-Fourier Motion

Sperling (1989) and Chubb and Sperling (1989) also distinguished between Fourier and non-Fourier apparent motion stimuli. If the space-time plots of one-dimensional spatial patterns contain oriented intensity contours, then their spatiotemporal signal is said to be Fourier in nature. Fourier stimuli can be detected by linear filters followed by half-wave rectification and standard motion analysis. If the plots do not contain oriented energy, then the stimuli are said to contain non-Fourier motion. They can be detected with nonlinear filters followed by full-wave rectification and standard motion analysis. Chubb and Sperling (1989) argued that, for some second-order stimuli, the short-range and long-range mechanisms can produce different results. For example, the perceived direction of motion reverses as the observer moves closer or farther from their , display, which is a variant of the reverse phi illusion of Anstis and Rogers (1975). In this stimulus, a grating of vertical bars are displaced to the left by an amount equal to 1/4 of the distance between two consecutive bars and the contrast of the bars is reversed. Subjects perceive motion to the left when observing the , display from nearby (Figure 1a) and motion to the right when observing the , display from afar (Figure 1b). In both cases, the strength of the perception is considerably weaker when compared to those

produced by the first-order and second-order stimuli described above. Chubb and Sperling (1989) argued that the far-view motion of , is detected by the short-range system and can be processed by a first-order Fourier mechanism, whereas the near-view motion of , is detected by a second-order mechanism and requires non-Fourier analysis and full-wave rectification. The model proposed herein utilizes a single processing stream to process both first-order and second-order motion stimuli. The model also suggests that various third-order motion stimuli (e.g., Lu and Sperling, 1995a) are due to a form-motion interaction between two or even three processing streams (Baloch and Grossberg, 1997; Francis and Grossberg, 1996a; Grossberg, 1991).

FIGURE 1

The present model analyses how monocular ON and OFF cells, at an early stage of magnocellular processing, respond through time to luminance increments and decrements before combining their outputs at *lightening cells* and *darkening cells* that the present modeling study predicts to exist. These two types of cells, which are predicted to exist early in the motion processing stream, play a role in the model similar to that of simple cells in the form processing stream. The lightening and darkening cells, in turn, input to spatially short-range filters which accumulate evidence for motion in a given direction. The pooled outputs from both lightening and darkening cells in a given direction mimmick human percepts of first-order and second-order motion in a variety of conditions. These results are consistent with recent experiments of Gellatly and Blurton (1996) showing that the spatiotemporal patterning of luminance increments and decrements through time determines these percepts, rather than distinct types of mechanisms. Our analysis hereby suggests that various second-order properties that have been attributed to a second processing stream are due to interactions between ON and OFF cells within a single processing stream.

2 Neural Substrate of Motion Processing

2.1 Parvocellular and Magnocellular Pathways

Some of the neural data that are clarified by model mechanisms are reviewed in this section and the next. Distinct ON and OFF channels for processing visual information arise at an early stage of retinal organization. Photoreceptors make direct synaptic contacts to the bipolar cells. Some bipolar cells are classified as the ON-center cells that are activated by direct illumination of cones. The OFF-center bipolar cells are inhibited by direct illumination of cones. Responses from the ON bipolar cells project to the ON ganglion cells while the responses from the OFF bipolar cells project to the OFF ganglion cells with amacrine cells mediating antagonistic interactions between the ON and OFF channels.

Enroth-Cugell and Robson (1966) found two distinct type of ganglion cells in the cats retina ganglia and classified them into into X and Y cells. The X cell small receptive fields (about 3 times smaller than the Y cells) and linear summation of spatial inputs are used in high-acuity vision and the processing of visual form. The Y cell larger receptive fields, nonlinear summation of spatial inputs, and rapidly conducting axons are used to process motion (Tolhurst, 1973). The sustained responses of X cells and the transient responses of Y cells lead to the alternate names sustained and transient cells, respectively (Clelland *et al.*, 1971). In macaque monkeys, ganglion cells have an analogous organization and the two major categories are labeled M and P cells. M cells respond more transiently than P cells to step changes in contrast, and the center of M cell receptive fields has a diameter 2-3 times larger than those of P cells.

Responses from the ganglion cells are projected to the LGN. In primates, the LGN is comprised of 6 layers numbered 1-6 from ventral to dorsal. Cells in the magnocellular layers 1 and 2 are larger,

and respond faster and more transiently than cells in the parvocellular layers 3-6. M ganglion cells project mainly to the magnocellular layers of the LGN and to a small portion of superior colliculus (Schiller and Malpeli, 1977). Cells at the parvocellular layers receive their inputs from P ganglion cells and respond in a more sustained way than the cells at magnocellular layers. Livingstone and Hubel (1988) have reported further differences between the magnocellular and parvocellular cells in terms of features like color, acuity, speed and contrast sensitivity.

Axons from LGN project primarily to layer 4C of cortical area V1. Layer 4C is subdivided in layers $4C\alpha$ and $4C\beta$. Projections from magnocellular layers of LGN contact layer $4C\alpha$ while those from parvocellular layers of LGN contact layer $4C\beta$. The segregation between parvocellular and magnocellular pathways found in LGN is thus maintained in V1.

From layer $4C\alpha$, magnocellular pathways involved in motion perception project to layer 4B, which then projects to cortical area MT, which is specialized to process visual motion (Albright, 1984; Maunsell and van Essen, 1983a; Newsome and Paré, 1988; Newsome *et al.*, 1985; Orban and Gulyas, 1988; van Essen, 1979; van Essen *et al.*, 1981; Zeki, 1974). Cells in MT thus have a predominantly magnocellular visual input (Maunsell *et al.*, 1990). Albright (1984) tested direction and orientation selectivity of V1 and MT cells, observing that virtually all cells in area MT were directionally selective and responses to first-order moving stimuli were stronger at area MT than at area V1. Albright (1992) showed that nearly all cells (99%) tested at area MT were selective to first-order motion and 87% of the same cells were also selective to second-order motion.

2.2 Cortical Responses to Motion After Parvocellular or Magnocellular Lesions

Schiller *et al.* (1990) tested the visual capacities of the magnocellular and parvocellular pathways and their projections. Seven rhesus monkeys were trained to perform visual detection discrimination tasks. In a control phase of the experiment, the animals were tested for contrast sensitivity, flicker detection, brightness discrimination, color, texture, pattern discrimination (same stimulus presented at different spatial frequency), shape perception, stereopsis, and motion. After the control phase, some monkeys had their parvocellular layers of LGN lesioned while some others had their magnocellular layers of LGN lesioned. The tests used during the control phase were repeated to observe the differences in their performance after lesions.

For the motion detection tasks, the monkeys were asked to fixate a point in the center of a screen. After fixation, a random array of spots filled the screen. In one small region (out of eight possible regions), the dots moved coherently. Detection was indicated by a direct saccade to the location of coherent motion. The results showed pronounced degradation in the performance of monkeys with magnocellular lesions while there was no change in the performance of monkeys with parvocellular lesions. Motion discrimination was further tested by changing the velocity or the direction of motion at one of the eight possible locations. Once again, monkeys with magnocellular lesions showed degradation in their performance. These results suggest that directional selectivity for continuous motion requires input from magnocellular transient cells, but not from parvocellular sustained cells.

Some other experiments have used the reversible inactivation of either magnocellular or parvocellular layers to examine their contribution to visual responses recorded in other areas of the visual cortex (Ferrera *et al.*, 1992; Maunsell *et al.*, 1992). In these experiments, inactivation was achieved by injecting either lidocaine or GABA. Results were quantified using a blocking index to compare responses before and after blocking: 0 corresponding to no effect and 1 to elimination of cortical response. In area MT, the blocking index after blocking the magnocellular layers of LGN was 0.75, while after blocking the parvocellular layers of LGN was 0.11.

Slaughter and Miller (1981) showed that injection of 2-amino-4-phosphonobutyrate (APB) produces prolonged hyperpolarization in retinal ON bipolar cells, making them unresponsive to light stimulation. Injection of APB had no effect on OFF cells. Schiller (1992) tested the effect of APB injection on the responses of directionally selective motion cortical cells (Figure 2). As a control, Schiller used a wide bright bar moving on a dark background over the receptive field of a directionally selective cortical cell before APB injection (Figure 3). The cell fired at the passage of both edges of the bar. After APB injection, the same cell fired only at the passage of the trailing edge of the bar. These results suggest that motion of the leading edge of a moving bright bar over a dark background is processed by the ON channel.

FIGURE 2

The model described herein simulates the psychophysical and neural data summarized above. In addition, the model predicts that monkeys with lesions in parvocellular layers, but not magnocellular layers, of the LGN should be able to detect and discriminate the correct direction of motion for second-order stimuli. Before describing the model, it is placed into a larger context by noting how it compares with other relevant motion models in the literature.

FIGURE 3

3 Gradient or Correlational Models of Motion Perception?

Most motion models fall into two categories: gradient models or correlational models. Gradient models detect data collected at single locations (Fennema and Thompson, 1979; Hadani *et al.*, 1980; Horn and Schunck, 1980; Limb and Murphy, 1975; Marr, 1982; Marr and Ullman, 1981) and employ a gating operation between a spatially oriented edge detector and a temporal luminance detector. If a cell representing a dark/bright edge (i.e., dark on the left side, bright on the right side) is activated when the temporal unit detects an increment of luminance at the edge location, the gating operation detects that the dark/bright edge is moving leftward (rightward in case of a bright/dark edge). Conversely, if the temporal unit detects a decrement of luminance at the dark/bright edge, the corresponding gating operation detects that the dark/bright edge is moving rightward (leftward in case of a bright/dark edge). Correlational models combine data that are separated both in space and time (Adelson and Bergen, 1985; Reichardt, 1961; van Santen and Sperling, 1984, 1985; Watson and Ahumada, 1985). In the original Reichardt (1961) detector, the delayed response from the left (right) filter is correlated with the response from the right (left) filter, and the output is the difference between these correlated responses.

The Motion Boundary Contour System (motion BCS) model incorporates aspects of both gradient and correlational models (Grossberg, 1991; Grossberg and Mingolla, 1993; Grossberg and Rudd, 1989, 1992); see Figure 4. In the motion BCS, the spatiotemporal visual signal is preprocessed by sustained and transient cells that elaborate properties of gradient models. The sustained cells have oriented receptive fields that generate responses to either dark/light or light/dark oriented stimuli, but not both. The activities of these simple cells are time-averaged and half-wave rectified to generate output signals. The transient cells have unoriented receptive fields that generate transient temporal responses in response to the onset or offset of stimuli, but not both. Their activities are also time-averaged and half-wave rectified to generate output signals. The outputs of these sustained and transient cells are then multiplied, or gated, at each position, as in the gradient models, to derive a local estimate of direction-of-motion. Outputs from gated cells sensitive to the same orientation and direction-of-contrast that lie along a given direction-of-motion are then combined

via short-range spatial filters (the analog of D_{max}) to accumulate evidence of motion in that direction. This correlational operation results in four types of cells, from all the gated combinations of light/dark or dark/light sustained cells and ON or OFF transient cells. All of these gated cells are sensitive to a particular direction-of-contrast as well as a particular direction-of-motion. A long-range spatial filter then correlates signals again by pooling outputs of gated cells that are sensitive to the same direction-of-motion. Long-range filtering pools signals from both directions-of-contrast, all orientations, and from both eyes. By combining the half-wave rectified output of simple cells that are sensitive to opposite contrast polarities, these “motion complex cells” carry out a full-wave rectification of the input. A contrast-enhancing competition then selects the cell or cells which receive the largest total inputs. The competition hereby votes for which direction has the most evidence. These competitively-sharpened long-range filter cells are the first true direction-of-motion cells in the model because they combine signals from all previous cell types that are sensitive to a particular direction-of-motion.

FIGURE 4

Grossberg and Rudd (1989, 1992) used this model to simulate many data about short-range and long-range apparent motion, including beta motion, gamma motion, delta motion, split motion, Ternus and reverse-contrast Ternus motion, brief flash speed-up, and aspects of Korté’s laws (Bartley, 1941; de Silva, 1926; Giaschi and Anstis, 1989; Kolers, 1972; Korté, 1915; Pantle and Picciano, 1976; Ternus, 1926/1950). Grossberg and Mingolla (1993) extended the model to two-dimensions to simulate how multiple moving orientations could all be pooled into a single direction-of-motion. Francis and Grossberg (1996a, 1996b) modeled how a $V2 \rightarrow MT$ pathway linking form processing in the $V1 \rightarrow V2$ cortical stream and motion processing in the $V1 \rightarrow MT$ cortical stream could be used to provide a complete simulation of Korté’s laws and related data about form-motion interactions (Cavanagh *et al.*, 1989; Kolers, 1972; Korté, 1915; Mather, 1988; Neuhaus, 1930; Ramachandran, 1985; von Grünau, 1979). This motion BCS model has thus been used to simulate a large set of data about short-range and long-range motion perception.

The data of Schiller *et al.* (1990) suggest, however, that the model needs to be refined. This is true because motion perception is spared when oriented sustained cells that are activated by the parvocellular layers of the LGN are blocked by APB. If oriented sustained cells are not needed for effective motion perception, then one needs to explain how processing that is based on the responses of transient cells alone can be used to generate precise estimates of object speed and direction without undermining the other explanations of the model. An initial effort to do this was reported in Nogueira *et al.* (1993). These results were followed by further model development in Chey *et al.* (1994, 1997) who simulated how a coherent representation of object direction and speed could be generated by signals contaminated by aperture ambiguities. By using a multiple-scale short-range filter whose larger scales tend to process higher speeds, Chey *et al.* (1994) simulated how speed estimates are influenced by input contrast, duration, dot density, and spatial frequency. Chey *et al.* (1997) showed how the addition of competition, long-range filters, and a directional grouping and attentive priming network can provide a solution to the aperture problem in which unambiguous feature tracking signals capture ambiguous aperture signals and attention can selectively prime a desired direction-of-motion. A schematic of this modified motion BCS is given in Figure 5, which indicates that oriented sustained cells are no longer used.

FIGURE 5

The present extension of the motion BCS in Figure 6a elaborates the design of the transient cells and how they activate the short-range filters. A key advance is that contributions from opponent

pairs of ON cells and OFF cells are modeled. Antagonistic rebounds, whereby offset of ON (or OFF) cell activity generates a transient onset of OFF (or ON) cell activity, play a central role in simulating data about second-order motion. These direct and rebound ON and OFF responses go through center-surround networks whose outputs are combined at lightening cells and darkening cells. These latter cells play a role much like that of simple cells in the form processing stream, in that they pool input from both ON cells and OFF cells to form responses that are sensitive to a prescribed polarity of change (Ferster, 1988; Gove *et al.*, 1995; Liu *et al.*, 1992). The outputs from these lightening and darkening cells then activate short-range filters which pool evidence for motion in a given direction.

Pooling the short-range filter contributions to a given direction from both lightening and darkening cells generates the motion directions that humans perceive in response to first-order and second-order stimuli under a variety of conditions. These results are simulated in Figures 10-14. More, however, is required of the model. It needs to be consistent with the larger motion BCS theory of Figure 5. In particular, the perceived motion directions need to survive the effects of long-range filtering. For this to happen in all cases, including the case when the , display is viewed from afar, it is sufficient to process the lightening and darkening cell outputs according to the same mechanisms in Figure 5 that were originally derived to explain other data, notably data about motion capture and long-range apparent motion; namely, the darkening and darkening cells outputs go through directional short-range filters to accumulate evidence for a given direction before competing across directions and then activating the long-range filters. The results of these simulations are shown in Figures 16-20. The remainder of the article explains in greater detail how these mechanisms generate the simulated percepts.

4 The Proposed Role of Lightening and Darkening Cells

The model will be described in two stages. First the stages through the lightening and darkening cells will be described in order to emphasize their key role in tracking the temporal pattern of luminance increments and decrements. Their outputs are then pooled to show all the simulated effects in the simplest possible framework. Then their outputs will be embedded into the larger motion BCS model to show how all these percepts emerge in a model that can also explain a wide variety of other motion data, including data about global motion capture, motion speed and direction, long-range apparent motion, and directional attentive priming.

Figure 6b shows more mechanistic details of the simpler version of the model that is schematized in Figure 6a. Level 1 of the model represents the visual input as bright or dark signals. These signals are fed to unoriented transient cell filters at Level 2. These filters detect temporal changes in the input and represent them at opponent ON and OFF transient cells. ON (OFF) cells fire at the onset of a bright (dark) stimulus or at the offset of a dark (bright) stimulus.

FIGURE 6

ON and OFF cells input to Level 3 via on-center off-surround kernels where they are organized into lightening and darkening cells. In Figure 6b, lightening cells are represented on the left and darkening cells on the right. Both ON and OFF cells contribute to the activation and deactivation of the lightening and darkening channels so that the segregation between ON and OFF channels is broken. This is in accord with neurophysiological evidence that the ON and OFF systems remain largely segregated up to the lateral geniculate nucleus and then converge in the striate cortex (Schiller, 1992). A lightening cell is excited by the ON cells in its center and the OFF cells in its surround and inhibited by the OFF cells in its center and the ON cells in its surround. Similarly,

a cell in the darkening channel is excited by the OFF cells in its center and the ON cells in its surround and inhibited by the ON cells in its center and the OFF cells in its surround. For example, the onset a bright spot in the absence of other signals makes a location appear brighter and as a consequence the area in its immediate neighborhood appears darker. This is represented by the cells in lightening channel at the spatial location corresponding to the stimulus center and by the cells in the darkening channel at the spatial locations corresponding to its surround. Therefore, the cells in lightening channel represent a local increase in brightness while the cells in the darkening channel represent a local increase in darkness.

It is important to distinguish the functions of two types of cross-talk in the model; namely, the cross-talk between ON and OFF cells in the lightening and darkening channels at Level 3 and and the cross-talk between the opponent ON and OFF channels at Level 2 (Figure 6b). At the opponent ON and OFF cells, cross-talk produces an antagonistic rebound of activity in the opposite channel at the offset of a signal. At the lightening and darkening cells, cross-talk permits pooling of opponent signals originating from the surround with the signals at the center. This allows the model to simulate reversed apparent motion from reversed luminance contrast. An illustration of how this is achieved, for example in the , display when observed from afar (Figure 1b), is given in Figure 7. A stimuli consisting of two simultaneous bright spots is presented in Frame 1 and switched off in Frame 2 (Figure 7a). The ON cells at the locations of the spots fire during Frame 1. These ON cells then excite the lightening cells at those locations and the darkening cells in their surround. The width of the activity in the darkening channel depends upon the size of the surround. When the spots are switched off, antagonistic rebound transiently turns on the corresponding OFF cells which, in turn, excite the darkening cells at the locations where the bright spots were removed and the lightening cells in their surround.

FIGURE 7

Figure 7b, shows the first three Frames of a segment of , display when seen from afar. As discussed above, when the bright spots are presented in Frame 1, the ON cells at the locations of the spots fire and in turn activate the lightening cells at those locations and the darkening cells in their surround. In Frame 2, the bright spots are removed and dark spots are presented to their left. The OFF cells in this case fire due to two different kind of processes: first, due to antagonistic rebound at the offset of the bright spots and, second, due to the onset of the dark spots. These OFF cells excite the lightening cells to the right of removed bright spots and to the left of dark spots. The lightening cells to the left of removed bright spots and to the right of dark spots, in addition to the excitatory signals from OFF cells in the surround, also receive inhibitory signals from OFF cells at those locations and therefore remain inactive. Hence, if the , display is viewed from a far enough distance, it allows the activities due to offset of bright spots and onset of dark spots to fall close to each other, and the subsequent processing stages of the model time-average and threshold these activities to represent rightward direction-of-motion. Similar arguments apply to the activities of darkening cells and subsequent time frames.

When the , display is viewed from nearby (Figure 7c), the lightening (darkening) cell activations due to antagonistic rebound of bright (dark) spots and the onset of dark (bright) spots fall some distance away from each other (how far depends on the size of surround regions of lightening and darkening cells). These activities when time-averaged, thresholded, and pooled by subsequent processing stages represent leftward direction-of-motion.

The outputs from lightening and darkening cells are then fed to their respective short-range spatial filters at Level 4. These spatially averaged activities are thresholded and pooled at Level 5 in the simple version of model in Figure 6b. The mathematical equations of the model are given

in Section 5. The reader who wishes to study simulations of model performance first can skip to Sections 6 and 9, where we use the model to simulate neurophysiological data about the effects of anatomical lesions and APB injections on motion processing, and psychophysical data about the reversal of perceived motion direction with distance from the stimulus (Chubb and Sperling, 1989). Sections 11 and 12 discuss other data and models which suggest that various first-order and second-order motion percepts are processed by a single processing stream. In particular, experiments of Lu and Sperling (1995a), among others, on second-order motion can naturally be explained by the model. The model also indicates at what neurophysiological stages these explanations can be tested by subsequent experiments.

5 Mathematical Description of Model Dynamics

5.1 Level 1: Input Representation

Level 1 of the model registers the input pattern and directs it to the ON and OFF cells of Level 2 (Figure 6b). Let s_i^+ represent the response to a bright stimulus and s_i^- to a dark stimulus at the i^{th} location. Then

$$s_i^+ = \begin{cases} 1.0 & \text{when the bright stimulus is on} \\ 0.0 & \text{otherwise} \end{cases}, \quad (1)$$

and

$$s_i^- = \begin{cases} 1.0 & \text{when the dark stimulus is on} \\ 0.0 & \text{otherwise} \end{cases}. \quad (2)$$

5.2 Level 2: Unoriented Transient Filter

Level 2 detects temporal changes in the input signal via ON and OFF unoriented transient cells. ON (OFF) cells fire either at the onset of an increase (decrease) in luminance or, via antagonistic rebound, at the offset of a decrease (increase) in luminance. A gated dipole circuit is used to represent these opponent transient changes (Grossberg, 1972). Such a circuit has previously been used to model transient responses to visual cues under a variety of conditions (Carpenter and Grossberg, 1981; Francis *et al.*, 1994; Francis and Grossberg, 1996b; Gaudio, 1994; Grossberg, 1976; Ögmen and Gagné, 1990). In both the ON and OFF channels, chemical transmitters gate signals in their pathways in such a way as to attempt to maintain unbiased transduction. Their slow rates of habituation and recovery determines antagonistic rebounds in the circuit.

Figure 8 illustrates the functioning of such a gated dipole circuit. Initially, when no phasic inputs are present, both channels receive equal tonic arousal signals γ_u . Therefore, activities u_1 and u_2 are equal. They cancel each other due to opponent interaction and both channels remain subthreshold. When a phasic input, s^+ , due to presentation of bright stimulus is turned on, u_1 receives both tonic and phasic inputs. Activity u_1 gets larger than activity u_2 and neurotransmitter v_1 habituates, or inactivates, slowly. Since u_1 responds faster than v_1 , initially u_3 becomes larger than u_4 and u_5 starts firing above threshold, resulting in u^{ON} signal. When neurotransmitter v_1 is sufficiently habituated, u^{ON} becomes subthreshold although the stimulus remains on. When the bright stimulus is removed, u_1 and u_2 receive only tonic input. Since neurotransmitter v_1 was inactivated during presentation of the bright stimulus s^+ , its value is now less than that of the neurotransmitter v_2 . Therefore u_4 becomes larger than u_3 . This results in a positive response at u_6 and an OFF response u^{OFF} is generated via an antagonistic rebound. Signal u^{OFF} becomes zero after v_1 accumulates back to its equilibrium value. Similar arguments apply for the onset of a

dark stimulus. In summary, an ON cell fires at the onset of a bright stimulus and at the offset of a dark stimulus, whereas an OFF cell fires at the offset of a bright stimulus and the onset of a dark stimulus.

FIGURE 8

The ON-channel of the dipole responds to a net increase s_i^+ in the luminance, as in (1), while the OFF-channel responds to a net decrease s_i^- , as in (2):

ON-Channel Input Stage

Let

$$\frac{d}{dt}u_{1i} = -A_2u_{1i} + s_i^+ + \gamma_u, \quad (3)$$

where u_{1i} is the activity of the ON channel, s_i^+ is the signal from Level 1 as described in (1), and γ_u is a tonic arousal level.

Off-Channel Input Stage

Let

$$\frac{d}{dt}u_{2i} = -A_2u_{2i} + s_i^- + \gamma_u, \quad (4)$$

where u_{2i} and s_i^- have analogous definitions and γ_u is the same arousal level as in (3).

ON-Transmitter Transmitter Production-Inactivation

Let v_{1i} be the habituating transmitter that multiplies, or gates, the half-wave rectified signal $[u_{1i}]^+$ in the ON channel. Transmitter v_{1i} varies more slowly than u_{1i} via the equation:

$$\frac{d}{dt}v_{1i} = B_2(1 - v_{1i}) - C_2[u_{1i}]^+v_{1i}, \quad (5)$$

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

OFF-Transmitter Transmitter Production-Inactivation

Similarly, the slowly varying habituating transmitter v_{2i} in the OFF channel gates the OFF channel signal $[u_{2i}]^+$ via the equation:

$$\frac{d}{dt}v_{2i} = B_2(1 - v_{2i}) - C_2[u_{2i}]^+v_{2i}. \quad (6)$$

Equations (5) and (6) control the level of available neurotransmitter. An increase in signal u_i increases the inactivation and release of transmitter via the mass action term $-C_2[u_i]^+v_i$. Transmitter accumulates to the maximum value of 1.0 via term $B_2(1 - v_i)$ at rate B_2 .

Transmitter-Gated ON-Activation

The transmitter-gated signal $u_{1i}^+v_{1i}$ activates the next stage of ON channel processing:

$$\frac{d}{dt}u_{3i} = -A_2u_{3i} + D_2[u_{1i}]^+v_{1i}. \quad (7)$$

Transmitter-Gated OFF-Activation

The same thing happens in the OFF channel:

$$\frac{d}{dt}u_{4i} = -A_2u_{4i} + D_2[u_{2i}]^+v_{2i}. \quad (8)$$

Normalized Opponent ON-Activation

Competition between the ON and OFF channels determines the opponent ON and OFF activations:

$$\frac{d}{dt}u_{5i} = -A_2u_{5i} + (E_2 - u_{5i})u_{3i} - (F_2 + u_{5i})u_{4i}. \quad (9)$$

Normalized Opponent OFF-Activation

Similarly,

$$\frac{d}{dt}u_{6i} = -A_2u_{6i} + (E_2 - u_{6i})u_{4i} - (F_2 + u_{6i})u_{3i}, \quad (10)$$

where A_2 is the passive decay rate, and E_2 and F_2 are the upper and lower bounds of ON and OFF cell activation.

ON Output:

These output are then thresholded and rectified. This forms the input for Level 3.

$$u_i^{ON} = [u_{5i} - , u]^+. \quad (11)$$

OFF Output:

Similarly,

$$u_i^{OFF} = [u_{6i} - , u]^+. \quad (12)$$

5.3 Level 3: Lightning and Darkening Channels

Two center-surround networks process the ON and OFF outputs from Level 2. This operation spatially contrast enhances the ON and OFF output signals. Let w_i^L and w_i^D represent the activity of a cell at the i^{th} position in the lightning and darkening channels, respectively.

Lightning Channel

$$\begin{aligned} \frac{dw_i^L}{dt} = & -A_3w_i^L + (B_3 - w_i^L) \left[\sum_j G_{ji}u_j^{ON} + \sum_j H_{ji}u_j^{OFF} \right] \\ & - (C_3 + w_i^L) \left[\sum_j H_{ji}u_j^{ON} + \sum_j G_{ji}u_j^{OFF} \right]. \end{aligned} \quad (13)$$

Darkening Channel

$$\begin{aligned} \frac{dw_i^D}{dt} = & -A_3w_i^D + (B_3 - w_i^D) \left[\sum_j G_{ji}u_j^{OFF} + \sum_j H_{ji}u_j^{ON} \right] \\ & - (C_3 + w_i^D) \left[\sum_j H_{ji}u_j^{OFF} + \sum_j G_{ji}u_j^{ON} \right], \end{aligned} \quad (14)$$

where G_{ji} and H_{ji} are the Gaussian center and surround kernels with parameter α_w controlling their amplitude and parameters σ_c and σ_s controlling their sizes:

$$G_{ji} = \frac{\alpha_w}{\sigma_c\sqrt{2\pi}} \exp\left[-\frac{(j-i)^2}{2\sigma_c^2}\right], \quad (15)$$

$$H_{ji} = \frac{\alpha_w}{\sigma_s \sqrt{2\pi}} \exp\left[\frac{-(j-i)^2}{2\sigma_s^2}\right]. \quad (16)$$

Equations (13) and (14) are shunting center-surround equations. The first term $-A_3 w_i$ controls passive decay. The second term, $(B_3 - w_i^L)[\sum_j G_{ji} u_j^{ON} + \sum_j H_{ji} u_j^{OFF}]$, in (15) describes the excitatory signals from the ON cells in the center ($\sum_j G_{ji} u_j^{ON}$) and the OFF cells in the surround ($\sum_j H_{ji} u_j^{OFF}$); that is, bright stimuli in the center or dark stimuli in the surround excite the lightening channel. Term $(B_3 - w_i^L)$ is a shunting term which sets the maximum possible activity of w_i^L at B_3 . The last term, $(C_3 + w_i^L)[\sum_j H_{ji} u_j^{ON} + \sum_j G_{ji} u_j^{OFF}]$, in (15) describes the inhibitory signals from the ON cells in the surround ($\sum_j H_{ji} u_j^{ON}$) and the OFF cells in the center ($\sum_j G_{ji} u_j^{OFF}$); that is, dark stimuli in the center and bright stimuli in surround inhibit the lightening channel. The shunting term $(C_3 + w_i^L)$ sets the minimum activity of w_i^L at $-C_3$. Similarly, the second term, $(B_3 - w_i^D)[\sum_j G_{ji} u_j^{OFF} + \sum_j H_{ji} u_j^{ON}]$, in (16) is the excitatory signal from the OFF cells in the center ($\sum_j G_{ji} u_j^{OFF}$) and the ON cells in the surround ($\sum_j H_{ji} u_j^{ON}$); that is, dark stimuli in the center or bright stimuli in the surround excite the darkening channel. Term $(C_3 + w_i^D)[\sum_j H_{ji} u_j^{OFF} + \sum_j G_{ji} u_j^{ON}]$, describes the inhibitory signals from the ON cells in the center ($\sum_j H_{ji} u_j^{OFF}$) and the OFF cells in the surround ($\sum_j G_{ji} u_j^{ON}$); that is, bright stimuli in the center or dark stimuli in the surround inhibit the darkening channel.

5.4 Level 4: Short-Range Spatial Filtering

The next operation pools the activations of cells that are spatially close by means of a short-range spatial filter. A separate filter is used for the lightening and darkening channels. A Gaussian kernel P_{ji} ensures that the contributions from adjacent neighbors are larger than the contributions from more distant neighbors. Let y_i^L and y_i^D denote the activity of the i^{th} filter in lightening and darkening channels, respectively.

Lightening Channel

$$\frac{dy_i^L}{dt} = -A_4 y_i^L + (B_4 - y_i^L) \sum_j P_{ji} [w_j^L]^+. \quad (17)$$

Darkening Channel

$$\frac{dy_i^D}{dt} = -A_4 y_i^D + (B_4 - y_i^D) \sum_j P_{ji} [w_j^D]^+. \quad (18)$$

The Gaussian kernel P_{ji} is defined by:

$$P_{ji} = \frac{\alpha_y}{\sigma_y \sqrt{2\pi}} \exp\left[\frac{-(j-i)^2}{2\sigma_y^2}\right]. \quad (19)$$

For simplicity, the Gaussian spatial filter is chosen to be of a single-scale and isotropic. This is sufficient for the cases where the stimuli generate motion in one dimension. Time-averaging followed by thresholding arranges data in the direction-of-motion. Multiscale short-range anisotropic spatial filters that accumulate evidence for motion in a particular direction are essential for a two-dimensional motion grouping system to detect object speed and direction (Chey *et al.*, 1995).

5.5 Level 5: Lightening-Darkening Pooling

The thresholded outputs from the short-range spatial filter are combined at Level 5. Let z_i be the activity of the i^{th} node. Then

$$z_i = [y_i^L - , y]^+ + [y_i^D - , y]^+ \quad (20)$$

The cells z_i at Level 5 are sensitive to direction-of-motion and insensitive to direction-of-contrast as in the Grossberg-Rudd model and its subsequent elaborations.

6 Model Simulations

6.1 Simulation Parameters and Layout

The model was used to simulate examples of first-order motion, first-order motion after blocking of ON cells, second-order motion, and , display second-order motion as seen from near and far. The parameters for all simulations were: $A_2 = 10.0$, $B_2 = 0.05$, $C_2 = 5.0$, $D_2 = 200.0$, $E_2 = 5000.0$, $F_2 = 5000.0$, $\gamma_u = 20.0$, $, u = 0.2$, $A_3 = 0.4$, $B_3 = 1.0$, $C_3 = 0.6$, $\alpha_w = 10.0$, $\sigma_c = 1.5$, $\sigma_s = 6.0$, $A_4 = 1.0$, $B_4 = 1.0$, $\alpha_y = 15.0$, $\sigma_y = 2.0$ and $, y = 0.73$.

There were 100 nodes at each layer, 11 time frames and each frame lasting 50 units of time. The equations were solved using a fourth-order Runge-Kutta algorithm on a Sun Sparcstation 5 computer. The results are shown as space-time plots. The horizontal axis corresponds to space while the vertical axis corresponds to time. Time evolves in the upward direction. When representing stimuli, white spots in the plot indicate bright stimuli, black spots indicate dark stimuli and grey spots indicate that no stimuli is present. When plotting cell responses at various levels of the model, brighter locations indicate higher cell activation with black indicating zero activity.

6.2 Simulation of Antagonistic Rebound

Before discussing simulations using space-time plots, we clarify how antagonistic rebounds look by describing two simulations of a Level 2 cell's transient ON and OFF responses: (1) a stimulus is presented and switched off; (2) it is replaced by a stimulus of opposite contrast when it is switched off. Figure 9a shows the results for the first case. Since ON and OFF channels are symmetric, presentation of a bright (s_i^+) or a dark stimulus (s_i^-) yields the same result in their respective channels. The time sequence of stimulus presentation is shown at the bottom of the plot. The ON and OFF transients have approximately the same maximum level of activation in this case. Figure 9b shows the results of the second case in which a stimulus is replaced by the stimulus of opposite contrast. In this case the OFF (ON) response for a bright (dark) stimulus is larger. In our simulation examples discussed next, the second situation is applicable only in case of second-order motion stimulus when bright locations are replaced by dark locations and vice versa. In all other cases, the bright and dark locations are simply switched off.

FIGURE 9

6.3 Simulation of First-Order Motion

A simple example of a first-order stimulus is a bright bar sliding horizontally. Figure 10a shows the spatiotemporal representation of cross-section of a bright bar sliding horizontally to the right over a grey background. The bright bar is 30 units wide and covers nodes 11-40 at Frame 1. Starting at Frame 2, the bar moves to the right at a speed of 5 units per frame.

FIGURE 10

Figure 10b shows the ON cell activity (u_i^{ON}) and Figure 10c the OFF cell activity (u_i^{OFF}). Figure 10d shows the response at the lightening channel (w_i^L) and Figure 10e at the darkening channel (w_i^D) of Level 3. Figures 10f and 10g show the thresholded rectified responses of the short-range spatial filter (Level 4) for the lightening ($[y_i^L - , y]^+$) and darkening ($[y_i^D - , y]^+$) channels, respectively. Figure 10h shows the pooled response (z_i) at Level 5. The result shows that the model is sensitive to the motion of both edges of the bright sliding bar (Schiller, 1992). Since the model is based on transient cell responses alone, it agrees with the findings of Schiller *et al.* (1990) that the magnocellular pathways are sufficient to detect and discriminate first-order continuous motion.

6.4 Simulation of First-Order Motion With Blocked ON Cells

APB injections block the responses of ON bipolar cells without affecting the responses of OFF cells (Slaughter and Miller, 1981). In the model, the outputs of ON and OFF channels are independent of each other. This clarifies how selective blocking of either channel alone can occur without affecting the other. Schiller (1992) found that blocking ON bipolar cells using APB makes direction-selective cortical cells in monkeys insensitive to the leading edge of a moving bright bar.

Blocking the output from the ON channel of Level 2 was achieved by raising the threshold θ_u in (13) such that the responses u_i^{ON} were always zero. The stimulus used was the moving bright bar of the first-order simulation described in Figure 11a. Figure 11b shows the ON cell activity (u_i^{ON}) and Figure 11c the OFF cell activity (u_i^{OFF}). The responses at the lightening channel (w_i^L) and the darkening channel (w_i^D) of Level 3 are shown in Figures 11d and 11e, respectively. No activity is present at the onset of the bar in Frame 1 or at the locations of the leading edge in both channels. Figures 11f and 11g show the thresholded rectified responses of the short-range spatial filter (Level 4) for the lightening ($[y_i^L - , y]^+$) and darkening ($[y_i^D - , y]^+$) channels, respectively. Figure 11h shows the pooled response (z_i) at Level 5. The results show that the model is insensitive to the leading edge of the moving bright bar when the ON channel is blocked. The model predicts that blocking the OFF bipolar cells would make direction-selective cortical cells insensitive to the trailing edge. Moreover, if a dark moving bar was used instead of a bright bar, then blocking ON (OFF) cells should make the system insensitive to the trailing (leading) edge.

FIGURE 11

6.5 Simulation of Second-Order Motion

A contrast-reversing noise field (Chubb and Sperling, 1988) was used as an example of second-order motion. Figure 12a shows spatiotemporal plot of this stimulus. A random pattern of ten contiguous bright and dark bars (10 units wide) is presented at Frame 1. Starting at Frame 2, these bars reverse their contrast one bar per frame from left to right. A motion to the right is observed.

FIGURE 12

Figure 12b shows the ON cell activity (u_i^{ON}) and Figure 12c the OFF cell activity (u_i^{OFF}). Figures 12d and 12e show the responses at the lightening channel (w_i^L) and the darkening channel (w_i^D) of Level 3, respectively. Figures 12f and 12g show the thresholded rectified responses of the short-range spatial filter (Level 4) for lightening ($[y_i^L - , y]^+$) and darkening ($[y_i^D - , y]^+$) channels, respectively. At the start of the experiment at Frame 1, a motion signal is seen at the bright locations in the lightening channel and the dark locations in the darkening channel. When a bar switches from bright to dark, its transient response is captured at the location of the bar in the lightening channel and in its surround in the darkening channel, and vice versa. These responses

are pooled at Level 5, as shown in Figure 12h. This output tracks the motion from left to right. The model also predicts that, as in first-order motion, macaque monkeys with lesions in the parvocellular layers of LGN should be able to detect second-order motion. This was not tested by Schiller *et al.* (1990).

Harris and Smith (1992) also used this stimulus to test if it would evoke optokinetic nystagmus (OKN) in an attempt to localize the site of detection of second-order motion. They found that, although the detection of correct direction of motion in this case was strong, it did not evoke OKN. As the first-order stimuli evoke OKN they concluded that the first-order and second-order stimuli are processed by two different channels. Our simulation results of this experiment suggest that the signal for OKN is perhaps tapped before the long-range filtering stage. For example, for first-order stimulus, the signals before the long-range filter stage already show correct directional preference (Figures 10f and 10g), whereas for second-order stimulus the signals are still ambiguous (Figures 12f and 12g). This ambiguity is resolved in the model at the long-range filter stage (Figure 12h).

This hypothesis about the possible site for signaling OKN is consistent with recent data showing that vergence eye movements are also processed at an early stage of visual processing. For example, Masson *et al.* (1997) used anticorrelated dense patterns to show that vergence eye movements derive their visual inputs from an early cortical processing stage. For example, they reported that: “These data indicate that the vergence eye movements initiated at ultrashort latencies result solely from locally matched binocular features, and derive their visual input from an early stage of cortical processing before the level at which depth percepts are elaborated” (p 283). The long-range filter stage in the model is proposed to occur in cortical area MT, which is the model site at which motion depth percepts are elaborated. It may thus be that several types of signals for the control of eye movements derive their visual inputs from processing stages earlier than those at which depth percepts are elaborated.

6.6 Simulation of , Display Motion - Near View

Figure 13a shows the spatiotemporal layout of the , display (Chubb and Sperling, 1989) when seen from nearby. A grating of bars, each 20 units wide, shifts to the left by 1/4 spatial distance between two consecutive bars. Therefore, the distance between two bars equals 80 units. Figure 13b shows the ON cell activity (u_i^{ON}) and Figure 13c the OFF cell activity (u_i^{OFF}). Figures 13d and 13e show the responses at the lightening channel (w_i^L) and the darkening channel (w_i^D) of Level 3, respectively. Figures 13f and 13g show the thresholded rectified responses of the short-range spatial filter (Level 4) for lightening ($[y_i^L - , y]^+$) and darkening ($[y_i^D - , y]^+$) channels, respectively. Figure 13h shows the pooled response (z_i) at Level 5. The output tracks motion to the left; i.e., in the direction of the motion of the grating.

FIGURE 13

6.7 Simulation of , Display Motion - Far View

Figure 14a shows the spatiotemporal plot of the , display as seen from afar. The display is shrunk 4 times relative to the display when seen from nearby (Chubb and Sperling, 1989). Each segment of the grating (dark or bright) is therefore 5 units wide and the distance between the centers of the two consecutive segments now equals 20 units. Figure 14b shows the ON cell activity (u_i^{ON}) and Figure 14c the OFF cell activity (u_i^{OFF}). Figures 14d and 14e show the response at lightening channel (w_i^L) and darkening channel (w_i^D) of Level 3, respectively. Although the grating is shifting to the left, the cells at this stage already begin to prefer the rightward direction of motion. This is due to the cross-talk between ON and OFF cells as described in Section 4. Figures

14f and 14g show the thresholded rectified responses of the short-range spatial filter (Level 4) for lightening ($[y_i^L - , y]^+$) and darkening ($[y_i^D - , y]^+$) channels, respectively. Since the activities due to off-surround contributions in a channel now fall within the effective bandwidth of the Gaussian short-range filter, their average over time is now organized to become sensitive to the rightward direction of motion and is further contrast-enhanced by thresholding. Figure 14h shows the pooled response (z_i) at Level 5. The output detects the motion to the right which is in the opposite direction of the motion of the grating.

FIGURE 14

7 Additional Motion BCS Model Mechanisms

The results in Figures 10-14 highlighted the role of lightening and darkening cells. We now embed these cells in the full motion BCS model of Figure 5 to show how they give rise to directionally sensitive motion output cells. When the opponent ON and OFF channels and the lightening and darkening cells of Figure 6a are embedded within the more comprehensive motion BCS model, all of the above percepts can again be simulated despite the smoothing effects of the long-range filter. The long-range filter is the stage at which contributions from lightening and darkening cells are finally pooled. These long-range filter cells are also the ones at which sensitivity to direction-of-motion and insensitivity to direction-of-contrast is finally achieved by pooling signals of opposite contrast polarity (Chey *et al.*, 1997; Grossberg and Rudd, 1989, 1992). Before that stage is reached, the unidirectional transient cell responses are progressively transformed into directional cells that are capable of using unambiguous feature tracking motion signals to capture ambiguous motion signals that arise due to the aperture problem and to thereby generate global representations of an object's speed and direction (Chey *et al.*, 1997). The relevant processing stages are as follows.

An early stage in the transformation of unidirectional transient responses uses an inhibitory veto mechanism (Emerson and Gerstein, 1977; Ganz, 1984; Goodwin *et al.*, 1975, Heggelund, 1984). Barlow and Levick (1965) first showed that inhibition was crucially involved in the function of directionally selective ganglion cells in the rabbit retina. They concluded that the directional selectivity of these cells was brought about through inhibitory lateral connections, probably mediated by retinal horizontal cells. These directionally-specific inhibitory connections veto responses in nearby cells, implementing a kind of logical NOT operation. The ganglion cells responded to single light flashes with much the same threshold as paired flashes presented in the direction that was not vetoed by inhibition. Gamma-aminobutyric acid (GABA) mediates inhibition in directional rabbit retina cells. Introduction of a GABA antagonist into the rabbit retina eliminates the selectivity of the previously directionally selective cells, causing them to respond equally well to both directions of movement (Ariel and Daw, 1982).

Evidence for inhibitory processes involved in directional selectivity has also been found in cat cortical cells. Hubel and Wiesel (1959, 1962) suggested that directional selectivity of simple cells could be explained by summation of responses from adjacent ON and OFF regions of the cell, where an ON region responded to the luminance increment and an OFF region responded to a luminance decrement. Such ON and OFF responses have been demonstrated in two classes of retinal ganglion cells (Kuffler, 1953) that converge at the simple cells in cortex (Schiller, 1982). However, a number of studies have later rejected the hypothesis that the temporal coincidence of these ON and OFF responses can explain directional selectivity (Emerson and Gerstein, 1977; Ganz, 1984; Goodwin *et al.*, 1975, Heggelund, 1984). For example, Goodwin *et al.* (1975) studied simple cells in cat striate cortex which showed ON and OFF receptive field regions for both stationary flashed stimuli and moving edges. The majority of these cells could not be correlated with the spatial arrangement

of their receptive fields and were independent of the width of the moving bar used as a stimulus, invalidating the spatial summation hypothesis. Like Barlow and Levick (1965), they concluded that inhibition in the non-preferred direction was primarily responsible for the direction selectivity.

Both Barlow and Levick (1965) and Goodwin *et al.* (1975) found directional selectivity to be contained within small sub-units of observed cell receptive fields. For example, Goodwin *et al.* reported that one cell was divided into 22 sub-units, each of which demonstrated the same directional selectivity of the cell as a whole. In fact, Goodwin *et al.* were unable to find non-directionally-selective subregions within the receptive field down to a displacement threshold of 1 minute of arc.

In summary, early directional selectivity appears to be based on inhibitory veto processing, as opposed to facilitatory or correlational operations. These processes seem to operate at a small scale in comparison to the size of individual receptive fields of directionally selective cells in either rabbit retina or cat cortical cells.

At what processing stage does such a directional veto mechanism operate? Chey *et al.* (1997) suggested why it occurs as part of transient cell processing prior to the short-range filter. It can here set up local directional estimates at directional transient cells before evidence for these estimates is spatially accumulated across a moving trajectory by directionally-sensitive short-range filter cells. We suggest herein that these directional transient cells operate upon the outputs of the lightening and darkening channels before the outputs are, in turn, processed by the short-range filters.

Competition across direction within each channel then acts to enhance the outputs of directional short-range filter cells that have few directional competitors at a given position, while attenuating outputs of directional cells with many directional competitors, without disrupting speed estimates. A divisive, or shunting, competition across direction and scale accomplishes this by computing the ratio of competing activities (Grossberg, 1980, 1983). Unambiguous feature tracking signals are hereby boosted relative to ambiguous signals, and ambiguous signals are biased towards a direction of motion that is perpendicular to a line's orientation.

The long-range motion filter then pools signals from multiple orientations and contrast polarities in a prescribed direction-of-motion. It is the model processing stage that generates cells that are truly directional selective, and is proposed to occur in cortical area MT where cells with similar receptive field properties have been reported (Albright *et al.*, 1984; Maunsell and van Essen, 1983a; Newsome *et al.*, 1983; Zeki, 1974). This processing stage also pools signals from both eyes (Grossberg and Rudd, 1992). It hereby achieves the depth selectivity of MT cells (Bradley *et al.*, 1995; Maunsell and van Essen, 1983b) and helps to explain how long-range apparent motion can occur with dichoptically presented stimuli (Gengerelli, 1948; Spigel, 1968).

The directional grouping and attentional priming stage of Figure 5 was not simulated because its role is not important in processing the displays that are being simulated. Extensive simulations showed that the ordering of directional transient cells, competition, and short-range filtering could be varied without disrupting the main qualitative results. They are quite robust.

8 Mathematical Description of the Motion BCS

Equations (1) - (16) above, through the lightening and darkening cells, are the same as before. The subsequent processing stages, summarized in Figure 15, are defined as follows. Directional transient cells are derived through the intervention of directional interneurons (Chey *et al.*, 1997).

FIGURE 15

Directional Interneurons

Directional interneuron activity ξ_i^v time-averages the lightening (or darkening) cell output:

$$\frac{d\xi_i^v}{dt} = -\xi_i^v + [w_i^v - , w]^+, \quad (21)$$

where $v = L$ for lightening and $v = D$ for darkening.

Directional Transient Cells - Left Direction

The directional transient cell activity x_i^{vL} that prefers left direction-of-motion receives excitatory input from the lightening (or darkening) cell which is vetoed by the directional interneuron activity offset by one unit in the left direction:

$$\frac{dx_i^{vL}}{dt} = -A_5 x_i^{vL} + B_5 [w_i^v - , w]^+ - C_5 [\xi_{i-1}^v]^+, \quad (22)$$

Directional Transient Cells - Right Direction

Similarly, the directional transient cell activity x_i^{vR} that prefers right direction-of-motion receives excitatory input from the lightening (or darkening) cell which is vetoed by the directional interneuron activity offset by one unit in the right direction:

$$\frac{dx_i^{vR}}{dt} = -A_5 x_i^{vR} + B_5 [w_i^v - , w]^+ - C_5 [\xi_{i+1}^v]^+, \quad (23)$$

Short-Range Spatial Filters

Short-range spatial filter activity $y_i^{v\nu}$ performs space and time averaging of directional cell responses. A Gaussian kernel P_{ji} ensures that the contributions from adjacent neighbors are larger than the contributions from more distant neighbors:

$$\frac{dy_i^{v\nu}}{dt} = -A_6 y_i^{v\nu} + (B_6 - y_i^{v\nu}) \sum_j P_{ji} [x_j^{v\nu}]^+, \quad (24)$$

and

$$P_{ji} = \frac{\alpha_y}{\sigma_y \sqrt{2\pi}} \exp \left[\frac{-(j-i)^2}{2\sigma_y^2} \right], \quad (25)$$

where $v\nu = LL$ for lightening-left, $v\nu = DL$ for darkening-left, $v\nu = LR$ for lightening-right, and $v\nu = DR$ for darkening-right. These activities are half-wave rectified to generate output signals

$$Y_i^{v\nu} = [y_i^{v\nu} - , y]^+. \quad (26)$$

Directional Competition

Competition occurs between left and right directional cells that obey membrane equations. This competition computes the ratio of competing activities in lightening and darkening channels. For simplicity, it is assumed that the competition acts quickly. Its activities are thus computed at steady state and half-wave rectified to yield the output signals:

$$\Upsilon_i^{LL} = \frac{[Y_i^{LL} - Y_i^{LR}]^+}{\beta_Y + Y_i^{LL} + Y_i^{LR}}, \quad (27)$$

$$\Upsilon_i^{DL} = \frac{[Y_i^{DL} - Y_i^{DR}]^+}{\beta_Y + Y_i^{DL} + Y_i^{DR}}, \quad (28)$$

$$\Upsilon_i^{LR} = \frac{[Y_i^{LR} - Y_i^{LL}]^+}{\beta_Y + Y_i^{LL} + Y_i^{LR}}, \quad (29)$$

and

$$\Upsilon_i^{DR} = \frac{[Y_i^{DR} - Y_i^{DL}]^+}{\beta_Y + Y_i^{DL} + Y_i^{DR}}, \quad (30)$$

Long-Range Spatial Filters

Long-range spatial filter activity z_i^ν separately pools the outputs from the lightening and darkening channels in both the left and right directions. A Gaussian kernel q_{ji} ensures that the contributions from adjacent neighbors are larger than the contributions from more distant neighbors, as in:

$$\frac{dz_i^\nu}{dt} = -A_7 z_i^\nu + (B_7 - z_i^\nu) \sum_j q_{ji} (\Upsilon_j^{L\nu} + \Upsilon_j^{D\nu}), \quad (31)$$

where $\nu = L$ for left, $\nu = R$ for right, and

$$q_{ji} = \frac{\alpha_z}{\sigma_z \sqrt{2\pi}} \exp \left[\frac{-(j-i)^2}{2\sigma_z^2} \right]. \quad (32)$$

The output is thresholded at z and half-wave rectified to generate output signals

$$Z_i^\nu = [z_i^\nu - z]^+. \quad (33)$$

9 Motion BCS Simulations

The same stimuli as in Figures 10-14 are now simulated in Figures 16-20. The effects of each processing stage are shown here. The simulation parameters for earlier stages were as before and for motion BCS were: $A_5 = 10.0$, $B_5 = 10.0$, $C_5 = 50.0$, $w = 0.1$, $A_6 = 1.0$, $B_6 = 1.0$, $\alpha_y = 15.0$, $\sigma_y = 1.5$, $y = 0.1$, $\beta_Y = 0.0001$, $A_7 = 1.0$, $B_7 = 1.0$, $\alpha_z = 15.0$, $\sigma_z = 5.0$ and $z = 0.6$. In Figures 16-20, activities of left and right directional transient cells for the lightening channel (x_i^{LL} and x_i^{LR}) are shown in (a) and (b), respectively, and for the darkening channel (x_i^{DL} and x_i^{DR}) in (c) and (d). The left and right short-range filter cells for the lightening channel (y_i^{LL} and y_i^{LR}) are shown in (e) and (f), respectively, and for the darkening cells (y_i^{DL} and y_i^{DR}) in (g) and (h). The pooled activity of left direction cells after competition ($\Upsilon_i^{LL} + \Upsilon_i^{DL}$) is plotted in (i) and right direction ($\Upsilon_i^{LR} + \Upsilon_i^{DR}$) in (j). The left and right long-range filter output (Z_i^L and Z_i^R) are plotted in (k) and (l), respectively. The main thing to note is whether energy is concentrated in the output of the left, Z_i^L , or right, Z_i^R , long-range filter.

FIGURES 16 - 20

10 A Single Processing Stream for First-Order and Second-Order Motion?

A number of researchers have suggested that first-order and second-order motion stimuli are processed by independent pathways. The psychophysical evidence for these arguments include scale-dependent direction of motion reversal in the σ display (Chubb and Sperling, 1989), first-order

and second-order motion percepts in a multi-frame motion sequence (Ledgeway and Smith, 1994; Mather and West, 1993), different temporal sensitivities for first-order and second-order motion stimuli (Derrington, 1994; Derrington and Badcock, 1985; Derrington *et al.*, 1993; Holliday and Anderson, 1994; Smith and Ledgeway, 1997a), first-order but not second-order motion detection at the absolute detection threshold (Smith and Ledgeway, 1997b; Watson *et al.*, 1980), first-order but not second-order motion activation of the optokinetic eye movement system (Harris and Smith, 1992), and a small phase-dependence during direction judgment experiments on superimposed Fourier and non-Fourier stimuli (Lu and Sperling, 1995a). Nevertheless, psychophysical experiments on adaptation and sensitivity studies provide evidence that first-order and second-order stimuli are processed by single processing stream (Ledgeway and Smith, 1995; Turano, 1991; Turano and Pantle, 1989). Taub *et al.* (1997) have conducted experiments with varying degrees of nonlinearity in non-Fourier motion stimuli and compared velocity discrimination judgments for first-order and second-order stimuli. Their findings are consistent with a single processing stream.

There is also neurophysiological evidence in support of single processing stream. Albright (1992) found that 87% of the cells in area MT that respond to the first-order stimuli also respond to the second-order stimuli. The model proposed herein utilizes a single processing stream to process both first-order and second-order motion stimuli. The model hereby clarifies why cells in area MT can respond to both first-order and second-order motion stimuli (Albright, 1992). In particular, both first-order and second-order motion stimuli are processed monocularly (Lu and Sperling, 1995a), whereas cells in MT are already binocularly sensitive. The model mechanisms that process these stimuli occur prior to the binocular fusion of information that is proposed to occur at a long-range spatial filter that converges on model MT cells.

Johnston and Clifford (1995) have convincingly argued that a single processing stream is sufficient to simulate a number of motion percepts that others have used to argue for multiple processing channels. Their model is based upon formal Taylor series expansions of image brightness around a point of interest. These expansions are used in conjunction with integral operations to provide “a least squares estimate of image speed based on measures of how the image brightness and its derivatives are changing with respect to space and time” (p 1123). The present approach directly develops a neural model of the magnocellular brain mechanisms that subserve motion perception. It is not yet clear how the two approaches can be linked.

The Motion BCS model that is developed herein has shown that various second-order properties that have been attributed to a second processing stream may be due to interactions between ON and OFF cells within a single processing stream. We wish to emphasize the logical force of this demonstration. It suggests that various earlier arguments about the existence of different first-order and second-order streams are logical nonsequiturs. Different motion *properties* do not imply different motion *processes*. Given that our analysis is also linked to known thalamocortical ON and OFF cell properties, which have not been incorporated into earlier models, the question of whether separate processing streams process first-order and second-order motion needs to be approached with renewed caution.

To make these demonstrations, the present study proposes a model of lightening and darkening cells and refines the first few processing stages of a neural architecture, called the Motion BCS, that has previously been used to simulate many other data about visual motion perception (Chey *et al.*, 1994, 1997; Grossberg and Mingolla, 1993; Grossberg and Rudd, 1989, 1992). For example, by the time signals in the motion BCS are processed by the short-range filters, they can do pre-attentive feature tracking. Subsequent stages of the motion BCS model include a long-range spatial filter at which multiple orientations, contrast polarities and inputs from both eyes converge to achieve true directionally selective cells. These cells feed on attentive directional grouping stage that uses the directional feature tracking signals to achieve global motion capture and attentive grouping of

motion signals (Chey *et al.*, 1997). These attentive mechanisms were not needed to simulate the data considered herein.

We wish to emphasize that, although only one motion processing stream is needed to explain the first-order and second-order motion percepts that are analyzed herein, these results are not meant to imply that multiple processing streams do not operate in other situations, or that interactions between these streams cannot influence motion percepts. On the other hand, these other processing streams are often devoted to the processing of stimulus *form*, not motion. Various modifications of motion stimuli can cause different combinations of motion, as well as form, mechanisms to be engaged. Thus the question of whether multiple streams influence first-order and second-order motion percepts needs to carefully address the functional role of these streams from the broader perspective of visual perception, not only their possible immediate influence on a relatively narrow set of motion percepts.

For example, Lu and Sperling (1995a, 1995b) have provided experimental evidence for a third-order motion system that requires feature tracking. A number of other experimentalists have also emphasized the role of feature tracking signals (Bowns, 1996; Castet *et al.*, 1993; Wallach, 1976). Lu and Sperling (1995a) noted that this system is slower than the first-order and second-order motion systems, operates interocularly as well as monocularly, requires much more stimulus contrast than first-order and second-order stimuli, and requires both bottom-up processing including interactions between form and motion pathways and top-down attentional priming. We have elsewhere argued that such “third-order motion” percepts are, in reality, form-motion interactions that help to join complementary processing properties of the form and motion processing streams (Baloch and Grossberg, 1997; Francis and Grossberg, 1996a; Grossberg, 1991). Namely, the form stream uses precise *orientational* estimates to form emergent 3-D boundary and surface representations at precisely calibrated depths from an observer, but exhibits poor *directional* tracking properties. The motion system sacrifices precise orientational and stereo estimates to achieve good directional tracking. The form-motion interaction between cortical areas V2 and MT is proposed to help the motion system better track emergent forms in depth. Our model of this form-motion interaction has been used to simulate motion percepts that occur when two spatially overlapping shapes that are presented discretely in time appear to transform smoothly from one shape to another, such as the line motion illusion (Hikosaka *et al.*, 1993a, 1993b), motion induction (Faubert and von Grünau, 1992, 1995; von Grünau and Faubert, 1994) and transformational apparent motion (Tse *et al.*, 1996).

Thus the fact that only one stream is needed to process some first-order and second-order motion percepts is not meant to imply that more than one stream may not be engaged to process other motion percepts. On the other hand, these streams may not all be *motion* processing streams, and their interaction may be aimed at functionally compensating for complementary weaknesses of the form and motion processing streams.

11 How Multiple-Scale ON and OFF Processing Help to Explain Other Second-Order Motion Data

The cross-talk between ON and OFF cells in lightening and darkening channels, competition between directional cells, and short-range and long-range filter stages can be used to explain other data that have previously been proposed to imply the existence of separate processing streams for processing Fourier and non-Fourier stimuli (Chubb and Sperling, 1989; Ledgeway and Smith, 1994; Lu and Sperling, 1995a). Arguments for separate streams are often based upon an elaborated Reichardt model as the main processing stage for motion processing. For example, Lu and Sperling

(1995a, 1995b) have reported data to identify three separate processing streams (or systems) for motion processing: a first-order system for Fourier stimuli such as moving luminance modulations, a second-order system for non-Fourier stimuli such as moving texture-contrast modulations, and a third-order system that tracks features. The first-order and second-order streams are identified to be monocular and bottom-up while the third-order stream is both monocular and binocular and can be influenced by top-down attentional priming (Lu and Sperling, 1995a).

Lu and Sperling (1995a) mainly dealt with four kinds of motion stimuli: luminance stimuli that are first-order; texture stimuli that are second-order; depth stimuli that are presented interocularly (stereo); and motion-motion modulated stimuli that require feature tracking. In some experiments, these motion stimuli were modulated with a pedestal (a stationary sine wave). The main purpose of modulating luminance and texture stimuli with a pedestal was to remove features, thereby preventing these stimuli from being processed by a third-order tracking system. Their temporal-frequency sensitivity tests (Experiment 1) for these four kinds of stimuli show that both first-order and second-order systems are fast while the third-order system for depth and motion modulated stimuli is comparatively slow. The pedestaled stimuli tests (Experiments 2 and 3) further confirmed the results from Experiment 1. These data agree with a single processing stream model for first-order and second-order motion stimuli such as ours. The third-order system is slow as it tracks features in the stimuli and requires interstream interactions such as between V2 and MT (Baloch and Grossberg, 1997; Chey *et al.*, 1997; Francis and Grossberg, 1996a; Grossberg, 1991).

The relative phase dependence test (Experiment 4) examined directional judgments for superimposed Fourier and non-Fourier gratings as a function of relative spatial phase. If pedestaled luminance modulated and texture modulated stimuli moved in opposite directions (pedestaled motion transparency test), then no motion was observed. This result is consistent with a single stream system for processing of both first-order and second-order stimuli. In particular, it agrees with our model's proposal that competition between opponent direction cells occurs no later than the long-range filter stage. This model property is supported by neurophysiological data suggesting the opponent motion direction cells compete no later than cortical area MT (Albright *et al.*, 1984).

The main evidence suggesting that first-order and second-order systems utilize separate channels came from the tests in which luminance-modulated and texture-modulated stimuli were presented together drifting in the same direction (Experiment 4). In one case, both stimuli had the same spatial frequency and temporal frequency. This case tested the hypothesis that both texture-modulated and luminance-modulated stimuli are computed by a first-order system with linear processing and half-wave rectification. In the other case, the parameters of the contrast-modulated stimulus remained the same, but the spatial and the temporal frequency of the luminance-modulated stimulus were only half that of the first case. This case tested the hypothesis that both stimuli are processed by a second-order system with nonlinear processing and full-wave rectification. The full-wave rectification would double the frequency of the luminance-modulated stimulus, thereby matching it to the texture-modulated stimulus. Interestingly, the stimuli in these cases were not modulated by the pedestal. The percent of correct direction-of-movement judgments of both stimuli together was determined for eight relative spatial phases. The test hypothesis was that, if both stimuli were processed by a single channel before the motion is computed, then their combined magnitude would depend on their relative phase. For example, stimuli with the same frequency but with opposite phase would cancel each other. Therefore, absence of any phase dependence would mean that these stimuli are processed by separate systems. To a good first approximation, they found no relative phase dependence in the first case, and a slight phase dependence in the second case, which they attributed to sine-wave luminance-modulated stimuli leaking into the second-order system. They concluded that separate channels compute motion for luminance-modulated and texture-modulated stimuli and these two streams are combined at a higher-level to arrive at a single-valued represen-

tation of motion direction.

These conclusions are generally true for motion energy models based on elaborated Reichardt motion detectors, including those where the spatio-temporal signal is preprocessed by nonlinear filters followed by full-wave rectification. Taub *et al.* (1997) nonetheless were able to successfully challenge the Lu and Sperling (1995a) conclusion that their data imply that separate first-order and second-order processing streams necessarily exist. Taub *et al.* (1997) proposed a single processing stream model and used psychophysical experiments to estimate a functional form of its nonlinear processing to handle a large number of data, including the phase-dependence data of Lu and Sperling (1995a). They proposed the nonlinear processing to be *compressive* and *asymmetric*; *i.e.*, it treats positive and negative contrasts differently. They also observed that if a large spatial filter was used before the nonlinear transformation, then the function appears to be more nearly linear. Taub *et al.* (1997) also noted that Lu and Sperling (1995a) “measured the fraction of correct judgments, not motion energy, and the measurements were made in the range of 86%-90%-correct performance. In this range, the fraction correct is likely to be a compressive function of motion energy”, thus further reducing the likelihood of observing the phase dependence psychophysically.

A key difference between our model and the elaborated Reichardt model, among others, is that our model predicts the existence of lightening and darkening cells that transiently respond to local increases in brightness or darkness, respectively. We now note similarities between the way lightening and darkening cell process their inputs and the nonlinear transformation proposed by Taub *et al.* (1997). In our model, on-center off-surround processing in the lightening and darkening channels allows the information from both ON and OFF cells in the center to interact with ON and OFF cells in the surround. The shunting competition causes spatial distribution of activity. The motion directions are computed independently in the lightening and darkening channels before they are pooled at the long-range filter stage. The inhibitory inputs are shunted at a value that is less than the excitatory inputs, thereby leading to asymmetric processing of excitatory and inhibitory signals. Also, the surround Gaussian filter is proposed to be spatially large, which allows the inputs to be presented linearly. Mathematical details and parameter values are given below. Thus, the neural architecture of our model is supported by the experimental findings of Taub *et al.* (1997). They do not identify the existence of lightening and darkening cells. Nor do they propose the nonlinearity to occur at ON and OFF cells stage. On the other hand, Taub *et al.* (1996) do note that the nonlinear transformation could happen in two pathways that are mirror-images of each other. The lightening and darkening cells of our model does support (approximately) mirror-image processing due to ON and OFF processing.

Taub *et al.* (1997) also noted that “the Fourier and non-Fourier gratings were presented in alternate rows of raster display; this spatial separation would reduce interactions among the two kinds of gratings simply because they might tend to stimulate separate detectors”. In our model, the anisotropic spatial filters tuned in the direction-of-drift, which was the same for both the luminance-modulated and texture-modulated stimuli in the phase-dependence test, can detect the motion much more independently than could isotropic spatial filters, thus also helping to keep the phase-dependence small. All of these factors conspire to generate small phase dependence without necessarily processing first-order and second-order stimuli independently. Finally, if luminance and texture stimuli have different spatial and temporal frequencies, the different selectivities of multiple spatial scales and interscale competition would allow the system to compute the correct direction and speed of motion (Chey *et al.*, 1997).

Lu and Sperling (1995a) also performed interocular tests (Experiments 5 and 6) and observed that first-order and second order motion were monocular processes, while depth and modulated motion stimuli required feature tracking and were not effected by interocular presentations. If this is true, then first-order and second-order processing must be handled before MT. This is consistent

with our model's hypothesis that lightening and darkening cells exist at an early stage before the long-range filter that converges on model MT cells. As per feature tracking to be both monocular and binocular, this is obvious in our model because feature tracking occurs before the long-range filter stage at which binocular fusion occurs.

A number of studies have used interleaved first-order and second-order motion stimuli (Ledgeway and Smith, 1994; Mather and West, 1993). In these experiments, one frame of first-order stimulus is replaced by a frame of a second-order stimulus and the subjects are asked to report the direction of perceived motion. For example, Ledgeway and Smith (1994) used luminance-modulated bit noise as first-order stimulus and contrast-modulated bit noise as a second-order stimulus. They varied the spatial displacement as a fraction of spatial period of the modulating sine wave between frames and subjects were asked to report the direction of perceived motion. The results became ambiguous when the spatial shift between the frames was equal to quarter of the wave-length; i.e., the effective phase shift between the stimuli of same type equaled half of wavelength. They argued that, if first-order and second-order stimuli were processed by separate channels, the effective phase shift of 0.5 wavelength within the independent channels would cause ambiguous motion direction perception.

The authors noted that, because the two stimuli are not alike, their difference could result in motion cancellation in single channel models that use some sort of nonlinearity at an early stage of processing. Such a nonlinearity, which exists in essentially all motion models today, could also explain their data without requiring the assumption of two separate channels. In order to eliminate this possibility, they tried to make the two stimuli close in appearance and alter their modulation depth and intensity. However, these modifications influence the global nature of the second-order signal, making it similar to the first-order signal (e.g., they become similar in their global visibility).

The model described here is capable of capturing these differences at spatially local levels. In particular, these local differences in image intensities, though small, are captured by the lightening and darkening cells, which process information from surround as well as center, and can capture the transient effects of relative changes in the local brightness levels (see Section 4). This is qualitatively similar to our explanation of the , display when seen from near and far (Chubb and Sperling, 1989). In those experiments, the grating shift was accompanied by an intensity contrast reversal. In interleaved experiments, the phase shift between frames is accompanied by replacing a contrast-modulated stimulus with a luminance-modulated stimulus, or vice versa. The local changes between these two types of stimuli are captured by the lightening and darkening cells. When the shift equals a quarter of the wavelength, as in the , display, it generates equivalent signals for both directions at the long-range filter stage, where the global motion signals are processed. Similarly, Mather and West (1993) used intensity kinematograms as first-order stimuli and texture kinematograms as second-order stimuli in their version of interleaved experiments. The same arguments apply in those cases as well.

A number of studies have reported that temporal sensitivity is worse for second-order motion than first-order motion and therefore suggest that two processing streams are necessary (Derrington, 1994; Derrington and Badcock, 1985; Derrington *et al*, 1993; Holliday and Anderson, 1994; Smith and Ledgeway, 1997a). However, selective adaptation studies strongly suggest a single-processing channel (Turano, 1991; Turano and Pantle, 1989). Turano (1991) reported that "the results support the view that signals generated from luminance-domain stimuli and from contrast-domain stimuli are processed by a common motion mechanism" (p 455). Holliday and Anderson (1994) also found that adaptation tuning curves for first-order and second-order stimuli are similar at high temporal frequencies (more than 4 Hz) with maximal post adaptation near 12 Hz. For example, they concluded "the results are consistent with the hypothesis that fast second-order motion is detected by Fourier-type mechanisms, preceded by a nonlinearity, and slow second-order motion is

detected by a process involving a comparison of local luminance features” (p 165). Lu and Sperling (1995a) have classified these data into three types of motion. They noted that the first-order and second-order are fast, while the third-order system is slow. As discussed above, these so called third-order stimuli may require more processing time because the percept requires form-motion interstream interactions such as between V2 and MT (Baloch and Grossberg, 1997; Chey et al., 1997; Francis and Grossberg, 1996a; Grossberg, 1991).

Other experiments such as those by Chubb and Sperling (1988) on drift-balanced stimuli and Harris and Smith (1992) on OKN were discussed above in the light of model simulations.

12 Summary

A neural model of motion perception based upon magnocellular dynamics is developed to provide a unified explanation of key first-order and second-order motion percepts. This motion Boundary Contour System model does not invoke parallel Fourier and non-Fourier pathways. A single processing stream is sufficient if opponent ON and OFF channels exist that are capable of antagonistic rebound and that combine their results at suitably defined lightening and darkening cells whose outputs are pooled to derive a motion percept. The model thus does not require separate channels, with one detecting only first-order motion while the other detects only second-order motion. This result is in agreement with the data of Albright (1992) showing that most of the cells at area MT are selective to both first-order and second-order motion stimuli. In addition:

1. The model explains why monkeys with lesions of the parvocellular but not magnocellular layers of the lateral geniculate nucleus are capable of detecting the correct direction of first-order motion (Schiller *et al.*, 1990). The model also predicts that such selective blocking would not effect the detection of the correct direction of second-order motion.
2. The model computes the correct direction of motion of a contrast-reversing noise field as an example of a second-order motion stimulus (Chubb and Sperling, 1988).
3. The model computes the opposite directions of motion for the , display when it is seen from near and far (Chubb and Sperling, 1989).
4. The model computes the correct direction of motion and tracks both edges of a moving bright bar moving on a background. The model tracks only the trailing edge of a moving bright bar if its ON channel is blocked (Schiller, 1992). The model predicts that blocking OFF bipolar cells would make direction-selective cortical cells insensitive to the trailing edge. Moreover, if a dark moving bar was used instead of a bright bar, then blocking ON (OFF) cells should make the direction-selective cells insensitive to the trailing (leading) edge.
5. The model qualitatively explains the data of Lu and Sperling (1995a) on how directional judgments depend upon relative spatial phase, or spatial and temporal frequency.
6. The model predicts the existence of two testable classes of cells in the magnocellular processing stream, lightening and darkening cells, that transiently respond to local increase in brightness and darkness, respectively, no later than cortical area MT.

The present version of the motion Boundary Contour System incorporates transient OFF cell rebounds and lightening and darkening cells into a neural theory of motion perception that has earlier simulated a wide variety of data about short-range and long-range apparent motion (Grossberg and Rudd, 1989, 1992), percepts of object speed and direction (Chey, *et al.*, 1994, 1997), and

form-motion interactions (Baloch and Grossberg, 1996, 1997; Francis and Grossberg, 1996a). The present results hereby show how a small set of properly configured neural mechanisms can be used to unify the explanation of a diverse set of motion data.

References

- Adelson, E.H., and Bergen, J.R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, **2**(2), 284-299.
- Albright, T.D. (1984). Direction and orientation selectivity of neurons in visual area MT of the macaque. *Journal of Neurophysiology*, **52**, 1106-1131.
- Albright, T.D. (1992). Form-cue invariant motion processing in primate visual cortex. *Science*, **255**, 1141-1143.
- 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.
- Anstis, S.M. and Rogers, B.J. (1975). Illusory reversal of visual depth and movement during changes in contrast. *Vision Research*, **15**, 957-961.
- Ariel, M. and Daw, N.W. (1982). Pharmacological analysis of directionally sensitive rabbit retinal ganglion cells. *Journal of Physiology*, **324**, 161-185.
- Baker, C.L., and Braddick, O.J. (1982). The basis of area and dot number effects in random dot motion perception. *Vision Research*, **22**, 1253-1259.
- Baloch, A. and Grossberg, S. (1996). Neural dynamics of morphing motion. *Investigative Ophthalmology and Visual Science*, **37**, 3419.
- Baloch, A. and Grossberg, S. (1997). A neural model of high-level motion processing: Line motion and formotion dynamics. *Vision Research*, **37**, 3037-3059.
- Barlow, H.B. and Levick, W.R. (1965). The mechanism of directionally selective units in rabbit's retina. *Journal of Physiology*, **178**, 477-504.
- Bartley, S.H. (1941). **Vision, a study of its basis**. New York: Van Nostrand Reinhold.
- Bischof, W.F., and Di Lollo, V. (1990). Perception of directional sampled motion in relation to displacement and spatial frequency: evidence for a unitary motion system. *Vision Research*, **30**(9), 1341-1362.
- Bowns, L. (1996). Evidence for a feature tracking explanation of why Type II plaids move in the vector sum direction at short durations. *Vision Research*, **36**, 3685-3694.
- Braddick, O.J. (1974). A short-range process in apparent motion. *Vision Research*, **14**, 519-527.
- Bradley, D.C., Qian, N., and Anderson, R.A. (1995). Integration of motion and stereopsis in middle temporal cortical area of macaques. *Nature*, **373**, 609-611.
- Carpenter, G.A. and Grossberg, S. (1981). Adaptation and transmitter gating in vertebrate photoreceptors. *Journal of Theoretical Neurobiology*, **1**, 1-42. Reprinted in S. Grossberg (Ed.), **The adaptive brain, Volume I**. Amsterdam: Elsevier/North-Holland.
- Castet, E., Lorenceau, J., Shiffrar, M. and Bonnet, C. (1993). Perceived speed of moving lines depends on orientation, length, speed, and luminance. *Vision Research*, **33**(14), 1921-1936.
- Cavanagh, P., and Mather, G. (1989). Motion: The long and short of it. *Spatial Vision*, **4**, 103-129.
- Cavanagh, P., Arguin, M. and von Grünau, M. (1989). Interattribute apparent motion. *Vision Research*, **29**, 1197-1204.
- Chang, J.J., and Julesz, B. (1983a). Displacement limits, directional anisotropy and direction versus form discrimination in random-dot kinematograms. *Vision Research*, **23**(6), 639-646.
- Chang, J.J., and Julesz, B. (1983b). Displacement limits for spatial frequency filtered random-dot kinematograms in apparent motion. *Vision Research*, **23**(12), 1379-1385.
- Chey, J., Grossberg, S., and Mingolla, E. (1994). Neural dynamics of motion processing and speed discrimination. Technical Report CAS/CNS-TR-94-030. Boston, MA: Boston University.

- Chey, J., Grossberg, S., and Mingolla, E. (1997). Neural dynamics of motion grouping: From aperture ambiguity to object speed and direction. *Journal of Optical Society of America A*, **14**(10), 2570-2594.
- Chubb, C., and Sperling, G. (1988). Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America A*, **5**(11), 1986-2007.
- Chubb, C., and Sperling, G. (1989). Two motion perception mechanisms revealed through distance-driven reversal of apparent motion. *Proceedings of the National Academy of Sciences - USA*, **86**, 2985-2989.
- Clelland, B.G., Dubin, M.W., and Levick, W.R. (1971). Sustained and transient neurons in the cat's retina and lateral geniculate nucleus. *Journal of Physiology*, **217**, 473-496.
- Derrington, A. (1994). Analysis of the motion of contrast-modulated patterns. *Investigative Ophthalmology and Visual Science*, **35**, 1406.
- Derrington, A.M. and Badcock, D.R. (1985). Separate detectors for simple and complex cells. *Vision Research*, **25**, 1869-1878.
- Derrington, A.M., Badcock, D.R., and Henning, G.B. (1993). Discriminating the detection of second-order motion at short stimulus durations. *Vision Research*, **37**, 1785-1794.
- de Silva, H.R. (1926). An experimental investigation of the determinants of apparent visual movement. *American Journal of Psychology*, **37**, 469-501.
- Emerson, R.C. and Gerstein, G.L. (1977). Simple striate neurons in the cat: II. Mechanisms underlying directional asymmetry and directional selectivity. *Journal of Neurophysiology*, **40**(1), 136-155.
- Enroth-Cugell, C., and Robson, J.G. (1966). The contrast sensitivity of retinal ganglion cells of the cat. *Journal of Physiology*, **187**, 517-552.
- Faubert, J. and von Grünau, M. (1992). The extent of split attention and attribute priming in motion induction. *Perception*, **21**, 105b.
- Faubert, J. and von Grünau, M. (1995). The influence of two spatially distinct primers and attribute priming on motion induction. *Vision Research*, **35**(22), 3119-3130.
- Fennema, C.L., and Thompson, W.B. (1979). Velocity determination in scenes containing several moving objects. *Computer Graphics and Image Processing*, **9**, 301-315.
- Ferrera, V.P., Nealey, T.A., and Maunsell, J.H.R. (1992). Mixed parvocellular and magnocellular geniculate signals in visual area V4. *Nature*, **358**(27), 756-758.
- Ferster, D. (1988). Spatially opponent excitation and inhibition in simple cells of the cat visual cortex. *Journal of Neuroscience*, **8**, 1172-1180.
- Francis, G. and Grossberg, S. (1996a). Cortical dynamics of form and motion integration: Persistence, apparent motion and illusory contours. *Vision Research*, **36**(1), 149-173.
- Francis, G. and Grossberg, S. (1996b). Cortical dynamics of boundary segmentation and reset: persistence, afterimages, and residual traces. *Perception*, **25**, 543-567.
- Francis, G., Grossberg, S., and Mingolla, E. (1994). Cortical dynamics of feature binding and reset: control of visual persistence. *Vision Research*, **34**, 1089-1104.
- Ganz, L. (1984). Visual cortical mechanisms responsible for direction selectivity. *Vision Research*, **24**, 3-11.
- Gellatly, A. and Blurton, A. (1996). What are the mechanisms of rivalrous first-order and second-order motions? *Perception*, **25**, **Supplement**, 8-9.
- Gengerelli, J.A. (1948). Apparent movement in relation to homonymous and heteronymous stimulation of the cerebral hemispheres. *Journal of Experimental Psychology*, **38**, 592-599.

- Gaudio, P. (1994). Simulations of X and Y retinal ganglion cell behavior with a nonlinear push-pull model of spatiotemporal retinal processing. *Vision Research*, **34**, 1767-1784.
- Giaschi, D., and Anstis, S. (1989). The less you see it, the faster it moves: Shortening the "on-time" speeds up the apparent motion. *Vision Research*, **29**, 335-347.
- Goodwin, A.W., Henry, G.H., and Bishop, P.O. (1975). Direction selectivity of simple striate cells: Properties and mechanisms. *Journal of Neurophysiology*, **38**, 1500-1523.
- Gove, A., Grossberg, S. and Mingolla, E. (1995). Brightness perception, illusory contours, and corticogeniculate feedback. *Visual Neuroscience*, **12**, 1027-1052.
- Grossberg, S. (1972). A neural theory of punishment and avoidance, II: Quantitative theory. *Mathematical Biosciences*, **15**, 253-285.
- Grossberg, S. (1976). Adaptive pattern classification and universal recoding: II. Feedback, expectation, olfaction, illusions. *Biological Cybernetics*, **23**, 187-202.
- Grossberg, S. (1980). How does a brain build cognitive code? *Psychological Review*, **87**, 1-51.
- Grossberg, S. (1983). The quantized geometry of visual space: The coherent computation of depth, form and lightness. *The Behavioral and Brain Sciences*, **6**, 625-657.
- Grossberg, S. (1991). Why do parallel cortical systems exist for the perception static form and moving form? *Perception and Psychophysics*, **49**, 117-141.
- Grossberg, S., and Mingolla, E. (1993). Neural dynamics of motion perception: Direction fields, apertures, and resonant grouping. *Perception and Psychophysics*, **53**(3), 243-278.
- Grossberg, S., and Rudd, M.E. (1989). A neural architecture for visual motion perception: Group and element apparent motion. *Neural Networks*, **2**, 421-450.
- Grossberg, S., and Rudd, M.E. (1992). Cortical dynamics of visual motion perception: Short- and long-range motion. *Psychological Review*, **99**(1), 78-121.
- Hadani, B., Ishai, G., and Gur, M. (1980). Visual stability and space perception in monocular vision: Mathematical model. *Journal of the Optical Society of America*, **70**(1), 60-65.
- Harris, L.R. and Smith, A.T. (1992). Motion defined exclusively by second-order characteristics does not evoke optokinetic nystagmus. *Visual Neuroscience*, **9**, 565-570.
- Heggelund, P. (1984). Direction asymmetry by moving stimuli and static receptive field plots for simple cells in cat striate cortex. *Vision Research*, **24**, 1316-??.
- Hikosaka, O., Miyauchi, S., and Shimojo, S. (1993a). Focal visual attention produces illusory temporal order and motion sensation. *Vision Research*, **33**(9), 1219-1240.
- Hikosaka, O., Miyauchi, S., and Shimojo, S. (1993b). Voluntary and stimulus-induced attention detected as motion sensation. *Perception*, **22**, 517-526.
- Holliday, I.E. and Anderson, S.J. (1994). Different processes underlie the detection of second-order motion at low and high temporal frequencies. *Proceeding of Royal Society of London B*, **257**, 165-173.
- Horn, B.K.P., and Schunck, B.G. (1980). Determining optical flow. MIT A.I. Memo 572. Cambridge, MA.
- Hubel, D.H. and Wiesel, T.N. (1959). Receptive fields of single cells in the cat's striate cortex. *Journal of Physiology (London)*, **148**, 574-591.
- Hubel, D.H. and Wiesel, T.N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology (London)*, **160**, 106-154.
- Johnston, A. and Clifford, C.W.G. (1995). A unified account of three apparent motion illusions. *Vision Research*, **8**, 1109-1123.
- Kolers, P.A. (1972). **Aspects of motion perception**. Elmsford, NY: Pergamon Press.
- Korté, A. (1915). Kinematoskopische Untersuchungen. *Zeitschrift für Psychologies*, 194-296.

- Kuffler, S.W. (1953). Discharge patterns and functional organization of mammalian retina. *Journal of Neurophysiology*, **16**, 37-68.
- Lappin, J.S., and Bell, H.H. (1976). The detection of coherence in moving random-dot patterns. *Vision Research*, **16**, 161-168.
- Ledgeway, T. and Smith, A.T. (1994). Evidence for separate motion-detecting mechanisms for first- and second-order motion in human vision. *Vision Research*, **34**, 2727-2740.
- Ledgeway, T. and Smith, A.T. (1995). Effects of adaptation to second-order motion on perceived speed. *Investigative Ophthalmology and Visual Science*, **36**, 53.
- Liu, Z., Gaska, J.P., Jacobson, L.D., and Pollen, D.A. (1992). Interneuronal interactions between members of quadrature phase and anti-phase pairs in the cat's visual cortex. *Vision Research*, **32**, 1193-1198.
- Limb, J.O., and Murphy, J.A. (1975). Estimating the velocity of moving objects in television signals. *Computer Vision and Image Processing*, **4**, 311-327.
- Livingstone, M., and Hubel, D. (1988). Segregation of form, color, movement, and depth: Anatomy, physiology, and perception. *Science*, **240**, 740-749.
- Lu, Z.-L., Sperling, G. (1995a). The functional architecture of human visual motion perception. *Vision Research*, **35**, 2697-2772.
- Lu, Z.-L., Sperling, G. (1995b). Attention-generated apparent motion. *Science*, **377**, 237-239.
- Marr, D. (1982). **Vision**, San Francisco: Freeman.
- Marr, D., and Ullman, S. (1981). Directional sensitivity and its use in early visual processing. *Proceedings of the Royal Society of London, Series B*, **211**, 151-180.
- Masson, G.S., Busetini, C., and Miles, F.A. (1997). Vergence eye movements in response to binocular disparity without depth perception. *Nature*, **389**, 283-286.
- Mather, G. (1988). Temporal properties of apparent motion in subjective figures. *Perception*, **17**, 729-736.
- Mather, G. and West, S. (1993). Evidence of second-order motion detectors. *Vision Research*, **33**, 1109-1112.
- Maunsell, J.H.R., Nealey, T.A., and DePriest, D.D. (1990). Magnocellular and parvocellular contributions to responses in the middle temporal visual area (MT) of the macaque monkey. *Journal of Neuroscience*, **10**, 3323-3334.
- Maunsell, J.H.R., Nealey, T.A., and Ferrera, V.P. (1992). Magnocellular and parvocellular contributions to neuronal responses in monkey visual cortex. *Investigative Ophthalmology and Visual Science*, **33**, 901.
- Maunsell, J.H.R., and van Essen, D.C. (1983a). Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *Journal of Neurophysiology*, **49**, 1127-1147.
- Maunsell, J.H.R., and van Essen, D.C. (1983b). Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. *Journal of Neurophysiology*, **49**, 1148-1167.
- Nakayama, K., and Silverman, G.H. (1984). Temporal and spatial characteristics of the upper displacement limit for motion in random dots. *Vision Research*, **24**(4), 293-299.
- Neuhaus, W. (1930). Experimentelle untersuchung der scheinbewegung. *Archiv für die gesamte Psychologie*, **75**, 315-458.
- Newsome, W.T., and Paré, E.B. (1988). A selective impairment of motion processing following lesions of the middle temporal visual area (MT). *Journal of Neuroscience*, **8**, 2201-2211.

- Newsome, W.T., Gizzi, M.S., and Movshon, J.A. (1983). Spatial and temporal properties of neurons in macaque MT. *Investigative Ophthalmology and Visual Science*, **24**, 106.
- Newsome, W.T., Wurtz, R.H., Dursteler, M.R., and Mikami, A. (1985). Deficits in visual motion processing following ibotenic acid lesions of the middle temporal visual area of the macaque monkey. *Journal of Neuroscience*, **5**, 825-840.
- Nogueira, C.A.M., Mingolla, E., and Grossberg, S. (1993). Computation of first order and second order motion by a model of magnocellular dynamics. *Investigative Ophthalmology and Visual Science*, **34**, 1029.
- Ögmen, H. and Gagné, S. (1990). Neural network architecture for motion perception and elementary motion detection in the fly visual system. *Neural Networks*, **3**, 487-506.
- Orban, G.A., and Gulyas, B. (1988). Image segregation by motion: Cortical mechanisms and implementation in neural networks. In R. Eckmiller and C. von der Malsburg (Eds.), **Neurocomputers**, Heidelberg: Springer-Verlag, 149-159.
- Pantle, A., and Picciano, L. (1976). A multistable movement display: Evidence for two separate motion systems in human vision. *Science*, **193**, 500-502.
- Ramachandran, V. (1985). Apparent motion of subjective surfaces. *Perception*, **14**, 127-134.
- Reichardt, W. (1961). Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In W.A. Rosenblith (Ed.), **Sensory communication** (pp. 303-317). New York: Wiley
- Schiller, P.H. (1982). Central connections in the retinal ON- and OFF-pathways. *Nature*, **297**, 580-583.
- Schiller, P.H. (1992). The On and Off channels of The visual system. *Trends in Neurosciences*, **15**, 86-92.
- Schiller, P.H., Logothetis, N.K., and Charles, E.R. (1990). Functions of the color-opponent and broad-band channels of the visual system. *Science*, **343**, 68-70.
- Schiller, P.H., and Malpeli, J. (1977). Properties of tectal projections of monkey retinal ganglion cells. *Journal of Neurophysiology*, **40**, 428-445.
- Slaughter, M.M., and Miller, R.F. (1981). 2-Amino-4-phosphonobutyric acid: A new pharmacological tool for retina research. *Science*, **211**, 182-184.
- Smith, A.T. and Ledgeway, T. (1997a). Sensitivity of second-order motion as a function of drift temporal frequency and viewing eccentricity. *Investigative Ophthalmology and Visual Science*, **38**, 401.
- Smith, A.T. and Ledgeway, T. (1997b). Separate detection of moving luminance and contrast modulations: fact or artifact? *Vision Research*, **37**, 45-62.
- Sperling, G. (1989). Three stages and two systems of visual processing. *Spatial Vision*, **4**, 183-207.
- 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, 103-121.
- Taub, E., Victor, J.D., and Conte, M.M. (1997). Nonlinear preprocessing in short-range motion. *Vision Research*, **37**, 1459-1477.
- Ternus, J. (1950). Experimentelle Untersuchungen über phänomenale Identität. In W. D. Ellis (Ed. and Trans.), **A sourcebook of Gestalt psychology**. New York: Humanities Press. (Original work published 1926.)
- Tolhurst, D.J. (1973). Separate channels for the analysis of shape and the movement of a moving visual stimulus. *Journal of Physiology*, **231**, 385-402.

- Tse, P., Cavanagh, P., and Nakayama, K. (1996). The role of parsing in high level motion processing. In Watanabe, T. (Ed.), **High level motion processing**, Cambridge, MA: MIT Press, in press.
- Turano, K. (1991). Evidence for a common motion mechanism of luminance- and contrast-modulated patterns: selective adaptation. *Perception*, **20**, 455-466.
- Turano, K., and Pantle, A. (1989). On the mechanism that encodes the movement of contrast variations: velocity discrimination. *Vision Research*, **29**, 207-221.
- van Doorn, A.J., and Koenderink, J.J. (1984). Spatiotemporal integration in the detection of coherent motion. *Vision Research*, **24**(1), 47-53.
- van Essen, D.C. (1979). Visual areas of the mammalian cerebral cortex. *Annual Review of Neuroscience*, **2**, 227-263.
- van Essen, D.C., Maunsell, J.H.R., and Bixby, J.L. (1981). The middle temporal visual area in the macaque: myeloarchitecture connections, functional properties and topographic organization. *Journal of Comp. Neurology*, **199**, 293-326.
- van Santen, J.P.H., and Sperling, G. (1984). Temporal covariance model of human motion perception. *Journal of the Optical Society of America A*, **1**(5), 451-473.
- van Santen, J.P.H., and Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America A*, **2**(2), 300-321.
- von Grünau, M. (1979). The involvement of illusory contours in stroboscopic motion. *Perception and Psychophysics*, **25**, 205-208.
- von Grünau, M. and Faubert, J. (1994). Intraattribute and interattribute motion induction. *Perception*, **23**, 913-928.
- Wallach, H. (1976). **On Perception**. New York, NY: Quadrangle Press.
- Watson, A.B., and Ahumada, A.J., Jr. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America A*, **2**(2), 322-342.
- Watson, A.B., Thompson, P.G., Murphy, B.J., and Nachmias, J. (1980). Summation and discrimination of gratings moving in opposite direction. *Vision Research*, **20**, 341-347.
- Zeki, S.M. (1974). Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *J. Physiol. London*, **236**, 549-573.

FIGURE CAPTIONS

Figure 1: Spatiotemporal representation of σ Display (Chubb and Sperling, 1989). (a) Near view. (b) Far view. Space is plotted on the horizontal axis and time on the vertical axis.

Figure 2: (a) The motion task used in Schiller *et al.* (1990). Monkeys fixate a point in the middle of the screen that is filled with random dots. When dots in a certain position begin moving coherently, monkeys are trained to saccade to that position. (b) Experimental results: Results before any lesion are represented as control. Parvocellular lesions do not produce any deficit in performance. Magnocellular lesions reduce performance to chance. (Adapted from Schiller *et al.*, 1990)

Figure 3: Effect of APB on cortical directionally selective cells. (a) A wide bright bar slides to the right through a cell's receptive field (RF), which is represented by the small circle. Bar edges are coded according to their spatial contrast (i.e., dark side on the left indicates a dark-light or DL edge, while dark side on the right indicates a light-dark or LD edge). (b) Luminance at the receptive field increases as the leading edge of the bar reaches the receptive field. Luminance decreases as the trailing edge reaches the receptive field. Before APB injection, the cell fires to both edges. Edge LD is the first one to cross the receptive field. After APB injection, the cell fires only at the passage of the trailing edge that indicates a decrease in luminance. (c) The bright bar now moves from right to left. (d) Now the first edge to cross the receptive field is the DL edge instead. Before APB injection, the cell again fires at passage of both edges. Responses are not as strong as when the bar was moving to the right, indicating that this cell is more selective to rightward motion than to leftward motion. After APB injection, the cell fires only at the passage of the trailing edge. (Adapted from Schiller, 1992)

Figure 4: Schematic of the Grossberg-Rudd motion model (Grossberg and Rudd, 1989, 1992).

Figure 5: Schematic of the Chey et al. motion model (Chey, Grossberg, and Mingolla, 1994, 1997).

Figure 6: (a) Model processing stages. (b) Model schematic. The bright and dark stimuli are represented at Level 1. A gated dipole detects the unoriented ON and OFF signals at Level 2. These signals are grouped into the lightening and darkening channels at Level 3 via an on-center off-surround network. The lightening channel is shown at the left. Level 4 is the short-range spatial filter and Level 5 pools signals from both channels.

Figure 7: Examples of center-surround processing in lightening and darkening cells. For stimuli, white implies a bright spot, black implies a dark spot, and grey implies no input. While for ON, OFF, lightening, and darkening cells, white implies active and grey in-active cell locations. (a) Onset and offset of bright spots. (b) A segment of σ display when observed from afar. (c) A segment of σ display when observed from nearby.

Figure 8: A qualitative representation of the functioning of a gated dipole as unoriented transient filter. See text for details.

Figure 9: Simulation results of an unoriented transient filter. The stimuli trace is shown at the bottom of each plot. (a) A stimulus is switched on generating ON transient and then switched off generating OFF transient due to rebound. (b) A stimulus is switched on generating ON transient and then replaced by a stimulus of opposite contrast generating OFF transient due to the combine

effect of rebound and opposite contrast phasic input.

Figure 10: Simulation results of first-order motion. (a) Stimulus, (b) u^{ON} , (c) u^{OFF} , (d) w_i^L , (e) w_i^D , (f) y_i^L , (g) y_i^D , (h) z_i . Variable z_i represents the rightward motion of both the leading and trailing edges.

Figure 11: Simulation results of first-order motion with blocked ON channel. (a) Stimulus, (b) u^{ON} , (c) u^{OFF} , (d) w_i^L , (e) w_i^D , (f) y_i^L , (g) y_i^D , (h) z_i . Variable z_i represents the rightward motion of the trailing edge.

Figure 12: Simulation results of second-order motion. (a) Stimulus, (b) u^{ON} , (c) u^{OFF} , (d) w_i^L , (e) w_i^D , (f) y_i^L , (g) y_i^D , (h) z_i . Variable z_i represents the rightward motion of the second-order stimulus.

Figure 13: Simulation results of , display - Near view. (a) Stimulus, (b) u^{ON} , (c) u^{OFF} , (d) w_i^L , (e) w_i^D , (f) y_i^L , (g) y_i^D , (h) z_i . Variable z_i represents the leftward motion of the near , display.

Figure 14. Simulation results of , display - Far view. (a) Stimulus, (b) u^{ON} , (c) u^{OFF} , (d) w_i^L , (e) w_i^D , (f) y_i^L , (g) y_i^D , (h) z_i . Variable z_i represents the rightward motion of the far , display.

Figure 15: Schematic of motion Boundary Contour System.

Figure 16: Simulation results of motion BCS for first-order motion. (a) x_i^{LL} , (b) x_i^{LR} , (c) x_i^{DL} , (d) x_i^{DR} , (e) y_i^{LL} , (f) y_i^{LR} , (g) y_i^{DL} , (h) y_i^{DR} , (i) $\Upsilon_i^{LL} + \Upsilon_i^{DL}$, (j) $\Upsilon_i^{LR} + \Upsilon_i^{DR}$, (k) Z_i^L , and (l) Z_i^R . Variable Z_i^R represents the rightward motion of both the leading and trailing edges.

Figure 17: Simulation results of motion BCS for first-order motion with blocked ON channel. (a) x_i^{LL} , (b) x_i^{LR} , (c) x_i^{DL} , (d) x_i^{DR} , (e) y_i^{LL} , (f) y_i^{LR} , (g) y_i^{DL} , (h) y_i^{DR} , (i) $\Upsilon_i^{LL} + \Upsilon_i^{DL}$, (j) $\Upsilon_i^{LR} + \Upsilon_i^{DR}$, (k) Z_i^L , and (l) Z_i^R . Variable Z_i^R represents the rightward motion of the trailing edge.

Figure 18: Simulation results of motion BCS for second-order motion. (a) x_i^{LL} , (b) x_i^{LR} , (c) x_i^{DL} , (d) x_i^{DR} , (e) y_i^{LL} , (f) y_i^{LR} , (g) y_i^{DL} , (h) y_i^{DR} , (i) $\Upsilon_i^{LL} + \Upsilon_i^{DL}$, (j) $\Upsilon_i^{LR} + \Upsilon_i^{DR}$, (k) Z_i^L , and (l) Z_i^R . Variable Z_i^R represents the rightward motion of the second-order stimulus.

Figure 19: Simulation results of motion BCS for for , display - Near view. (a) x_i^{LL} , (b) x_i^{LR} , (c) x_i^{DL} , (d) x_i^{DR} , (e) y_i^{LL} , (f) y_i^{LR} , (g) y_i^{DL} , (h) y_i^{DR} , (i) $\Upsilon_i^{LL} + \Upsilon_i^{DL}$, (j) $\Upsilon_i^{LR} + \Upsilon_i^{DR}$, (k) Z_i^L , and (l) Z_i^R . Variable Z_i^L represents the leftward motion of the near , display.

Figure 20: Simulation results of motion BCS for , display - Far view. (a) x_i^{LL} , (b) x_i^{LR} , (c) x_i^{DL} , (d) x_i^{DR} , (e) y_i^{LL} , (f) y_i^{LR} , (g) y_i^{DL} , (h) y_i^{DR} , (i) $\Upsilon_i^{LL} + \Upsilon_i^{DL}$, (j) $\Upsilon_i^{LR} + \Upsilon_i^{DR}$, (k) Z_i^L , and (l) Z_i^R . Variable Z_i^R represents the rightward motion of the far , display.