1995-11

Neural Dynamics of Motion Grouping: From Aperture Ambiguity to Object Speed and Direction

Neural dynamics of motion grouping: From aperture ambiguity to object speed and direction

Jonathan Chey, Stephen Grossberg, and Ennio Mingolla,

November, 1995

Technical Report CAS/CNS-1995-031

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

NEURAL DYNAMICS OF MOTION GROUPING: FROM APERTURE AMBIGUITY TO OBJECT SPEED AND DIRECTION

Jonathan Chey†, Stephen Grossberg‡and Ennio Mingolla§

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

Journal of the Optical Society of America, A, in press

Keywords: Vision, motion perception, directional grouping, speed perception, aperture problem, attention, visual cortex, neural network

Running Head: Neural Dynamics of Motion Grouping

†Supported in part by the Advanced Research Projects Agency (ONR N00014-92-J-4015), Air Force Office of Scientific Research (AFOSR F49620-92-J-0225), the National Science Foundation (NSF IRI-90-00530), the and the Office of Naval Research (ONR N00014-91-J-4100).

‡Supported in part by the Advanced Research Projects Agency (ONR N00014-92-J-4015), Air Force Office of Scientific Research (AFOSR F49620-92-J-0499), and the Office of Naval Research (ONR N00014-91-J-4100 and ONR N00014-95-1-0657, and ONR N00014-95-1-0409).

§Supported in part by the Air Force Office of Scientific Research (AFOSR F49620-92-J-0499). and the Office of Naval Research (ONR N00014-91-J-0597 and ONR N00014-95-1-0409).

¶Acknowledgments: The authors wish to thank Cynthia Bradford and Diana Meyers for their valuable assistance in the preparation of the manuscript.

Abstract

A neural network model of visual motion perception and speed discrimination is developed to simulate data concerning the conditions under which components of moving stimuli cohere or not into a global direction of motion, as in barberpole and plaid patterns (both Type 1 and Type 2). The model also simulates how the perceived speed of lines moving in a prescribed direction depends upon their orientation, length, duration, and contrast. Motion direction and speed both emerge as part of an interactive motion grouping or segmentation process. The model proposes a solution to the global aperture problem by showing how information from feature tracking points, namely locations from which unambiguous motion directions can be computed, can propagate to ambiguous motion direction points, and capture the motion signals there. The model does this without computing intersections of constraints or parallel Fourier and non-Fourier pathways. Instead, the model uses orientationally-unselective cell responses to activate directionally-tuned transient cells. These transient cells, in turn, activate spatially short-range filters and competitive mechanisms over multiple spatial scales to generate speed-tuned and directionally-tuned cells. Spatially long-range filters and top-down feedback from grouping cells are then used to track motion of featural points and to select and propagate correct motion directions to ambiguous motion points. Top-down grouping can also prime the system to attend a particular motion direction. The model hereby links low-level automatic motion processing with attention-based motion processing. Homologs of model mechanisms have been used in models of other brain systems to simulate data about visual grouping, figure-ground separation, and speech perception. Earlier versions of the model have simulated data about short-range and long-range apparent motion, second-order motion, and the effects of parvocellular and magnocellular LGN lesions on motion perception.

1 Introduction: Global Capture of a Moving Object's Direction and Speed

When an object moves, aperture ambiguity and image or detector noise often prevent all but a small subset of its image features, such as its bounding contours, from generating unambiguous motion direction cues. Despite this fact, object percepts often seem to pop-out with a well-defined motion direction and speed. The present article develops a neural model of how this feat is accomplished by the brain.

Figure 1

Five general design principles motivate the model's development. The first is that unambiguous *feature tracking* signals capture and transform ambiguous motion signals into coherent representations of object direction. Two classical examples of feature tracking are shown in Figure 1. A second principle is that object direction *and* speed are both emergent properties of this process.

These considerations lead to the model's central design problem: What type of feature tracking processes can select unambiguous direction *and* accurate speed signals from ambiguous motion signals? For example, consider the horizontal motion of both a vertical and a tilted line that move at the same speed. Suppose that the unambiguous feature tracking points at the line ends capture the ambiguous motion signals near the line middle. The preferred ambiguous motion direction and speed are normal to the line's orientation. For a vertical line, the speed of the feature tracking signals at the line ends equals that of the preferred ambiguous speed near the line middle. For a tilted line, the preferred ambiguous speed is less than that of the feature tracking speed. If the speed of the line is judged using a weighted average of feature signals and ambiguous signals, then the tilted line will be perceived to move slower than the vertical line, as found by Castet *et al.* (1993); see Figure 2d. These data also show that ambiguous speeds have a greater effect as line length increases when the line is briefly viewed. Feature tracking signals at the line ends thus propagate inwards along the line to capture ambiguous motion speeds and directions. Since capture takes longer to complete when lines are longer, ambiguous motion signals have a larger effect on longer lines.

Figure 2

Our model simulates data of Castet *et al.* (1993). It also simulates how the barberpole illusion (Wallach, 1976) is produced, how it is affected by configurational changes, and how plaid patterns move both coherently and incoherently. In particular, the model provides explanations of when moving plaid patterns cohere or do not (Adelson and Movshon, 1982; Kim and Wilson, 1993; Lindsey and Todd, 1995), how contrast affects their perceived speed and direction (Stone, Watson and Mulligan, 1990), and why movement of Type 2 patterns differs from those of Type 1 patterns (Ferrera and Wilson, 1990, 1991; Yo and Wilson, 1992). Bowns (1996) has provided experimental evidence against the Yo and Wilson (1992) hypothesis that Type II plaids move in the vector sum direction because of temporal delay between Fourier and non-Fourier information, and argues in favor of a feature tracking

explanation.

These simulations build upon earlier ones that employ a subset of model mechanisms that do not include its motion capture process (Chey *et al.*, 1994). The earlier model incorporates a third design principle: perceived motion direction and speed are a collective property of multiple populations of cells with different receptive field sizes, or scales. This multiscale network models the short-range motion process of Braddick (1974), including the fact that the short-range motion limit D_{max} depends on the spatial frequency content of the image (Anderson and Burr, 1987; Burr *et al.*, 1986; Nakayama and Silverman, 1984, 1985; Petersik *et al.*, 1983). In the model, larger scales preferentially process higher speeds. A key problem in designing this multiscale short-range motion filter is to keep the largest scale from winning at all motion speeds, because it has a larger receptive field that can attain a higher activation level. Once this problem is solved, the model simulates how visual speed perception, discrimination, and reaction time are affected by stimulus contrast, duration, dot density, and spatial frequency (Brown and Sekuler, 1980; Castet *et al.*, 1993; de Bruyn and Orban, 1988; Diener *et al.*, 1976; Ferrera and Wilson, 1990; Maunsell and van Essen, 1983; Mashour, 1964; Orban *et al.*, 1984, 1986; Stone and Thompson, 1992; Thompson, 1982; Tynan and Sekuler, 1982; Watamaniuk *et al.*, 1993).

A fourth design principle concerns how nearby contours of different orientation and contrast polarity that are moving in the same direction cooperate to generate a pooled motion direction signal. This process is modeled by a long-range motion filter that has been used to simulate data about long-range apparent motion (Kolers, 1972), including beta, gamma, delta, reverse, split, Ternus, and reverse-contrast Ternus motion and Korté's laws (Francis and Grossberg, 1996a; Grossberg and Rudd, 1989, 1992).

A fifth design principle concerns motion capture. The key problem here is how to capture the correct motion direction without distorting its speed estimate. Surprisingly, motion capture can be achieved by a long-range grouping network that also allows attention to prime a desired motion direction. In other words, motion capture, which seems to be an automatic and preattentive process, is carried out by the same circuit that permits top-down attention to selectively focus on a desired direction (Groner *et al.*, 1986; Sekuler and Ball, 1977; Stelmach *et al.*, 1994). Cavanagh (1992) has described an attention-based motion process, in addition to "low-level" or automatic motion processes, and has shown that it provides accurate velocity judgments. The present work models how the attentive process and the motion capture process are linked, and thereby explains how the attentive process yields accurate velocity judgments.

This capture-and-priming circuit obeys the same rules as attentive priming circuits that have been used to model other brain processes, such as early form vision, visual object recognition, auditory streaming, and speech perception (Grossberg, 1995). 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.*, 1995, 1996) have also been simulated using these model motion mechanisms as they interact with visual forms (Baloch and Grossberg, 1996a, 1996b).

Unlike models of Yo and Wilson (1992) or Castet *et al.* (1993), the present model does not postulate the existence of distinct channels for processing Fourier and non-Fourier signals, or for processing feature tracking and locally ambiguous motion signals. Instead, a single hierarchically organized processing stream reacts appropriately to both feature track-

ing and ambiguous motion signals in a context-sensitive fashion. This processing stream is incorporated into larger multistream model systems to explain data about form-motion interactions and visual search (Baloch and Grossberg, 1996b; Francis and Grossberg, 1996a; Grossberg, 1991, 1994; Grossberg, Mingolla, and Ross, 1994). Multiple streams are not, however, needed to explain the data that are considered here. The stages that are needed will first be functionally motivated before technical details are addressed.

2 Short-Range and Long-Range Motion Processing

Several types of data suggest that both spatially short-range and long-range mechanisms are involved in motion processing. One classical source are data about short-range and long-range apparent motion. Short-range apparent motion typically operates over distances under 15 minutes of visual angle and at inter-stimulus intervals under 100ms (Braddick, 1974, 1980). Long-range apparent motion operates over distances up to several degrees, and over inter-stimulus intervals up to 500ms (Kolers, 1972). Following the proposal of Grossberg and Rudd (1989, 1992), the present model further develops the hypothesis that both short-range and long-range motion are mediated by a hierarchically organized process within a single processing stream.

Various other percepts than apparent motion also suggest that initial localized motion measurements are followed by a long-range motion pooling stage. Plaid patterns (Adelson and Movshon, 1982), typically consist of two overlapping sinusoidally or square-wave modulated luminance gratings. Such a pattern has two possible visual interpretations: as a pair of independently moving components or a single coherently moving plaid. Evidence for a motion pooling process underlying coherent plaid motion has been obtained through experiments showing that the direction of motion of the individual components and the plaid as a whole respond differently to stimulus parameter changes. For example, Adelson and Movshon (1982) showed that contrast thresholds for the detection of plaid components alone were lower than contrast thresholds for the detection of coherently moving plaids. Ferrera and Wilson (1987) showed elevations in detection thresholds for plaid patterns following masking by plaids with different component angles, suggesting that masking reduced the response of mechanisms tuned to the plaid motion direction but not to the component directions. Movshon *et al.* (1985) identified motion-sensitive cells in visual area MT that were either tuned to the direction of individual components or to the direction of the plaid.

3 How is Ambiguous Motion Resolved?

Given that a long-range motion stage operates on the outputs of short-range motion detectors, how are the two stages combined to generate a global motion signal for a moving form? Adelson and Movshon (1982) suggested that plaid motion is determined by the *intersection of constraint lines* of each component. The intersection-of-constraints (IOC) solution does not, however, always correctly describe motion data. Ferrera and Wilson (1990, 1991) constructed several exceptions to the IOC solution using Type 2 plaids for which the motion vector given by the IOC lies outside of the arc formed by the motion vectors perpendicular to the two components (Figure 3a). Such plaids more rigorously test

the IOC solution than do Type 1 plaids, whose IOC solution direction lies within the two component directions. Ferrera and Wilson showed that, in certain situations, the motion of Type 2 plaids is biased towards directions normal to the component orientations and that their perceived speed is slower than predicted by the IOC. Rubin and Hochstein (1993) showed that when viewing moving lines through an aperture, observers misjudged direction away from the IOC solution and towards the vector sum of directions perpendicular to the component orientations.

Figure 3

Perceived Type 1 plaid directions can also deviate from the IOC solution. Stone *et al.* (1990) showed that when the two components are of unequal contrast, the plaid direction is biased towards the direction normal to the component with higher contrast. They modified the IOC rule to account for this bias by using an additional processing stage which makes a speed estimate that depends on component contrast. Stone and Thompson (1992) showed, that the speed of high contrast gratings is perceived as faster than those of lower contrast. Therefore, it is reasonable to assume that a higher contrast component will have a faster perceived speed, biasing the IOC calculated direction towards that component's motion vector.

Given that the IOC solution cannot easily explain all the data, notably the observed biases in Type 2 plaid motion, what alternatives are there? Plaid direction could be calculated as the vector sum of component directions; that is, as the sum of motion vectors orthogonal to the orientation of the components. However, vector summation is also insufficient. For example, the perceived motion direction of Type 2 plaids tends towards the IOC direction after a suitable duration (see below).

An alternative explanation of perceived plaid motion can given in terms of feature tracking. In plaid displays, the intersections of the component gratings are trackable features that provide unambiguous motion information that is the same as the IOC specified motion. Such a feature tracking explanation is, however, hard to distinguish from the IOC explanation since it, by itself, it does not explain deviations from IOC-computed directions any better than IOC.

The importance of feature tracking points in motion integration has been supported by studies showing that feature signals can be differentiated from ambiguous motion signals by their different responses to changes in stimulus parameters. For example, Lorenceau *et al.* (1993) measured discrimination performance in the perception of translating lines of different orientations, contrasts and lengths. Discrimination worsened at low contrast, long length and short duration. They postulated higher contrast thresholds and longer integration times for the mechanisms that respond to feature signals. Similarly, Yo and Wilson (1992) showed that the bias in perceived motion direction of Type 2 patterns towards the vector sum of the component directions was reduced over time, which they interpreted as suggesting longer integration times for the feature motion detectors.

Mingolla *et al.* (1992) studied the integration of directional information using a display with small apertures, each of which contained a single moving bar. By varying bar length, they controlled whether feature information was present in each aperture. Bars whose end points were visible had correctly perceived motions; otherwise motion was ambiguous, as

in the classic aperture problem. Mingolla et al. (1992) found that features substantially affected the perceived direction of motion. These results suggest that feature information is critical in plaid perception and that IOC constraints are not calculated from ambiguous motion signals without feature motion information.

The results of Mingolla *et al.* (1992) and Castet *et al.* (1993) suggest that perceived directions combine direction signals normal to component orientations and feature tracking signals. Although these results show that feature information can overrule ambiguous motion signals, In addition, the perceived motions of Type 2 plaids show that feature signals do not always completely dominate ambiguous motion signals. In addition, Lorenceau and Shiffrar (1992) showed that motion information is not integrated easily when features signaled different motion directions in each aperture, but is integrated when only ambiguous motion signals filled the apertures. Thus the two sorts of motion signals appear to be averaged, with feature signals given more importance than ambiguous signals.

4 Motion Coherence and Incoherence

In certain situations, the visual system does not combine motion signals across space, as when plaid components do not cohere. Adelson and Movshon (1982) proposed that coherence is determined by the relative spatial frequencies of the components. Kim and Wilson (1993) argued that the difference in orientation or direction of motion of two gratings determines their coherence, with gratings having closer orientations more likely to cohere. Williams and Sekuler (1984) earlier showed that coherent motion can be observed in fields of dots moving in a range of directions if that range is sufficiently small. Stoner *et al.* (1990) emphasized relative depth cues derived from the luminance values where the gratings intersected. In their experiments, if the luminances signaled that the gratings were transparent, then they tended not to cohere. Other static depth cues also affect the probability of grating coherence (Trueswell and Hayhoe, 1993).

Grating coherence can thus depend on many cues. Although depth cues are not included here, Grossberg (1994a, 1997) has modeled how depthful surfaces, including transparent surfaces, may separate occluding and occluded forms. Baloch and Grossberg (1996a, 1996b) and Francis and Grossberg (1996a) have modeled how these forms may modulate motion percepts. These form-motion interactions, are not invoked here. Instead, we simulate data wherein the relative orientations of components determines perceived coherence.

5 Model Overview: Short-Range Spatial Processing

The model stages are functionally motivated in this section and the next.

Figure 4

5.1 Multiscale Short-Range Spatial Filter

Figure 4 summarizes the model's five main processing stages. Figure 5 plots how the first three stages respond to one-dimensional inputs as a function of their speed (Chey

et al., 1994). First, transient cells react to changes in visual inputs. Cells sensitive to image transients occur at several stages of brain motion processing, including the y cells of the retina (Enroth-Cugell and Robson, 1966; Hochstein and Shapley, 1976a, 1976b). The hypothesis that transient cells input to the motion system is consistent with data showing that lesions of magnocellular LGN players greatly disrupt motion processing but lesions of the parvocellular layers do not (Logothetis *et al.*, 1990). The transient cells input to a spatially short-range multi-scale filter that models the short-range motion process of Braddick (1974, 1980). Each curve of Figure 5 (Level 2) plots the maximum activity of one filter size as a function of input speed. The largest scale wins at every speed because each filter responds as soon as it receives any input. To achieve a more selective response, larger scales should require more input in order to fire. This problem is corrected as follows.

First, each filter is given a positive output threshold that increases with filter size; that is, output threshold is self-similar across scale. Self-similar thresholds have also been used to model multiscale groupings of visual (Grossberg, 1987) and speech (Cohen and Grossberg, 1986) signals. They seem to be a rather general principle of neural design. Figure 5 (Level 2, thresholded output) shows the maximal thresholded output of each scale as a function of input speed. Now each scale wins within a different speed range that increases with scale size.

Figure 5

These monotonically increasing curves are not, however, well enough separated to form usable tuning curves. Competition (across position within each scale, and across scale at each position) achieves this goal, see Figure 5 (Level 3). A weighted average of the outputs of these tuning curves was used in Chey *st al.* (1994) to simulate data about how speed perception, discrimination, and reaction time are affected by stimulus contrast, duration, dot density, and spatial frequency.

In summary, the first few stages of the model use transient cells that feed a multiscale short-range motion filter whose larger scales selectively process higher speeds due to the combined action of self-similar thresholds (at the low end of their tuning curves) and competition (at the high end).

5.2 From Undirectional to Directional Processing

In the simulation of Figure 5, the moving target was one-dimensional. In the full two-dimensional model that is developed here, a central problem is how to convert undirectional transient cells responses into directionally-sensitive responses. An early stage in this transformation uses an inhibitory veto mechanism (Emerson and Gerstein, 1977; Ganz, 1984; Goodwin *et al.*, Heggelund, 1984). Barlow and Levick (1965) first showed that inhibition led to directionally selective ganglion cells in the rabbit retina. These inhibitory connections veto responses in nearby cells, using a logical NOT operation. Gamma-aminobutyric acid (GABA) mediates inhibition in directional rabbit retina cells, and GABA antagonists eliminate directional selectivity (Ariel and Daw, 1982). These ganglion cells respond to single light flashes with much the same threshold as paired flashes presented in the direction that was not vetoed by inhibition.

Evidence for inhibitory processes in directional selectivity has also been found in cat

cortex. Hubel and Wiesel (1959, 1962) suggested that directional selectivity of simple cells could be explained by summing responses at adjacent ON and OFF regions of the cell. Moreover, ON and OFF retinal ganglion cells (Kuffler, 1953) converge at cortical simple cells (Schiller, 1982). However, later studies rejected this hypothesis (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 the cells' directional selectivity did not correlate with the spatial arrangement of their receptive fields and was 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 directional selectivity.

Both Barlow and Levick (1965) and Goodwin *et al.* (1975) found directional selectivity within small sub-units of observed cell receptive fields. For example, Goodwin *et al.* reported that one cell was divided into 22 subunits, 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 processes, which operate at a small scale compared to receptive field sizes of directionally selective cells in either rabbit retina or cat cortical cells.

At what processing stage does such a directional veto mechanism operate? Consistent with the above data, we suggest that it occurs as part of transient cell processing, prior to the short-range filter. This hypothesis minimizes the sensitivity of the veto mechanism to stimulus form. Since the vetoing signal is spatially offset from the cell that it is vetoing, a cell can be erroneously activated (not vetoed) if it occurs at line ends or corners when the veto mechanisms lies beyond the end of the stimulus. Such problems are reduced if vetoing occurs before short-range filtering, since the impact of a small number of false directional signals is reduced by subsequent spatial averaging.

As noted below, the veto mechanism is designed so that directional transient cells respond just as well at fast speeds as at slow speeds. Responses of directional transient cells to a tilted line moving to the right are shown in Figure 6a. These responses are ambiguous with respect to the direction of motion of the line. They only constrain it within 180 degrees. The aperture problem is clearly visible here and there is, as yet, no representation of stimulus speed.

Figure 6

5.3 Short-Range Filter and Feature Tracking

Figure 6a shows that the very short spatial range over which vetoing operates is insufficient to generate a reliable feature tracking signal. Vetoing eliminates the wrong direction, but it does not boost processing in the right direction. Such a boost is needed because a small set of feature tracking signals needs to overwhelm the effects of a much larger set of ambiguous signals (Figure 1). The short-range filter is assumed to be spatially anisotropic to accumulate evidence that a feature is moving in a given direction and thereby boost its

feature tracking signals.

Figure 6b and 6c show how the anisotropic short-range filter and its self-similar threshold transform the directional transient cell outputs of Figure 6a in response to a tilted line moving to the right. The thresholded outputs of the horizontally oriented filters at the line ends respond best, and *only* to the correct direction of motion at these feature tracking points. All other positions still experience the aperture problem.

5.4 Competition

As in the one-dimensional model, competition converts cell responses into true tuning curves. Intrascale competition again occurs across space within each scale. It does so only in the direction of its short-range filter, cross-directional inhibition could severely impair speed estimates. Interscale competition again occurs at each position within each direction. These competitive interactions do not, however, sufficiently boost the amplitude of feature tracking signals relative to ambiguous signals.

What is needed is a form of competition that enhances the activities of directional cells that have few directional competitors at a given position, attenuates activities of directional cells with many directional competitors, and does not disrupt speed estimates. A divisive, or shunting, competition across direction and scale accomplishes this by computing the ratio of competing activities (Grossberg, 1980, 1983). Competition occurs across direction and scale because many speeds are represented, and thus scales activated, at each ambiguous point (Figure 1). Directional competition also increases with the directional difference. Its net effects are two-fold: Unambiguous feature tracking signals are boosted relative to ambiguous signals, and ambiguous signals are biased towards a direction of motion that is perpendicular to a line's orientation.

6 Model Overview: Long-Range Spatial Processing

6.1 Long-Range Filter

The long-range motion filter pools signals from multiple orientations and contrast polarities in a prescribed direction of motion (Grossberg and Rudd, 1992). It is the model processing stage that generates cells which are truly directionally selective, and is proposed to occur in cortical area MT where similar cells occur (Albright *et al.*, 1984; Allman *et al.*, 1985; Maunsell and van Essen, 1983a; Newsome *et al.*, 1983; Zeki, 1974). This processing stage also pools signals from both eyes, as do MT cells (Bradley *et al.*, 1995; Maunsell and van Essen, 1983b), and explains how long-range apparent motion occurs with dichoptically presented stimuli (Gengerelli, 1948; Spigel, 1968). Long-range spatial averaging further boosts the relative advantage of feature tracking signals, especially at object corners (Figure 7). It also reduces directional and speed biases due to incomplete activation of the inhibitory kernels of the intrascale competition, while preserving speed estimates elsewhere.

Figure 7.

6.2 Long-Range Directional Grouping and Attentional Priming

The grouping network enables the small set of stronger feature tracking signals to select consistent directional and speed signals form the large set of weaker ambiguous motion signals. The grouping cells interact via feedback with the long-range filter cells (Figure 1). In this way, the small advantage of feature tracking signals in a given spatial region can be amplified at the grouping cells and fed back to the long-range filter cells, where consistent directional and speed signals are selected and inconsistent ones suppressed. This selection process expands the region of consistent signals. Another pass through the feedback loop expands it further, and so on.

A feedback system requires more processing time than a feedforward system. Various data show that motion capture does take time. These include data of Castet *et al.* (1993) on the effects of a tilted line's length on its perceived speed (Figure 2), and data of Ferrera and Wilson (1990), simulated below, about how the perceived motion directions of Type 2 plaids change towards the feature motion direction over 60-150 msec. Wilson *et al.* (1992) interpreted this effect as evidence for a longer integration time in a pathway specialized for processing feature tracking signals. Data of Bowns (1996) contradict this interpretation. The present model proposes that this amount of time naturally arises within a single processing stream as a result of long-range grouping and feedback.

The feedback mechanism is implemented through a layer of *long-range directional grouping cells* (Figure 8). Such cells receive excitatory input over a wide spatial extent from all scales of long-range filter cells with similar directional preferences. Thus grouping cells pool over both space and direction. Such cooperative influences have been described in experiments containing elements moving in a range of different directions (Williams and Sekuler, 1984). A winning direction is then selected by mutual competition between grouping cells tuned to different directions.

Figure 8

Inhibitory feedback from the winning grouping cells is delivered to all long-range filter cells that are tuned to different directions. Through this suppression, grouping cells choose a direction, but not a speed. The speed choice is implicit in the direction choice, being whatever speed is represented by the surviving long-range filter cells. The original speed estimates of the winning direction are unaffected by this operation.

The grouping feedback can be implemented in at least two ways. One way distributes inhibition to all long-range filter cells that are tuned to other directions. In the other way, inhibitory feedback nonspecifically inhibits all directions, but is supplemented by specific excitatory feedback from grouping cells to long-range filter cells of the same direction (Figure 11). The specific excitation balances the nonspecific inhibition at that direction. The net effect is again off-surround inhibition. Here, though, there is no need to selectively grow inhibitory connections to all other directions. All inhibitory feedback is nonspecific and all excitatory feedback specific and reciprocal. This network realizes a matching rule that is familiar in Adaptive Resonance Theory, or ART (Carpenter and Grossberg, 1987; Grossberg, 1995). By the ART matching rule, top-down matching signals (here from grouping to filter cells) can prime a given direction while inhibiting all other directions.

During matching of bottom-up and top-down signals, only bottom-up signals that are confirmed by excitatory top-down prime can survive the inhibition. All other signals, here all other directions, are inhibited. This process of top-down priming realizes a type of attention. Higher cognitive processes can use this priming mechanism to track objects that are moving in attended directions.

The directional grouping circuit is proposed to occur in the ventral part of MST, which has large directionally tuned receptive fields that are specialized for detecting moving objects (Tanaka *et al.*, 1983). In this interpretation, MST_v can attentionally modulate MT, which is proposed to include long-range filter cells. Consistent with this proposal, Treue and Maunsell (1996) have shown that attention can modulate motion processing in cortical areas MT and MST in behaving macaque monkeys. O'Craven *et al.* (1996) have shown using fMRI that attention can modulate the MT/MST complex in humans. This interpretation predicts that MST_v cells make a directional choice whose feedback to MT overcomes aperture ambiguities and selects an object's true direction of motion.

7 Data Simulations

The model is used to simulate increasingly complex properties of psychophysical data, namely: how feature tracking and ambiguous signals are combined, how feature tracking signals interact with each other, and how feature tracking and ambiguous signals may not combine.

7.1 Line Motion

Castet *et al.* (1993) had observers compare the speed of tilted lines to that of a vertical comparison line undergoing horizontal movement. They varied both the orientation and length of the lines. Their data show two major effects (see Figure 2a). First, the perceived speeds of tilted lines are slower and the degree of speed bias increases with line tilt from vertical. Second, the magnitude of this bias increases with line length. Our simulations show how unambiguous feature tracking signals and ambiguous motion signals from line interiors are combined. In each simulation, a line length and orientation were chosen and the motion of the line with those characteristics was simulated. Speed data were collected at a fixed time after the simulated motion starts. A ratio was computed between the spatially averaged speed signal obtained from the simulation and that obtained from a simulation with a vertical line of the same length and speed. These ratios measure the perceived relative speed of each line to the vertically oriented line. Figure 2b shows these ratios for three orientations and three line lengths.

Figure 9

In Castet *et al.* (1993), each line was displayed for only a short duration (167ms). The length of this presentation is compatible with the hypothesis that feature tracking information has not yet fully propagated along the line and that speed biases are due to residual ambiguous motion signals present along the line length. This hypothesis suggests that the biases are due to incomplete processing and are therefore transient.

Figures 9 shows the evolution over time of long-range filter cell activities during motion

of a line tilted at 45 degrees from vertical. As the line moves, grouping cells become active and propagate the feature tracking signals along the line. Although the interior line signals originally indicate lower speeds perpendicular to the line's orientation, over time, the perceived direction and speed at each point becomes consistent along the whole line length. Figure 10 plots how perceived direction and speed change gradually over time.

Figure 10

7.2 Barberpole Motion

When a moving oriented line is viewed through a circular aperture, its motion is ambiguous. However, when a line is viewed through a rectangular aperture, the classic barberpole illusion is produced, whereby the line is perceived to move with the direction of the aperture's long axis (Wallach, 1976). This suggests that the visual system utilizes the feature tracking signals derived from the line endings at the edge of the aperture to determine line direction. See Hildreth (1984) and Nakayama and Silverman (1988) for supportive data.

Under certain conditions, the barberpole display can simultaneously support at least two distinct perceived motion directions: motion coincident with the long axis of the aperture and motion orthogonal to the line. For example, in a horizontal aperture with a diagonally oriented moving line, the line appears to move diagonally in the corner regions and horizontally in the central region (Figure 1b). This percept is most prevalent when one line is in the aperture at any time. Multiple moving lines tend to lock the percept into motion along the length of the aperture.

There are at least two possible explanations for the perceived diagonal motion in the corner of the barberpole display: The diagonal motion could be an average of the horizontal and vertical feature signals at each end of the line. Alternatively, the competing feature signals could cancel each other out, leaving ambiguous motion signals in the direction normal to line orientation.

One way to distinguish between these explanations is to alter the orientation of the line. According to the first explanation, this should result in at most a small shift towards the direction normal to the new orientation. According to the second explanation, such a change in orientation should also have little effect until it becomes large enough to favor one of the feature directions. Then the motion may be entirely captured by that direction. Informal observations suggest that the second explanation is correct: The percept is usually in the direction of the aperture edge whose orientation is most nearly perpendicular to the line. This result suggests that diagonal motion in the corner of barberpole displays results from inconclusive competition between different feature motion signals, in the sense that no feature gains dominance and propagates to across the moving line.

The difficulty of integrating feature signals in different directions of motion was also demonstrated by Lorenceau and Shiffrar (1992), who studied the integration of motion information between multiple apertures that revealed a portion of a translating diamond. The terminator motions from the different apertures could indicate different directions of motion, although the diamond moved rigidly as a whole. Observers had difficulty perceiving rigid diamond motion, suggesting that feature information could not be integrated across the apertures. Integration became more likely after stimulus manipulations reduced

the influence of terminator motions, such as adding jagged aperture edges, or using low contrast terminators relative to contour contrast.

Interference between multiple feature tracking signals is attributed in the model to grouping cell kernels that are sufficiently large to overlap several feature tracking signals. When small grouping cell kernels are used, there will be some locations near the line end that cohere with just the nearest feature tracking signal.

Figure 11 illustrates time slices of motion signals from a simulation of a moving line behind a rectangular aperture. Figure 12 shows how a small (5 degree) change in line orientation causes more rapid convergence to the feature direction closest to the vector perpendicular to the line's orientation. The larger the bias provided by the line orientation, the more rapidly this effect is felt, until the extreme case is reached where the line moves perpendicularly to an aperture edge.

Figure 11

Another way for one feature tracking direction to win in a corner region is to prime the grouping cells to a particular motion direction, and to influence the long-range filter cells via feedback. This situation can occur when multiple lines move behind the aperture. Initially, the motion of the line is ambiguous, but over time a winning direction emerges at the grouping cell level. If a second line enters the aperture before residual grouping cell activity decays, the previous winning direction can continue to win the competition. This priming effect suggests why corner diagonal movements are not observed when multiple lines occur in a barberpole display.

Figure 12

7.3 Plaid Motion

Kim and Wilson (1993) argued that relative line orientation is the prime determining factor in determining plaid coherence. There is, however, still much dispute regarding the importance of other stimulus parameters and the probabilities of observing coherent motion (Lindsey and Todd, 1995; Stoner *et al.*, 1990; Stoner and Albright, 1992; Trueswell and Hayhoe, 1993). Part of this dispute is due to the many stimulus configurations employed in plaid motion designs.

Kim and Wilson (1993) reported that coherent motion is almost always observed for component orientations within 45 degrees of the plaid direction and almost never for larger orientational differences. These data were collected using sine-wave gratings of different spatial frequencies. Lindsey and Todd (1995) reported that square wave gratings moved coherently at all relative orientations and that it was necessary to undergo prolonged viewing before incoherent motion could be observed at all. Lindsey and Todd also found that increasing orientational differences results in an increased probability of observing incoherent motion after adaptation.

The resilience of coherent motion percepts is not surprising when one considers that a plaid display contains only consistent feature tracking motion signals at the component intersections. The model assumes that incoherent motion results from failure of feature

tracking signals to dominate the grouping cell competition, freeing individual component motions to express directions normal to their orientations. Unlike the barberpole display, incomplete competition cannot arise since there are no conflicting feature tracking signals. The only competition comes from the ambiguous signals of the individual components.

If this is true, then why are 2-D plaid percepts *ever* observed to move incoherently? How can plaid displays exhibit incoherent motion when simple lines of a single orientation are always coherent with feature signals derived from their end points. Two elements of the model help to explain this. First, the feature tracking signals derived from plaid displays differ from those obtained from moving lines with visible end points. The intersection points of plaid components, although forming trackable features, may be less salient than line ends. For plaids formed from sine-wave gratings, the intersection points form amorphous blobs that have no clear edges or other feature points to track. This may explain why Kim and Wilson (1993) and others who use sine-wave gratings so often observe incoherent motion. Square-wave plaids have more sharply defined intersection points. When the luminance values at the intersections are added, the intersections form moving diamonds that are easily tracked. Luminance discontinuities are not necessary to observe coherent motion, as when the luminance of the moving diamond equals that of the components.

The simulations consider square-wave plaids with uniform luminance values. At each intersection of such a display (Figure 13b), there are four feature points, one at each corner of the diamond, that provide trackable features. To track both leading and trailing edges at an intersection point, the model uses both ON and OFF cells. ON cell transient responses are generated along the leading edge of both contours (Figure 13b), OFF responses along the trailing edges (Figure 13c).

Grossberg and Rudd (1989, 1992) proposed that the short-range filters process ON and OFF transient responses independently and that these channels are combined at the long-range filter stage. Until now, we have considered only a single set of transient responses without specifying whether they are ON, OFF, or a combination of both. Both channels are now used, and are processed independently until combined at the long-range filter. The resulting feature signals are smaller than those derived from line ends, because two orientations join at a corner. At such a corner, both long-range and short-range filters sum over both sides of the contour, resulting in a less specific feature signal. This helps to explain why plaid patterns are more likely to move incoherently than lines. Baloch *et al.*, (1996) have shown how these ON and OFF mechanisms simulate first-order and second-order motion percepts.

The second major factor controlling plaid incoherence, particularly for the square-wave gratings studied by Lindsey and Todd (1995), is that the viewer is exposed to the motion for a long time. Plaid pattern motion is highly repetitive and can thus fatigue motion detectors. Likewise, after prolonged viewing of barberpole displays, percepts can fluctuate from one feature direction to the other, also presumably as a result of adaptation to the prevailing direction.

7.3.1 Coherent Plaid Motion

Ferrera and Wilson (1987) classified plaids into three groups: Type 1 plaid direction lies in the arc between the directions normal to the components. Type 2 plaid direction

lies outside of this arc (Figure 3). Type 1 patterns are called symmetric or asymmetric, depending on whether the angles formed between the components and the direction of motion are the same for both components (symmetric) or not (asymmetric).

Figure 13

The simplest pattern, Type 1 symmetric, is simulated first. Figure 13 shows the stimulus configuration and the ON and OFF transient cell responses along its leading and trailing edges. Figure 14 shows long-range filter outputs generated from the combined ON and OFF channels. Figure 15a shows coherent plaid motion. The feature motion signals derived from the corner points have propagated across the whole plaid pattern, capturing motion into a common horizontal direction that corresponds to the motion direction of the plaid as a whole.

Figure 14

7.3.2 Incoherent Plaid Motion

Coherent plaid motion occurs when a single direction wins at the grouping cells. Incoherent motion occurs when no such winner is established. Under the assumption that adaptation to the perceived direction during coherent motion contributes to incoherent motion percepts, incoherent plaid motion can be simulated by reducing feature signal strengths to a level where they can no longer win the grouping cell competition.

Several possible mechanisms could become fatigued during coherent plaid motion. We assume that this adaptation occurs at only a single processing stage, the synapses connecting the long-range filters to the grouping cells. A homologous type of habituation has elsewhere been used to explain data about visual persistence, form-motion interactions, and aftereffects (Francis and Grossberg, 1996a, 1996b; Francis, Grossberg, and Mingolla, 1994; Grossberg, 1991). In the present application, after adaptation, each grouping cell receives a smaller input signal from the long-range filter cells that have been active. During coherent motion, grouping cells suppress the activity of all long-range filters but those whose directional preferences match the chosen direction of movement (Figure 8). Therefore, only these filter cells will become adapted, while all other filters will remain in an unadapted state.

Figure 15b simulates long-range filter activity with the same Type 1 symmetric plaid, but with adapted horizontal filter cell responses. Now there is no winner at the grouping cell level, so ambiguous long-range filter responses remain intact along the components. These component responses are hypothesized to correspond to incoherent motion.

Figure 15

As discussed above, it is difficult to qualitatively model data regarding plaid motion due to the variety of stimulus configurations and viewing conditions used and the variability in results obtained under different experimental regimes. However, other things being equal,

simulations should show that incoherent motion is correlated with the difference between component orientations. Simulations were run in which plaid patterns were tested with different levels of long-range filter cell adaptation. These simulations show that greater adaptation is required to produce incoherent motion for smaller differences between the component orientations (see Figure 16).

Figure 16

Type 1 Asymmetric plaids exhibit much the same behavior with respect to coherence and incoherence as do Type 1 Symmetric plaids. The critical variables are again the relative difference between the component orientations and the actual direction of motion. The model predicts that the relative orientation difference between the components is less important than the relative differences between the components and the feature tracking direction. Asymmetric plaid patterns could be used to test this hypothesis, since it is possible to vary these parameters independently. Such manipulations have not been reported in the literature.

7.3.3 Type 2 Plaid Motion

In Type 1 plaids, coherence is aided by the fact that the plaid motion direction lies between the two component directions. In Type 2 plaids, this is not the case and this difference presumably underlies the fact that biases have been reported in the perceived direction of Type 2 plaids, but not Type 1 (Ferrera and Wilson, 1990). A key feature of these biases is that their magnitude is duration-dependent (Figure 17a). Wilson, Ferrera and Yo (1992) modelled this bias as the result of a delay in a non-Fourier motion pathway, so that initial percepts are based only on the responses of Fourier motion pathways, which respond to the component orientations only. An explicit delay was required because model output is computed simultaneously with motion integration. Such a postulate of differential response time is unnecessary in the present model. Instead, the change in bias derives from the integration time needed for the grouping cells to become active and influence the long-range filter cells. The critical model property is that long-range filter cells become active before the feedback grouping mechanism selects a winning direction. During this time, model response is based on feedforward motion signals and is biased towards the component motion signals. Figure 17b plots the time course of the present model's perceived motion direction in response to a simulated Type 2 plaid.

Figure 17

Such directional biases have not been reported for Type 1 plaid patterns. They would presumably be small due to the similarity between the vector average of component directions and the plaid direction. The model suggests, however, that there may be an additional reason why biases are not reported for Type 1 asymmetric plaids: When grouping cells select a winning direction, grouping cell activity eliminates competing directional signals, leading to a gradual change in overall perceived direction towards the winning grouping cell's preferred direction. For Type 2 plaids, activity along both components

converges to the actual plaid direction from the same side. For Type 1 plaids, each component converges to the plaid direction from a different side. Although not modelled, it is reasonable to assume that component motions lead to a perceived coherent motion only if they are sufficiently similar. Thus, it may be that biases are reported in Type 2 motion for situations in which Type 1 plaids move incoherently.

Informal observations suggest that Type 2 plaids tend to move incoherently. Even when they move coherently, the blobs at the intersection point of the components are often segmented from the components themselves. However, Kim and Wilson (1993) reported that for a wide range of component directions, ranging from 18.4 to 71.6 degrees away from the plaid direction, Type 2 patterns moved coherently when different spatial frequency components were employed. Kim and Wilson modeled the propensity for Type 2 plaids to cohere by asserting that the proposed Fourier motion pathway would be inactive for components with such widely varying frequencies and thus their coherence would be based purely on the overlap of Fourier signals derived from the components. Because the component orientations are so similar, coherence in the Kim and Wilson model is almost always predicted for Type 2 patterns. However, such coherence is erroneously predicted to result in motion that is the vector sum of the component orientations due to the non-response of the non-Fourier motion pathway. In contrast, the present model predicts that coherent motion relies on the presence of a strong feature tracking signals. If such feature tracking signals are absent, for whatever reason, then ambiguous motion signals do not have a large enough magnitude to win competition at the grouping cell level and so no coherence is established.

7.3.4 Contrast Effects on Plaid Motion

Chey *et al.* (1994) simulated the observed changes in speed perception in response to stimulus contrast (Stone and Thompson, 1992). These results are extended herein to include the effects of grouping on contrast-sensitive plaid motion. As discussed above, Stone *et al.* (1990) used the fact that perceived speed increases with contrast to provide an IOC-based explanation for the reported biases in plaid motion with unequal contrast components. The model provides an alternative account of their data, as follows. Increases in stimulus contrast increase the energy or activation levels of the short-range filters. Velocity estimates are determined by the energy produced by the short-range filters. If the filters that detect the motion of the higher contrast grating are more highly activated, then their velocity estimates will dominate the directional percept, leading to the observed biases.

As was demonstrated for Type 2 patterns above, directional percepts in the model converge to the feature tracking direction over time. When symmetric plaid patterns are employed, such convergence is not noticeable because the energy contributed by each component is balanced by the other. However, when one component has a higher contrast than the other, this is no longer the case, and a temporary bias occurs to one component's direction of motion. These biases can explain the data reported in Stone *et al.* (1990) because, as in the Ferrera and Wilson (1990) study, their data were gathered with short duration viewing. Each plaid was viewed for 300ms, but was only at full contrast for 200ms. The authors did not report results using longer viewing periods, possibly because

the biases became less evident or non-existent after prolonged exposure.

The Stone *et al.* (1990) data were simulated by changing the input magnitudes used for each component grating in plaid input patterns, as was done Chey *et al.* (1994) to simulate stimulus contrast changes. These input magnitudes were varied in the same way as they were in the Stone *et al.* (1990) experiments; namely, by choosing a base input contrast/magnitude and then varying the component contrast ratios while keeping the total input contrast/magnitude constant. Figure 18 shows that the simulations provide a good qualitative fit to these data. These results used 1.5 simulation time steps, a slightly longer duration than was used in the Type 2 simulations. This also reflects the data: Ferrera and Wilson (1990) employed shorter durations than did Stone *et al.* (1990).

Figure 18

8 Discussion

The speed-sensitive motion Boundary Contour System that is further developed herein has, by now, been shown to simulate a wide range of behavioral and neural data about motion perception. The earliest version of the model was used to simulate data about short-range and long-range apparent motion, including data about beta, gamma, delta, reverse, split, Ternus and reverse-contrast Ternus motion and Korté's laws (Francis and Grossberg, 1996a; Grossberg and Rudd, 1989, 1992). It also suggested how long-range apparent motion mechanisms could be used to generate continuous attentive tracking signals in response to moving targets that are intermittently occluded by intervening objects (Grossberg, 1991, 1996).

The model was then adapted in Nogueira *et al.* (1993) to use only transient cell inputs to explain additional data about second-order motion, including the perceived reversal of motion direction with distance from the stimulus, and why monkeys with lesions of the parvocellular but not magnocellular layers of the lateral geniculate nucleus are capable of detecting the correct direction of second-order motion (Schiller *et al.*, 1990). Baloch *et al.* (1996) extend these results. These studies did not, however, fully exploit the multiple spatial scales of the model. Chey *et al.* (1994) showed how multiscale filtering and competition could be used to simulate further data concerning how visual speed perception and discrimination are affected by stimulus contrast, duration, dot density, and spatial frequency.

The present refinement of the model builds upon these results and those of Grossberg and Mingolla (1993) to show how the global aperture problem may be solved; namely, how long-range filtering and grouping mechanisms can transform the outputs of the speed-sensitive multiscale filter into a coherent representation of object speed and direction. The model shows how the motion capture process that carries out this transformation can also act like a top-down attentive mechanism for priming motion directions without disrupting speed estimates. The model hereby clarifies how low-level automatic motion detectors and attention-modulated motion grouping processes may work together to generate accurate velocity judgments (Cavanagh, 1992).

This capture-and-priming circuit is familiar in Adaptive Resonance Theory, where it

has been used to simulate data about early vision, visual object recognition, auditory streaming, and speech perception (Carpenter and Grossberg, 1987; Gove *et al.*, 1995; Govindarajan *et al.*, 1994; Grossberg, 1995; Grossberg *et al.* 1994). This connection with Adaptive Resonance Theory indicates how key model parameters may self-organize during development. More recently, the model has been incorporated into explanations of the line motion illusion and motion induction experiments (Baloch and Grossberg, 1996a, 1996b). In all these explanations, similar mechanisms of transient cell detection, short-range filtering, competition, and long-range filtering have been used. Thus the total explanatory power of the proposed mechanisms, and the experimental support for them, goes far beyond the data simulated herein.

In order to generate a globally consistent representation of object motion, the Motion BCS model takes account of both unambiguous feature tracking signals and ambiguous signals that are due to the aperture problem. The model differs in several ways from others that have postulated that both feature and non-feature motion signals are involved in motion perception (Castet *et al.*, 1993; Kim and Wilson, 1993; Wilson *et al.*, 1992). First, the model does not process the two types of signals using different mechanisms. Instead, each signal type is processed by the same mechanisms operating at different image locations. Observations that feature tracking signal processing has different characteristics from ambiguous motion processing (Lorenceau and Shiffrar, 1992; Yo and Wilson, 1992) are explained as a result of the long-range process rather than of a separate processing channel.

Feature tracking signals are amplified in the model by anisotropic filtering of transient motion signals and competitive processes. Supportive evidence for such mechanisms has been described for barberpole type motion by Power and Moulden (1992), who showed that aperture widths influenced perceived motion of translating lines. Additional evidence for the importance of anisotropic filtering comes from studies indicating that multiple dot flashes contribute to apparent motion percepts (Nakayama and Silverman, 1984) and that the luminance of, and distance between, dot flashes influences perceived apparent motion between those flashes (Allik, 1992).

The model explains various data by assuming that feature tracking motion signals dominate motion percepts and that deviations from feature tracking directions (Castet *et al.*, 1993; Yo and Wilson, 1992) are the result of incomplete feature signal integration. These explanations are compatible with the short durations used in experimental paradigms that show perceived directions of motion deviating from the feature motion directions. The model is also able to account for data showing that integration processes are affected by non-speed parameters, such as stimulus contrast (Stoner *et al.*, 1990)

The relative durations used in the various simulations are roughly in accord with the relative times utilized in the modelled experiments, as seen in Table 1. The Stone *et al.* (1990) experiment is somewhat discordant with the other data. In order to reproduce similar magnitude biases, the simulations were run for only 1.5 time units, although this experiment utilized longer exposure durations than did the Castet *et al.* (1993) or Yo and Wilson (1992) experiments. Despite the fact that the model made a number of simplifying assumptions for the sake of computational tractability, there are several more probable explanations for this discrepancy, including that Stone *et al.* (1990) used slower speed

Study	Stimulus Duration	Simulation Time	Simulation Duration
Castet *et al.* (1993)	167ms	3	160ms
Yo and Wilson (1992)	40-160ms	3	160ms
Stone *et al.* (1990)	200ms	1.5	80ms

Table 1: Experimental durations and simulation durations for three sets of experiments.

stimuli (2 degrees/second) than did Yo and Wilson (8/9.6 degrees per second) and lower luminances (100 cd/sq.m) than did Castet *et al.* (131 cd/sq. m). Both effects could be compensated by a somewhat longer presentation time. A systematic parametric study that included all the stimuli of the three experiments would be most helpful. For definiteness, a simulation time of 3 units was set equal to 160 ms in Table 1, but this conversion factor must remain uncertain until more systematic parametric data become available.

The perceived motion of coherently moving plaid patterns has a natural explanation using the motion of feature tracking signals. Incoherent plaid motion was attributed to the fact that feature signals are too weak to influence nearby ambiguous signals. Most reports of incoherent plaid motion use sine-wave gratings (e.g., Adelson and Movshon, 1982; Kim and Wilson, 1993) whose feature blobs at the intersection points result in less powerful feature signals than do the clearly defined diamonds located at the intersection points of square-wave gratings, which Lindsey and Todd (1995) have shown to move coherently until after a substantial period of adaptation.

Long-range grouping in the motion BCS model operates using large isotropic input kernels. This assumption is a simplification. Processing of more complex image scenes requires consideration of more complex input kernels, including inputs from segmentations performed by the static form system. Motion integration is affected by the form relationships between locations where feature tracking and ambiguous motion signals are generated. For example, Shiffrar *et al.* (1995) have shown that the perceived motion of barberpole displays was determined primarily by the motion of the line terminators, even in the presence of superimposed moving dots whose movement is also unambiguous. Assuming that the motion signals generated by the moving dots and line terminators are of similar strength, this suggests that the terminators exert a greater influence on the

perceived direction of motion of the lines because they are connected to the line, whereas the dots are spatially isolated from the lines. Similarly, Nakayama and Silverman (1988) showed that terminators attached to a moving curve were more dominant in determining perceived rigidity and that the influence of disconnected terminators was dependent on their distance from the curve. Baloch and Grossberg (1996a, 1996b) and Francis and Grossberg (1996a) have modeled how such form-motion interactions may occur.

Appendix: Network Equations and Parameters

The network is defined by differential equations that specify the time-varying activity or membrane potential of individual neurons or populations of neurons. These cells are typically controlled by a membrane or shunting equation (Douglas *et al.*, 1995; Grossberg, 1982; Hodgkin, 1964) which performs leaky integration of its inputs:

$$\frac{dx}{dt} = \alpha(-x + (\beta - \gamma x)E - (\delta + \epsilon x)I) \tag{1}$$

In equation (1), the activity, x, of a cell is driven by excitatory input and inhibitory input I. The integration rate of the cell is given by parameter α. Parameters γ and ϵ determine whether the cell responds to its inputs additively ($\gamma = \epsilon = 0$) or whether they are shunted. In the case of shunting input, the parameters $\beta\gamma^{-1}$ and $-\delta\epsilon^{-1}$ determine the maximal and minimal activity levels of the cell, respectively. Shunting automatically gain-controls cell responses to the inputs. Cell activity then tends to compute a Weber-law modulated contrast ratio of the inputs (Grossberg, 1980, 1983).

Each cell activity is denoted by a variable whose letter indicates the type of cell (e.g., transient, short-range filter, etc.) and whose subscript indicates the cell's spatial position. Superscripts denote directional tunings and scales. Thus, f_{ij}^{ds} denotes the activity of a short-range filter of directional preference d, scale s, located at spatial position (i, j). Notation $[x]^+ = \max(x, 0)$ denotes half-wave rectification and $[x]^t = [x - t]^+$ half-wave rectification with threshold at t. Notation $\|S\|$ indicates the size of a set S.

In all simulations, 16 directions and 4 scales were used. All inputs moved at a single constant speed. Equations were numerically integrated using Euler's method with a time step of 0.01. In all cases, the same parameter set was used. Simulations varied only in their size and input.

Within each direction, the network is functionally identical to the speed-sensitive model described in Chey *et al.* (1994) and so retains all the speed-tuning characteristics of that model. Figure 19 depicts all the model operations.

Level 0: Input. Stimulus forms with prescribed luminances moving at different speeds generate model inputs. See below for details.

Figure 19

Level 1: Transient Cell Network.

Change-Sensitive Receptors. The activity a_{ij} of change-sensitive receptors responds with a pulse of fixed duration (scaled to equal 1) and amplitude η to luminance variations of contrast η within their receptive fields. In simulations involving single moving lines, only ON responses to luminance increments are considered:

$$a_{ij}^{ON} = \begin{cases} 0 & \text{if } (t - t_{ON}) \geq 1 \\ \eta & \text{if } 0 < (t - t_{ON}) < 1 \end{cases}, \tag{2}$$

where t_{ON} is the most recent time at which luminance increased within the receptive field of the cell. In the plaid simulations, OFF responses to luminance decrements are also simulated:

$$a_{ij}^{OFF} = \begin{cases} 0 & \text{if } (t - t_{OFF}) \geq 1 \\ \eta & \text{if } 0 < (t - t_{OFF}) < 1 \end{cases}, \tag{3}$$

where t_{OFF} is the most recent onset time of a luminance decrement.

Transient Cells. Transient cell activities b_{ij} perform spatial and time averaging of receptor responses over a fixed size spatial kernel:

$$\frac{db_{ij}}{dt} = -b_{ij} + (1 - b_{ij}) \sum_{(x,y) \in B_{ij}} a_{xy}, \tag{4}$$

where set B_{ij} consists of 10 adjacent change-sensitive receptor positions. Sets B_{ij} are defined so that each transient cell draws from a non-overlapping set of receptors. Transient cell activity is shunted: its response rate increases with total receptor activity in set B_{ij}, but its amplitude is bounded by 1.

Directional Interneurons. Directional interneuron activity c_{ij}^{d} time-averages transient cell output (Figure 19). Each interneuron helps to create directional transient signals and so is assigned a direction d:

$$\frac{dc_{ij}^{d}}{dt} = -c_{ij} + b_{ij} - 10[c_{IJ}^{D}]^{+}. \tag{5}$$

Each interneuron receives excitatory input, b_{ij}, from a transient cell and inhibitory input, c_{IJ}^{D}, from an adjacent inhibitory interneuron. The coordinates (I, J) are offset by one unit from (i, j) in direction d. Mutual competition between interneurons prevents a directional interneuron of opposite direction from inhibiting directional transient cells of the correct direction at high speeds. The inhibitory input to an interneuron is therefore given by c_{IJ}^{D}, where D is the opposite direction from d given by $((d + 8) \bmod 16)$. Inhibition is set to be stronger than excitation in (5).

Directional Transient Cells. Directional transient cell activities e_{ij}^{d} with direction d receive excitatory input b_{ij} from transient cells, which is vetoed by directional interneuron activity c_{IJ}^{D}:

$$\frac{de_{ij}^{d}}{dt} = 10 \left(-e_{ij}^{d} + b_{ij}^{d} - 10[c_{IJ}^{D}]^{+} \right). \tag{6}$$

Level 2: Short-Range Filters. Short-range spatial filter activity f_{ij}^{sd} performs space and time averaging of directional transient cell responses. Each filter has a scale s, which determines receptive field size, and a direction d, which determines receptive field orientation and the directionally selective transient cells that activate it:

$$\frac{df_{ij}^{sd}}{dt} = 10 \left(-f_{ij}^{sd} + \sum_{(I,J) \in F_{ij}^{sd}} e_{IJ}^{sd} \right). \tag{7}$$

Set F_{ij}^{sd} includes $2s + 1$ receptors aligned in the orientation coincident with direction d and centered around position (i, j). Each short-range filter output is thresholded by an amount that increases linearly with filter size before being spatially blurred. The output is given

by:

$$g_{ij}^{ds} = \sum_{(I,J)\in G_{ij}^{d}} e^{-(i-I)^2-(j-J)^2} [f_{IJ}^{ds}]^{\frac{3s}{2}}. \tag{8}$$

Set G_{ij}^{d} contains five filters offset by distances ranging from -2 to 2 around position (i, j) in direction d. This short-range Gaussian filter reduces incomplete activation of spatial competitive kernels in the next stage.

Level 3: Competition Network.

Intra-Scale Competition. The activity of an intra-scale competitive cell h_{ij}^{ds} is determined by a feedforward center-surround competition with nearby short-range filters:

$$\frac{dh_{ij}^{ds}}{dt} = 10\left(-h_{ij}^{ds} + \frac{\sum_{(I,J)\in E_{ij}^{d}} g_{IJ}^{ds}}{\|E_{ij}^{d}\|} - \frac{\sum_{(I,J)\in I_{ij}^{d}} g_{IJ}^{ds}}{\|I_{ij}^{d}\|}\right). \tag{9}$$

Set E_{ij}^{d} defines the spatial locations that form the excitatory kernel of the cell. It includes 5 thresholded short-range filters offset from position (i, j) by -2 to 2 in direction d. Set I_{ij}^{d} defines the inhibitory kernel, which includes 6 filters offset by distances ranging from -5 to -3 and 3 to 5 in direction d.

Inter-Scale Competition. The inter-scale competition activity k_{ij}^{ds} refines speed tuning using a shunting equation. Excitatory input at scale s is given by the rectified activity $[h_{ij}^{ds}]^{+}$ of an intra-scale competitive cell of the same direction and scale raised to a power. Inhibitory input sums the rectified activities of all other scales raised to the same power:

$$\frac{dk_{ij}^{ds}}{dt} = -k_{ij}^{ds} + (1 - k_{ij}^{ds})\left([h_{ij}^{dt}]^{+}\right)^3 - (1 + k_{ij}^{ds})\frac{\sum_{t\neq s}\left([h_{ij}^{dt}]^{+}\right)^3}{\|\{t \neq s\}\|}. \tag{10}$$

The power function sharpens the tuning of each cell.

Inter-Directional Competition. The inter-directional competition activity l_{ij}^{ds} occurs at each position across direction. Excitatory input comes from the like-directional inter-scale competitive cell and inhibitory input from all other directions across all scales weighted by their mutual distance in directional space; namely,

$$\frac{dl_{ij}^{ds}}{dt} = 10\left(-l_{ij}^{ds} + 10[k_{ij}^{ds}]^{+} - \frac{1}{10} l_{ij}^{ds} \sum_{D\neq d}\sum_{t} \mid D - d \mid [k_{ij}^{Dt}]^{+}\right). \tag{11}$$

The shunting of inhibitory input to these cells ensures that relative activities across scales are maintained (Grossberg, 1980).

Level 4: Long-Range Filter. Long-range filter activity m_{ij}^{ds} space and time averages inter-directional competitive cell outputs within an elongated, directional receptive field. It also receives inhibitory feedback from long-range directional grouping cells that choose among the competing filter directions:

$$\frac{dm_{ij}^{ds}}{dt} = -m_{ij}^{ds} + \frac{\sum_{(x,y)\in M_{ij}^{d}} [l_{xy}^{ds}]^{+}}{\|M_{ij}^{d}\|} - 3(1 + m_{ij}^{ds}) \sum_{D\neq d} [o^{D}]^{+}. \tag{12}$$

Set M_{ij}^d contains 11 cells offset by distances from -5 to 5 in direction d from position (i,j). In related applications, the long-range filter is assumed to have a Gaussian profile (Grossberg and Rudd, 1992), which is omitted here for computational simplicity.

Level 5: Directional Grouping and Priming. Each grouping cell activity n_{ij}^d sums activity from long-range filters and competes with other grouping cells to select a winning direction, which is then used to select consistent long-range filter activities. Each such cell sums filter activities over a wide spatial range, a small range of directions, and all scales:

$$\frac{dn_{ij}^d}{dt} = \frac{1}{5}\left(-n_{ij}^d + (1-n_{ij}^d)N_{ij}^d - 10\sum_{D\neq d}[n_{ij}^D]^+\right), \tag{13}$$

where the excitatory input is given by:

$$N_{ij}^d = \frac{\sum_{(I,J)\in O_{ij}}\sum_D\sum_s X(\mid d-D\mid)\Psi(\mid D\mid)\left([m_{IJ}^{Ds}]^+\right)^2}{\|O_{ij}\|}. \tag{14}$$

Function X selects the range of directions over which the grouping cell summates. It is set equal to 1 for $|d-D|=0$, $\frac{1}{2}$ for $|d-D|=1$, and 0 otherwise. Set O_{ij} determines the domain of spatial averaging. In most simulations, it is set large enough to cover the entire image. The activity m_{IJ}^{Ds} of each long-range filter cell that excites a grouping cell is rectified and raised to the power 2. Every grouping cell competes with every other grouping cell at the same spatial position. Shunting ensures that each cell's activity can never grow above 1. This limits the total possible feedback strength and ensures that feedback strength cannot become excessively large.

Term $\Psi(D)$ controls the adaptation level of the long-range filter cell that signals motion in direction. Adaptation occurs from sustained activation of long-range filters with directional preference over a long time period. It reduces $\Psi(D)$, where D is the winning grouping cell direction. Term $\Psi(D)$ is initially set to 1 for all directions. To simulate adaptation, $\Psi(D)$ is set to a fractional value. To find the value at which incoherent motion is first observed, $\Psi(D)$ is initially set to one, then reduced by decrements of 0.05 until incoherent motion is observed; that is, until no winning direction is established at the grouping cells.

Line Inputs. Inputs consist of a temporal sequence of luminance patterns. Any change of luminance in these patterns from one time unit to the next triggers change-sensitive receptor activity. Three types of inputs were simulated: lines, lines moving behind apertures, and plaids. Simulations of moving lines specify a line length, orientation and speed. Lines move horizontally with speeds adjusted so that the horizontal component of line motion is the same. The line moves for a fixed time. An image array sufficiently large to contain the line at its start and end position is chosen.

To project a continuous line onto a two-dimensional grid of pixels, certain sampling problems must be faced, since in a regular two-dimensional Cartesian grid vertically and horizontally oriented short-range filters are not the same length as diagonally oriented filters. To overcome this problem, simulations calculated transient cell responses at offsets from short-range filter positions to ensure that all filters of the same scale were of the same size. Since this results in a large number of transient cells, simulations were simplified to

reduce processing time.

Transient cell responses were pre-calculated for each simulation, since their responses can be determined ahead of time from input speed. These responses are used to determine transient cell responses based on the time since the leading edge (for ON responses) or trailing edge (for OFF responses) of an input passed a given image location. These times are easily calculated from the lines' starting position, speed and orientation. Line simulations utilized only ON responses and continued for 4 time units. Line directions were sampled at time step 3 as described below. Four line orientations were used: 22.5, 45, 67.5 and 90 degrees from horizontal and three line lengths: 5, 13, and 26 units.

Lines Behind Aperture Inputs. Barberpole simulations used a line moving behind an aperture. Line terminators were never present in the image. The line started in the top left hand corner of the image and moved until it exited the bottom right hand corner. To simulate diagonal barberpole motion, the grouping cells spanned the length of the moving line. In any simulation in which grouping cells span the entire image, the activity of single grouping cell was calculated for each directional preference and used to reflect network activity,

Plaid Inputs. Plaid inputs were formed from two overlapping component lines. Line speeds were chosen so that, regardless of component orientations, the plaid moved horizontally at a given speed. A complete plaid simulation input is shown in Figure 13, including responses of ON and OFF directional transient cells calculated in the same way as for line inputs. The quantitative plaid simulations used only the ON plaid responses to the leading edge of a plaid to reduce processing time.

Three different relative component orientations were utilized when simulating coherent and incoherent plaid motion: 22.5, 45 and 67.5 degrees from horizontal. For each component orientation, a series of simulations were run in which $\Psi(D)$ was altered, where D designates rightward motion. These simulations started with $\Psi(D)$ at 1, at which all plaids moved coherently, and reduced it by units of .05 until incoherent motion was observed.

Changes in plaid component contrasts were simulated by altering receptor response magnitudes in the same way as was done for the one-dimensional simulations. These simulations assumed a base level of luminance, corresponding to receptor response magnitude, and then calculated a set of response magnitudes that corresponded to the simulated series of contrast ratios. These ratios ranged from $2^{0.5}$ to $2^{3.5}$. For each contrast ratio, a simulation was run and an image direction sampled at time unit 1.5.

Outputs. Energy is defined at each location within each direction by summing the long-range filter outputs from equation (12) as follows:

$$y_{ij}^{d} = \sum_{s \in S} m_{ij}^{ds}, \tag{15}$$

and then summing these responses across all directions:

$$y_{ij} = \sum_{d} y_{ij}^{d}, \tag{16}$$

The speed measure s_{ij}^{d} is derived from the long-range filter outputs within each direction

and position:

$$s_{ij}^{d} = \frac{\sum_{s \in S} m_{ij}^{ds} s}{y_{ij}^{d}} \tag{17}$$

In order to interpret motion signals as vectors, speed measures are combined over different directions. For each direction, a motion vector is defined as the vector extended in that direction with a magnitude equal to the speed measure:

$$v_{ij}^{d} = (s_{ij}^{d} \cos d,\ s_{ij}^{d} \sin d). \tag{18}$$

The sum of these vectors is the perceived motion at that location:

$$v_{ij} = \sum_{d} v_{ij}^{d}. \tag{19}$$

To determine perceived direction and/or speed of a moving object, a weighted average of motion vectors was taken across the image. Outputs were eliminated from positions whose total energy across all scales and directions was less than some threshold (set to 1 in all simulations):

$$w_{ij} = \begin{cases} (0,0) & \text{if } (y_{ij} < 1) \\ v_{ij} & \text{if } (y_{ij} \geq 1 \end{cases}. \tag{20}$$

The motion values w_{ij} were then multiplied by the energy y_{ij}, summed across position, and then divided by the total energy. The resulting total motion vector

$$v = \frac{\sum_{(i,j)} y_{ij} w_{ij}}{\sum_{(i,j)} y_{ij}} \tag{21}$$

was used to measure object speed and direction, as in Figure 10.

Parameters.

In a neural model such as ours, model complexity is determined by the number of processes (as in Figure 4), not by the number of parameters. In addition, model parameters do not just fit prescribed form-factors. Rather, data are derived as emergent properties of network interactions. Finally, one needs to assess how many data may be rationalized by a single set of processes. This being said, in equations (2) and (3), one parameter scales input amplitude. In the dynamical equations (4)–(14), most baseline processing rates were chosen equal to 1 for simplicity. All fast rates were chosen equal to 10, again for simplicity, since our goal herein was qualitative rather than quantitative data fits. The slower rate of directional grouping and priming in (13) was set to 1/5. There is a total of 11 non-unity parameters in these equations. In addition, 7 parameters determine the sizes of the receptive fields, and there were 16 directions and 4 scales. These parameters are robust just so long as reasonable relative sizes are observed; e.g., long-range filters have larger scales than short-range filters. Most of the directions were not critical in fitting the data curves. They illustrate how the model solves the aperture problem. Using these parameters, 33 data points were fit in Figures 2, 16, 17, and 18. Figures 5, 6, 7, and 9–15 simulated many hundreds of data points to describe the speed-sensitive tuning curves at the various model stages, and the temporal evolution of the motion capture process over entire vector fields of motion direction and speed vectors in response to visual forms moving through time

under various experimental conditions. The same processes have also been used to explain many other motion data sets, as reviewed in the Discussion.

9 References

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, 284–299.

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

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

Albright, T.D. and Desimone, R. (1987). Local precision of visuotopic organization in the middle temporal area (MT) of the macaque. *Experimental Brain Research*, **65**, 582–592.

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.

Allik, J. (1992). Competing motion paths in sequences of random dot patterns. *Vision Research*, **32**, 157–165.

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

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

Ariel, M. and Daw, N.W. (1982). Pharmacological analysis of directionally sensitive rabbit retinal ganglion cells. *Journal of Physiology*, **324**, 161–185.

Assad, J.A. and Maunsell, J.H.R. (1995). Neural correlates of inferred motion in primate posterior parietal cortex. *Nature*, **373**(6514), 518–521.

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

Baloch, A. and Grossberg, S. (1996b) A neural model of high-level motion processing: Line motion and formotion dynamics. Technical Report CAS/CNS-96-20. Boston, MA: Boston University.

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

Barlow, H.B. and Levick, W.R. (1965). The mechanism of directionally selective units in rabbit's retina. *Journal of Physiology*, **178**, 477–504.

Beck, J., Prazdny, K., and Rosenfeld, A. (1983). A theory of textural segmentation. In Beck, J., Hope, B., and Rosenfeld, A., editors, **Human and Machine Vision**. New York: Academic Press, pp. 1-38.

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.

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

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.

Brown, K. and Sekuler, R. (1980). Models of stimulus uncertainty in motion perception. *Psychological Review*, **87**, 435–469.

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

Carpenter, G.A. and Grossberg, S. (1987). A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics and Image Processing*, **37**, 54–115.

Carpenter, G.A. and Grossberg, S. (Eds.) (1991). **Pattern Recognition by Self-Organizing Neural Networks**. Cambridge, MA: MIT Press.

Carpenter, G. A. and Grossberg, S. (1993). Normal and amnesic learning, recognition, and memory by a neural model of cortico-hippocampal interactions. *Trends in Neurosciences*, **16**, 131–137.

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

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

Cohen, M. and Grossberg, S. (1986). Neural dynamics of speech and language coding: Developmental programs, perceptual grouping, and competition for short term memory. *Human Neurobiology*, **5**, 1-22.

de Bruyn, B. and Orban, G.A. (1988). Human velocity and direction discrimination measured with random dot patterns. *Vision Research*, **28**(12), 1323–1335.

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

Douglas, R.J., Koch, C., Mahowald, M., Martin, K.A.C., and Suarez, H.H. (1995). Recurrent excitation in neocortical circuits. *Science*, **269**, 981–985.

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. (1966). The contrast sensitivity of retinal ganglion cells of the cat. *Journal of Physiology, London*, **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.

Ferrera, V.P. and Wilson, H.R. (1987). Direction specific masking and the analysis of motion in two dimensions. *Vision Research*, **27**(10), 1783–1796.

Ferrera, V.P. and Wilson, H.R. (1990). Perceived direction of moving two-dimensional patterns. *Vision Research*, **30**(2), 273–287.

Ferrera, V.P. and Wilson, H.R. (1991). Perceived speed of moving two-dimensional patterns. *Vision Research*, **31**(5), 877–893.

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*, **26**, 543–567.

Francis, G., Grossberg, S., and Mingolla, E. (1994). Cortical dynamics of feature binding and reset: Control of visual persistence. *Vision Research*, **34**(8), 1089–1104.

Ganz, L. (1984). Visual cortical mechanisms responsible for direction selectivity. *Vision Research*, **24**(1), 3–11.

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

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

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.

Govindarajan, G., Grossberg, S., Wyse, L., and Cohen, M. (1994). A neural network model of auditory scene analysis and source segregation. Technical Report CAS/CNS-TR-94-039, Boston, MA: Boston University.

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

Grossberg, S. (1973). Contour enhancement, short term memory, and constancies in reverberating neural networks. In **Studies of Mind and Brain**, Chapter 8, pp. 334–378. Dordrecht: Reidel Press.

Grossberg, S. (1976). Adaptive pattern classification and universal recoding I: Parallel development and coding of neural features. *Biological Cybernetics*, **23**, 121–134.

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

Grossberg, S. (1982). **Studies of mind and brain**. Boston: Reidel Press.

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

Grossberg, S. (1986). The adaptive self-organization of serial order in behavior: Speech, language, and motor control. In E.C. Schwab, and H.C. Nusbaum, H.C. (Eds.), **Pattern Recognition by Humans and Machines, Volume 1: Speech Perception**. New York: Academic Press, pp.187-294.

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

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

Grossberg, S. (1995). The attentive brain. *American Scientist*, **83**, 438–449.

Grossberg, S. (1996). How is a moving target continuously tracked behind occluding cover? In T. Watanabe (Ed.) **High level motion processing**, Cambridge, MA: MIT Press.

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

Grossberg, S., Boardman, I., and Cohen, M. (1994). Neural dynamics of variable-rate speech categorization. Technical Report CAS/CNS-TR-94-038, Boston, MA: Boston University. *Journal of Experimental Psychology: Human Perception and Performance*, in press.

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

Grossberg, S. and Mingolla, E. (1985). Neural dynamics of perceptual grouping: Textures, boundaries and emergent segmentations. *Perception and Psychophysics*, **38**(2), 141–171.

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 Pessoa, L. (1995). Texture segregation, surface representation and figure-ground separation. Submitted for publication.

Grossberg, S. and Rudd, M.E. (1989). A neural architecture for visual motion perception: Group and element apparent motion. *Neural Networks*, **2(6)**, 421–450.

Grossberg, S. and Rudd, M.E. (1992). Cortical dynamics of visual motion perception: Short-range and long-range apparent motion. *Psychological Review*, **99**(1), 78–121.

Heggelund, P. (1984). Direction asymmetry by moving stimuli and static receptive field plots for simple cells in cat striate cortex. *Vision Research*, 24(1):1316.

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

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

Hildreth, E.C. (1984). **The Measurement of Visual Motion**. Cambridge, MA: MIT Press.

Hodgkin, A.L. (1964). **The conduction of the nervous impulse**. Liverpool, UK: Liverpool University.

Hochstein, S. and Shapley, R. (1976a). Quantitative analysis of retinal ganglion cell classifications. *Journal of Physiology, London*, **262**, 237–264.

Hochstein, S. and Shapley, R. (1976b). Linear and nonlinear spatial subunits in Y cat retina ganglion cells. *Journal of Physiology, London*, **262**, 265–284.

Horn, B.K.P. and Schunck, B.G. (1981). Determining optical flow. *Artificial Intelligence*, **17**, 185–203.

Hubel, D.H. and Wiesel, T.N. (1959). Receptive fields of single cells in the cats 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.

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

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

Kohonen, T. (1988). Self-Organization and Associative Memory. Berlin: Springer-Verlag.

Kolers, P.A. (1972). Aspects of Motion Perception. Oxford, UK: Pergamon Press.

Kuffler, S.W. (1953). Discharge patterns and functional organization of the mammalian retina. *Journal of Neurophysiology*, **16**, 37–68.

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

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

Lorenceau, J. and Shiffrar, M. (1992). The influence of terminators on motion integration across space. *Vision Research*, **32**(2), 263–273.

Lorenceau, J., Shiffrar, M., Wells, N., and Castet, E. (1993). Different motion sensitive units are involved in recovering the direction of moving lines. *Vision Research*, **33**, 1207–1217.

Marshall, J.A. (1990). Self-organizing neural networks for perception of visual motion. *Neural Networks*, **3**(1), 45–74.

Mashhour, M. (1964). **Psychophysical relations in the perception of velocity**, Stockholm: Almquist and Wiksell.

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 duration, speed, and orientation. *Journal of Neurophysiology*, **49**(5), 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**(5), 1148-1167.

Mikami, A., Newsome, W.T., and Wurtz, R.H. (1986). Motion selectivity in macaque visual cortex. II. Spatiotemporal range of directional interactions in MT and V1. *Journal of Neurophysiology*, **55**(6), 1328–1339.

Mingolla, E., Todd, J.T., and Norman, J.F. (1992). The perception of globally coherent motion. *Vision Research*, **32**(6), 1015–1031.

Movshon, J.A., Adelson, E.H., Gizzi, M.S., and Newsome, W.T. (1985). The analysis of moving visual patterns. In C. Chagas, R. Gattass, and C. Gross, (Eds.), **Pattern Recognition Mechanisms**, Volume 11 of **Experimental Brain Research Supplementum**. Berlin: Springer-Verlag, pp.117-151.

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

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

Nakayama, K. and Silverman, G.H. (1988). The aperture problem II: Spatial integration of velocity information along contours. **Vision Research**, **28**, 747–753.

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

Nogueira, C.A.M., Mingolla, E., and Grossberg, S. (1993). Computation of first order and second order motion by a model of magnocellular dynamics. *Investigative Opthalmology and Visual Science*, **34**(4), 1029.

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

Orban, G.A., de Wolf, J.D. and Maes, H. (1984). Factors influencing velocity coding in the human visual system. *Vision Research*, **24**(1), 33–39.

Orban, G.A., Kennedy, H., and Bulier, J. (1986). Velocity sensitivity and direction selectivity of neurons in areas V1 and V2 of the monkey: Influence of eccentricity. *Journal of Neurophysiology*, **56**(2), 462–480.

Payne, B.R., Berman, N., and Murphy, E.H. (1980). Organization of direction preferences in cat visual cortex. *Brain Research*, **211**, 445–450.

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

Power, R.P. and Moulden, B. (1992). Spatial gating effects on judged motion of gratings in apertures. *Perception*, **21**, 449–463.

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

Rodman, H.R. and Albright, T.D. (1987). Coding of visual stimulus velocity in area MT of the macaque. *Vision Research*, **27**(12), 2035–2048.

Rubin, N. and Hochstein, S. (1993). Isolating the effect of one-dimensional motion signals on the perceived direction of moving two-dimensional objects. *Vision Research*, **33**(10), 1385–1396.

Schiller, P.H. (1982). Central connections on the retinal ON- and OFF-pathways. *Nature*, **297**, 580–583.

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

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

Sereno, M.E. (1989). Learning the solution to the aperture problem for pattern motion with a Hebb rule. In **Advances in Neural Information Processing Systems 2**. San Diego, CA: Morgan-Kauffman Publishers, pp.468-476.

Sereno, M.E. (1993). **Neural Computation of Pattern Motion**. Cambridge, MA: MIT Press.

Shiffrar, M., Li, X., and Lorenceau, J. (1995). Motion integration across differing image features. Submitted for publication.

Snowden, R.J., Treue, S., and Andersen, R.A. (1992). The response of neurons in areas V1 and MT of the alert rhesus monkey to moving random dot patterns. *Experimental Brain Research*, **88**, 389–400.

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.

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

Stone, L.S. and Thompson, P. (1992). Human speed perception is contrast dependent. *Vision Research*, **32**(8), 1535–1549.

Stone, LS., Watson, A.B., and Mulligan, J.B. (1990). Effect of contrast on the perceived direction of a moving plaid. *Vision Research*, **30**(7), 1049–1067.

Stoner, G.R. and Albright, T.D. (1992). The influence of foreground/background assignment on transparency and motion coherency in plaid patterns. *Investigative Opthalmology and Visual Science*, **33**(5), 1050.

Stoner, G.R., Albright, T.D., and Ramachandran, V.S. (1990). Transparency and coherence in human motion perception. *Nature*, **344**, 153–155.

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

Thompson, P. (1982). Perceived rate of movement depends on contrast. *Vision Research*, **22**, 377–380.

Torre, V. and Poggio, T. (1978). A synaptic mechanism possibly underlying directional selectivity to motion. **Proceedings of the Royal Society of London, B**, **202**, 409–416.

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

Trueswell, J.C. and Hayhoe, M.M. (1993). Surface segmentation mechanisms and motion perception. *Vision Research*, **33**(3), 313–328.

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

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

Tynan, P.D. and Sekuler, R. (1982). Motion processing in peripheral vision: Reaction time and perceived velocity. *Vision Research*, **22**, 61–68.

Ullman, S. (1981). Analysis of visual motion by biological and computer systems. *IEEE Computer*, **14**, 57–69.

Ungerleider, L.G. and Mishkin, M. (1982). Two cortical visual systems. In D.J. Ingle, M.A. Goodale, and R.J.W. Mansfield (Eds.), **Analysis of visual behavior**. Cambridge, MA: MIT Press, pp.549–585.

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 der Malsburg, C. (1973). Self-organization of orientation selective fields in the striate cortex. *Kybernetik*, **14**, 85–100.

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.

Watamaniuk, S.N.J., Grzywacz, N.M., and Yuille, A.L. (1993). Dependence of speed and direction perception on cinematogram dot density. *Vision Research*, **33**(5/6), 849–859.

Wertheimer, M. (1961). Experimentelle Studienber das Sehen von bewegung. In T. Shipley (Ed.), **Classics in Psychology**. Philosophical Library. (Original work published 1912).

Williams, D.W. and Sekuler, R. (1984). Coherent global motion percepts from stochastic local motions. *Vision Research*, **24**, 55–62.

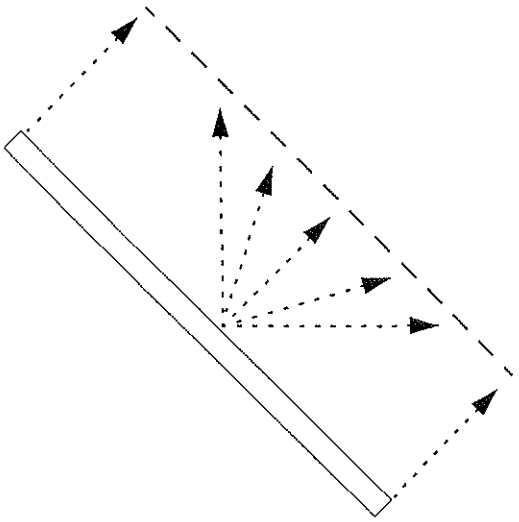
Wilson, H.R., Ferrera, V.P., and Yo, C. (1992). A psychophysically motivated model for two-dimensional motion perception. *Visual Neuroscience*, **9**, 79–97.

Yo, C. and Wilson, H.R. (1992). Perceived direction of moving two-dimensional patterns depends on duration, contrast and eccentricity. *Vision Research*, **32**(1), 135–147.

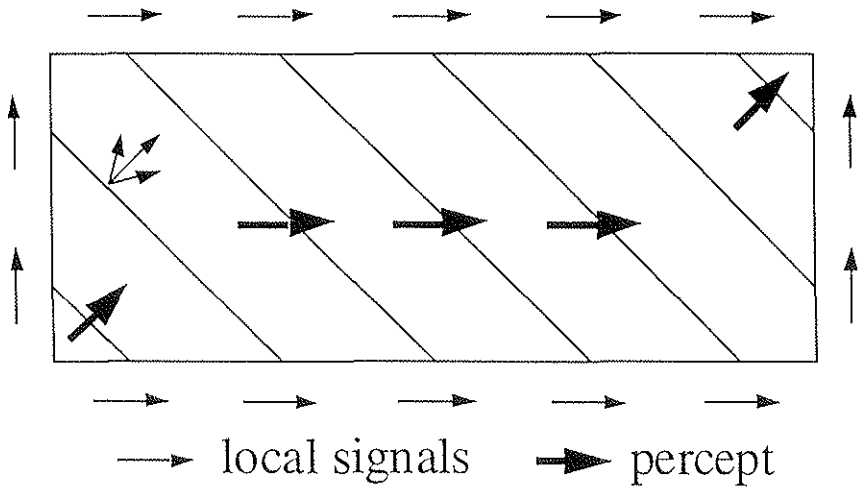
Yuille, A.L. and Grzywacz, N.M. (1988). A computational theory for the perception of coherent visual motion. *Nature*, **333**, 71–74.

Zeki, S.M. (1974). Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *Journal of Physiology (London)*, **236**, 546–573.

Figures



(a)



(b)

Figure 1. Feature tracking and motion capture: Often only a small subset of image features (e.g. lines ends) generate unambiguous motion direction signals due to aperture ambiguity and image or detector noise. We model how these unambiguous locations generate feature tracking signals that are used to capture ambiguous signals. (a) moving line; (b) barberpole illusion.

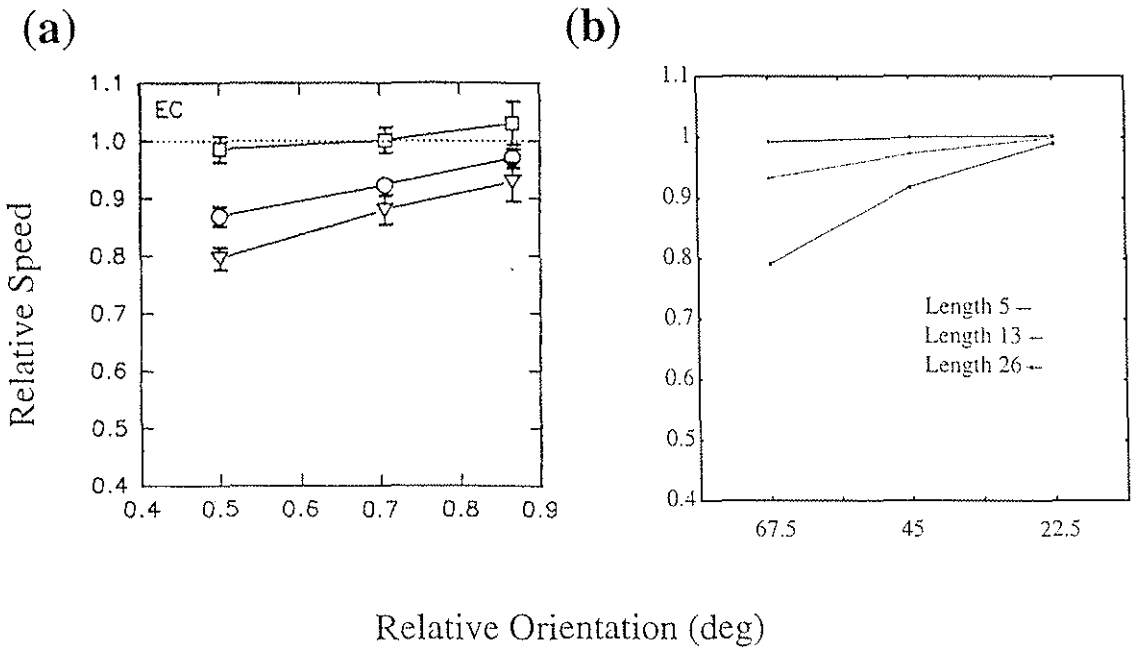


Figure 2. Effects of line length and orientation on perceived speed of horizontally moving lines. Relative perceived speed for three different line orientations and lengths are shown as percentages of the perceived speed of a vertical line of the same length. Figure (a) shows data from Castet *et al.* (1993, p. 1925). Each data line corresponds to a different line length (0.21, 0.88 and 1.76 degrees). The horizontal axis shows the ratio of the speed normal to the line's orientation relative to the actual translation speed. The three data points for each line length correspond to line angles of 60, 45 and 30 degrees from vertical respectively. The horizontal dotted line indicates a veridical speed perception, results below this line indicate a bias towards the perception of slower speeds. Figure (b) shows simulation results, also for three lengths and orientations. In both cases, perceived relative speed decreases with line length, and angle from vertical. Simulated lines use slightly different orientations than in the experiments so that the simulated input conforms to the Cartesian grid.

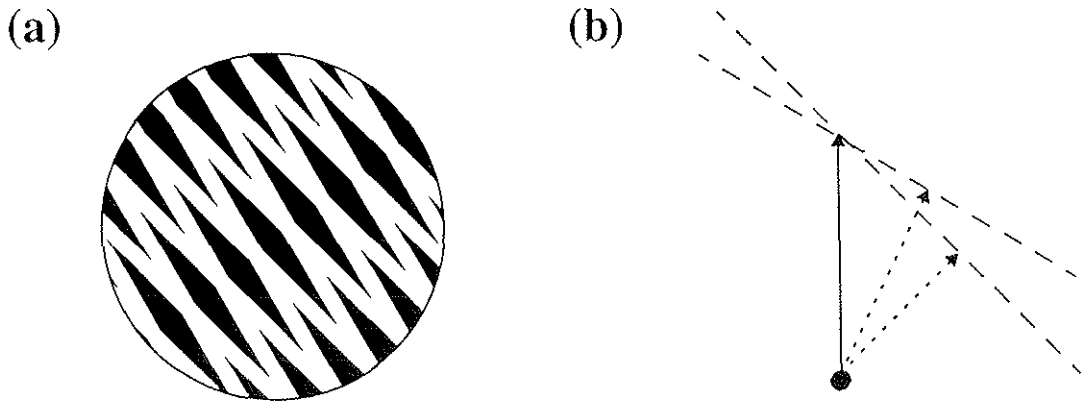


Figure 3. (a) In Type 2 plaids, the IOC solution lies outside the arc formed by the directions normal to the components. In (b) the IOC solution predicts motion upwards whereas the vector sum of the component directions lies between the directions of the two components.

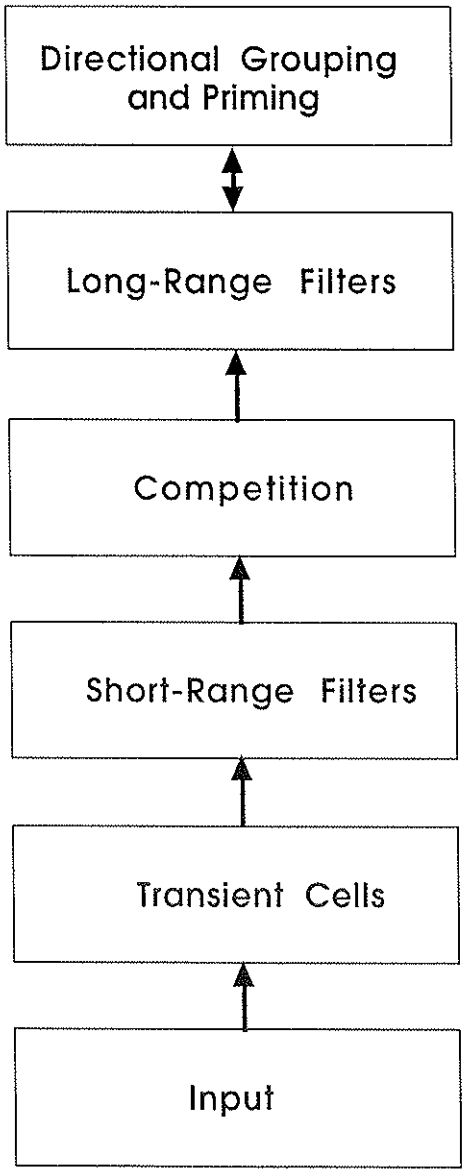


Figure 4. Model processing stages. See text for details.

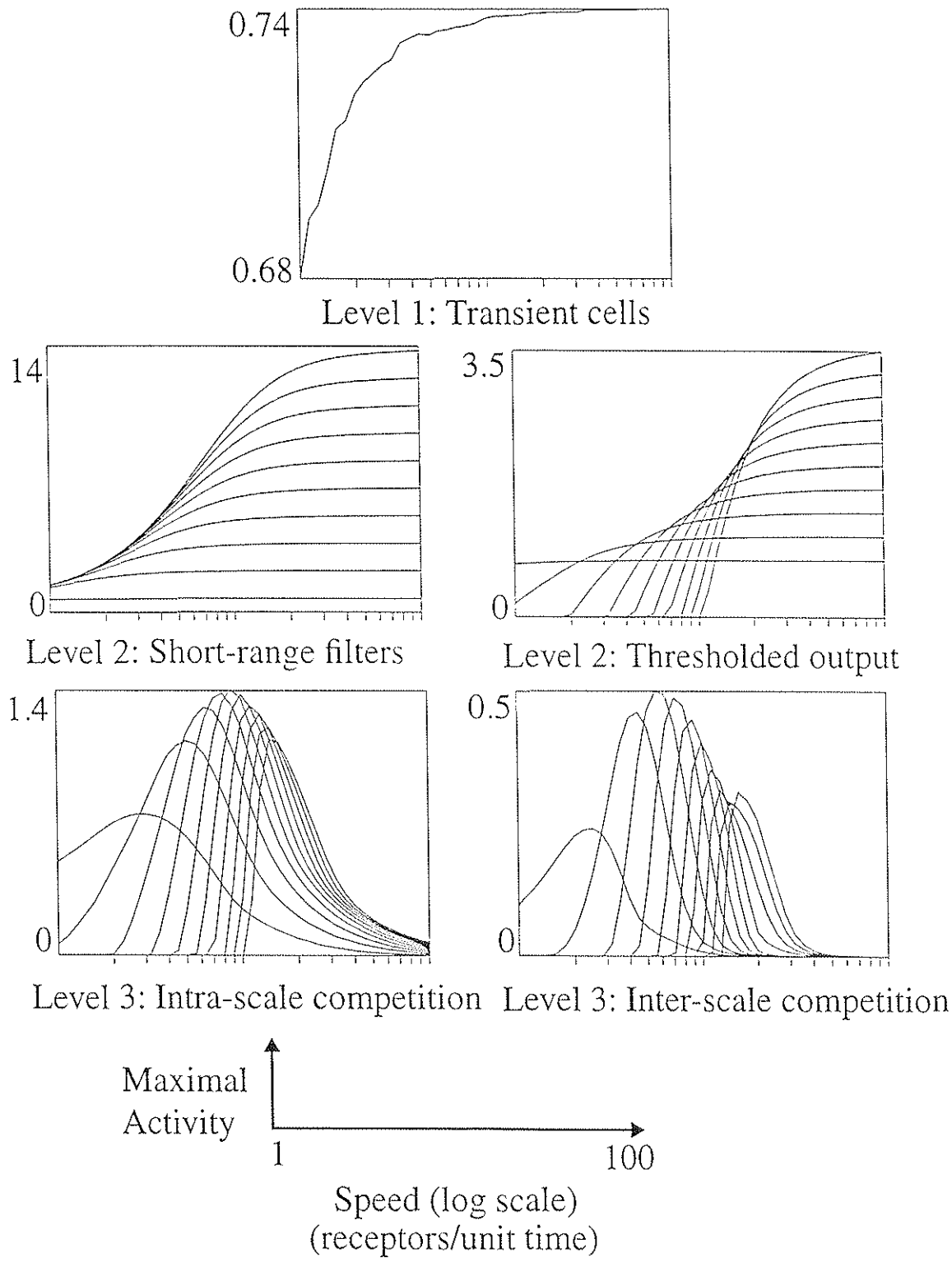
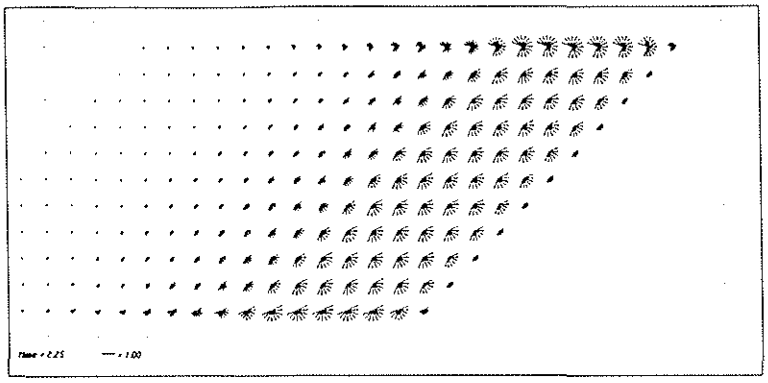
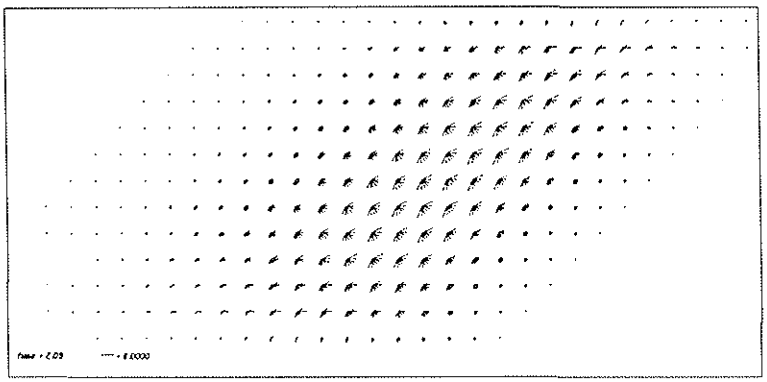


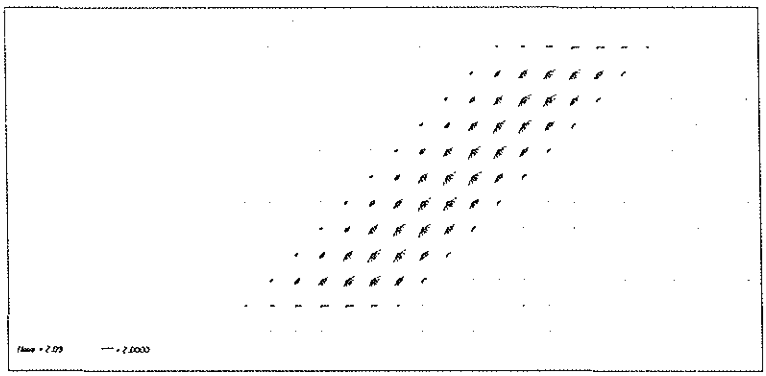
Figure 5. Maximal response of 1-D model cells to a range of simulated speeds. For levels where there are multiple spatial scales at each position, activities from different scales are shown as different curves superimposed on the same plot. The smaller scales tend to respond less vigorously to fast speeds. See text for details.



(a) Directionally sensitive transient cells

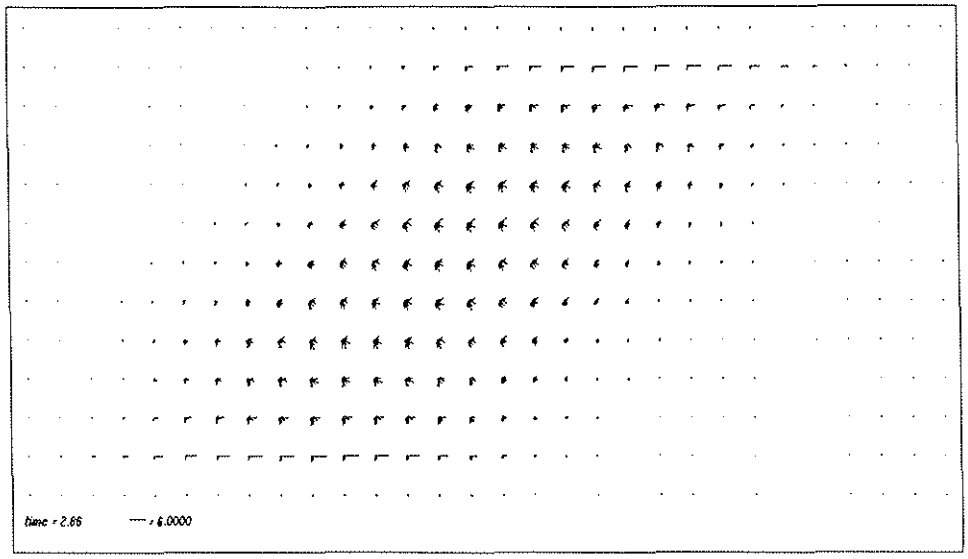


(b) Short-range filters

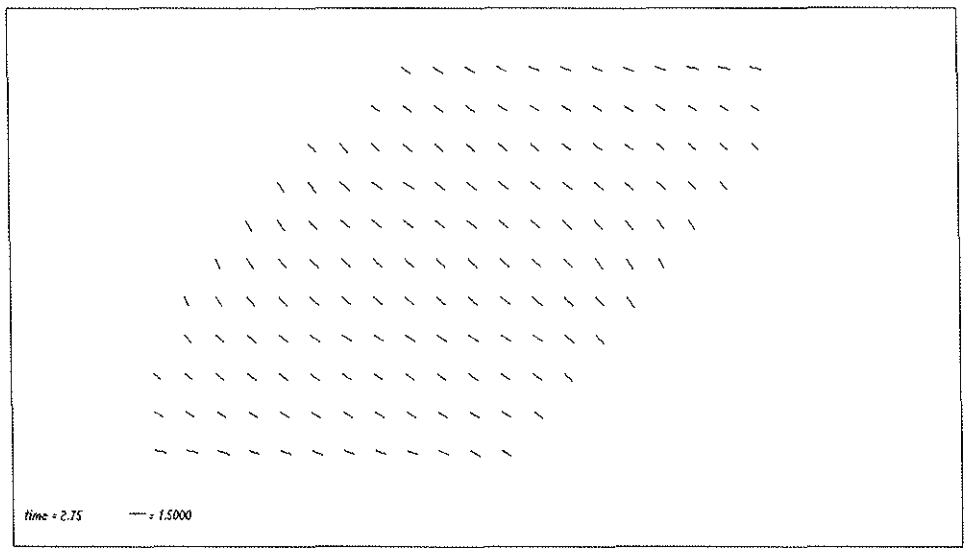


(c) Thresholded short-range filters

Figure 6. Simulated responses to moving tilted line. Line length in each direction codes the activity level of the maximally active directionally tuned scale at that location. Ambiguous filter activation occurs in the line interior and unambiguous responses at the line ends after thresholding in (c).



(a)



(b)

Figure 7. Simulations of long-range filtering: (a) shows maximal activities from each scale after long-range filtering. Filtering enhances activity at feature points relative to ambiguous motion signals. The ambiguous motion signals are relatively uniform along the length of the line. (b) shows the same activities interpreted as motion vectors.

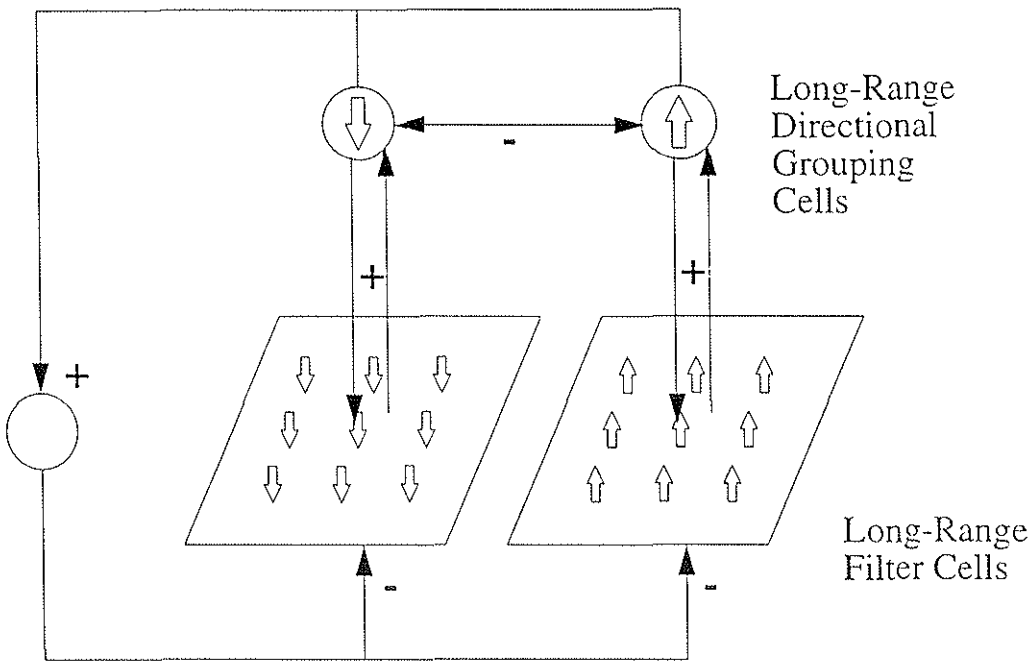


Figure 8. Long-range directional grouping and feedback in which each grouping cell nonspecifically inhibits all long-range filters while exciting cells with the same directional preference.

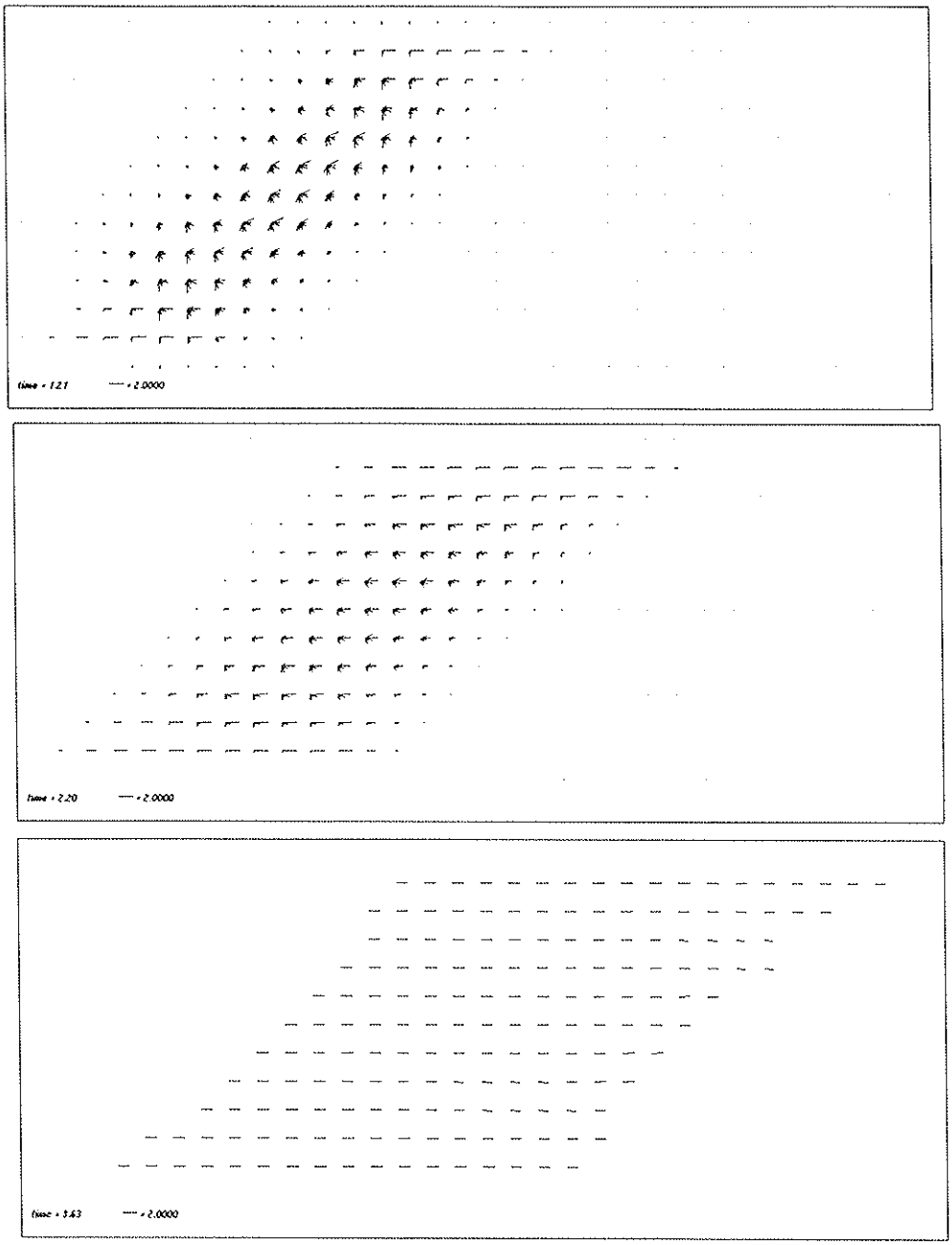


Figure 9. Maximal activities across all scales of the long-range filters during the rightward motion of a 45 degree line at three successive times. Ambiguous motion signals are eliminated by feedback inhibition from the long-range directional grouping cells.

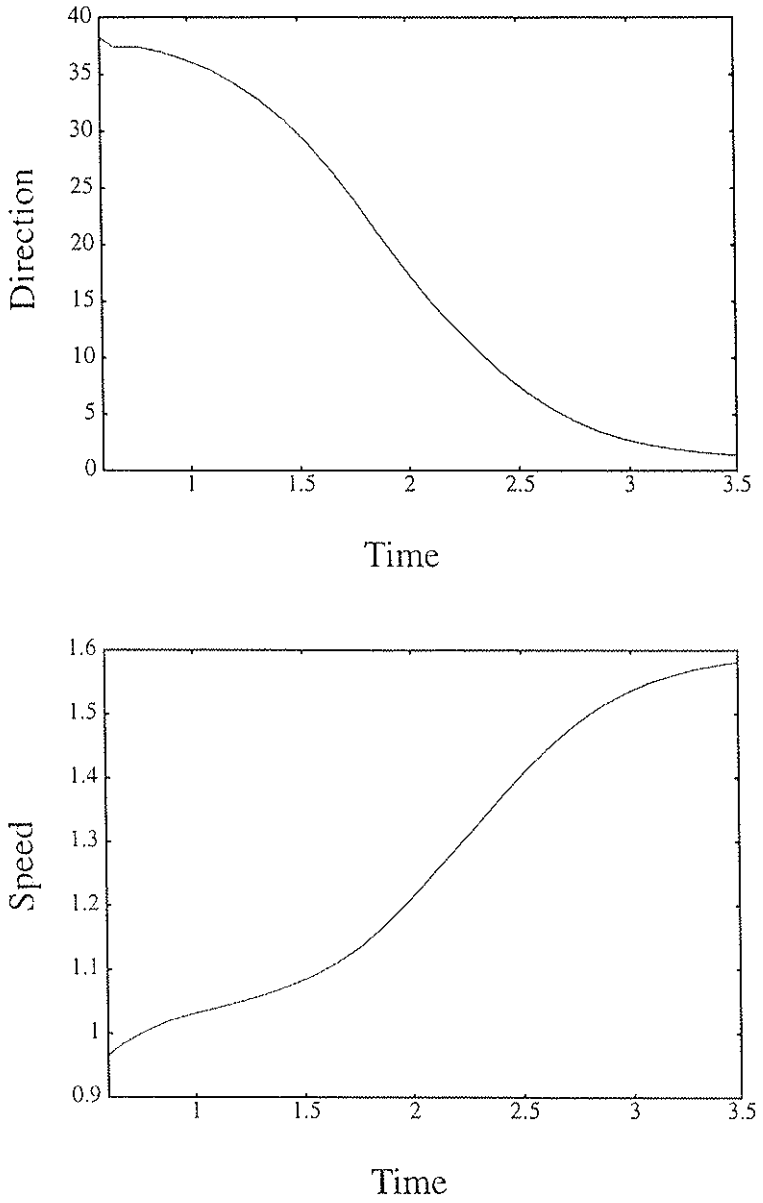


Figure 10. Motion direction and speed of a moving line derived from long-range filter activity over time (the same filter activities are shown in Figure 9). The top plot show how the perceived direction of a moving line gradually converges to its actual direction of motion (0 degrees from horizontal) after starting at a direction almost perpendicular to the line's orientation (45 degrees). The bottom plot shows how the perceived speed of the line gradually asymptotes over time.

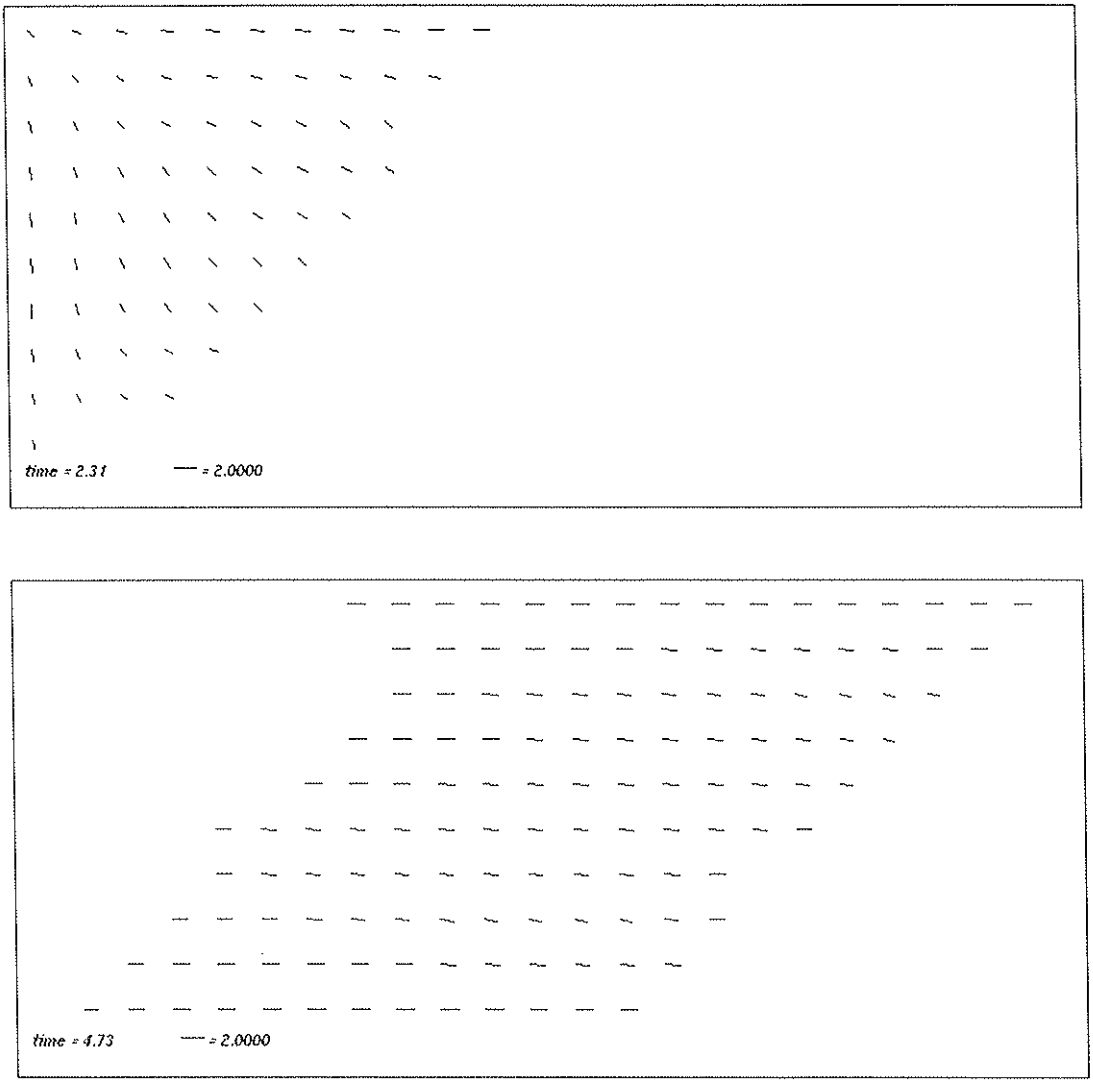


Figure 11. Time slices from the movement of a line behind a rectangular aperture illustrating the two phases of motion in the barberpole illusion. Each graphic shows long-range filter cell activity interpreted as motion vectors. In the first time slice, the line moves through the corner region and signals in the line interior indicate motion perpendicular to the line's orientation. Feature signals dominate the output near the aperture edges but do not dominate the ambiguous motion signals along the rest of the line. In the later time slice, motion is captured by feature motion signals in the direction coincident with the long axis of the aperture.

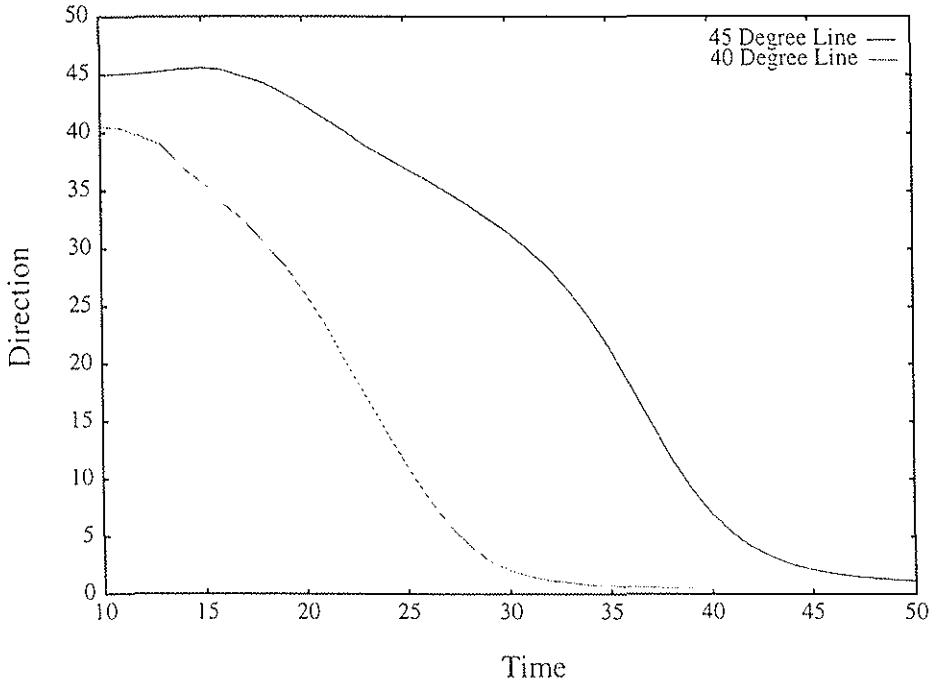
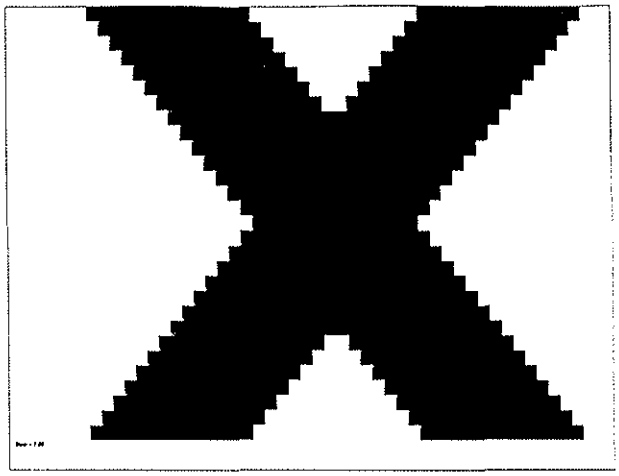
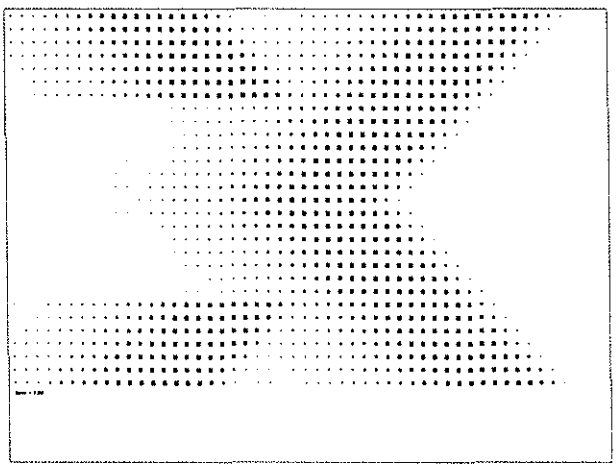


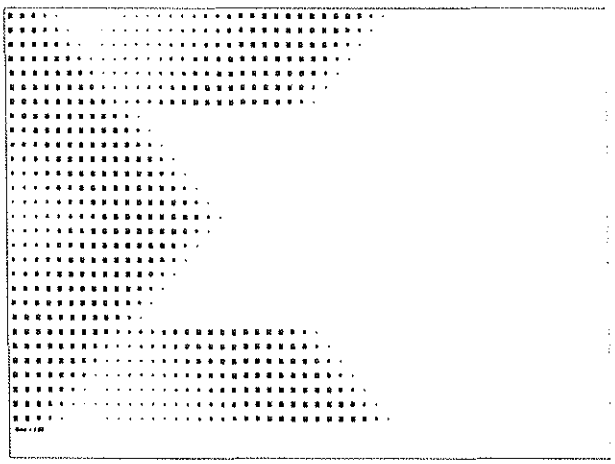
Figure 12. Perceived direction of lines moving behind a rectangular aperture (barberpole illusion) over time. Results are shown from two independent simulations utilizing two different line orientations. When the orientation of the line is tilted away from 45 degrees, the perceived direction of the line converges more rapidly to the direction coincident with the long axis of the aperture.



(a)



(b)



(c)

Figure 13. (a) Input luminance, (b) ON transient cell, and (c) OFF transient cell responses for a Type 1 symmetric plaid whose components are oriented at 45 degrees from vertical during rightward movement.

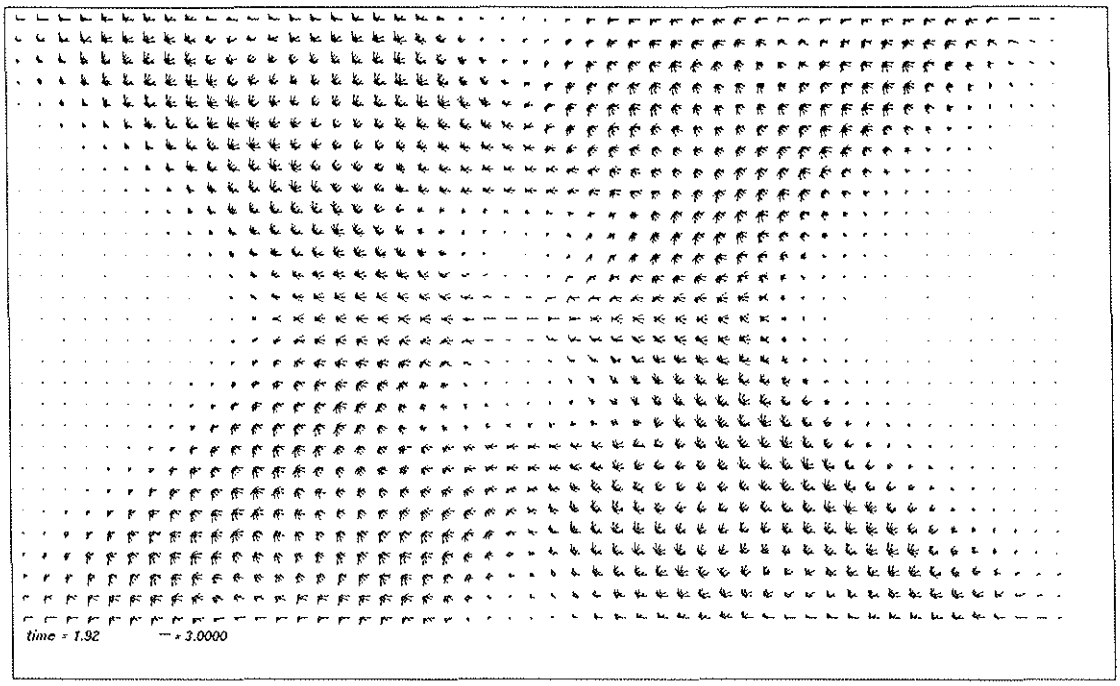
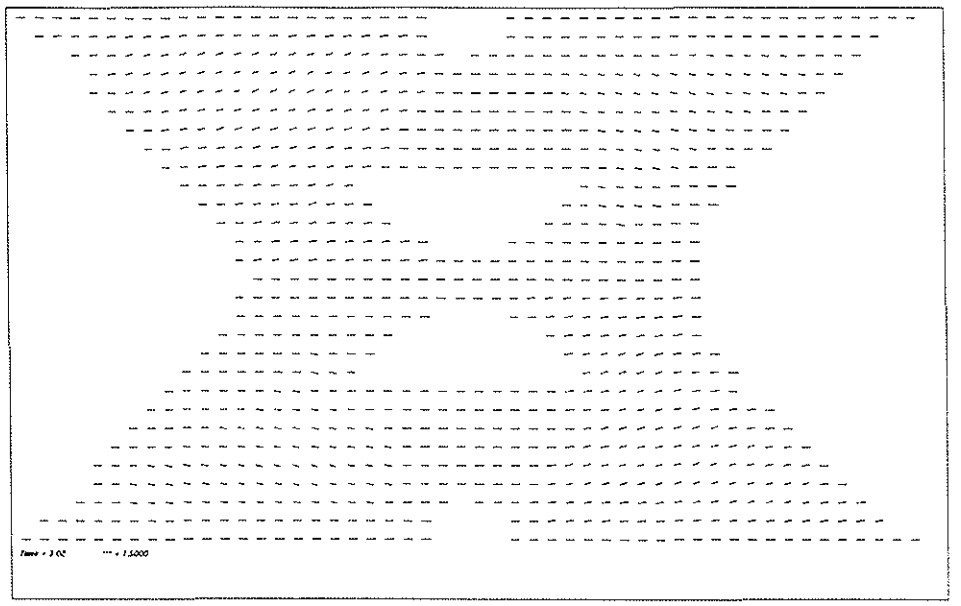
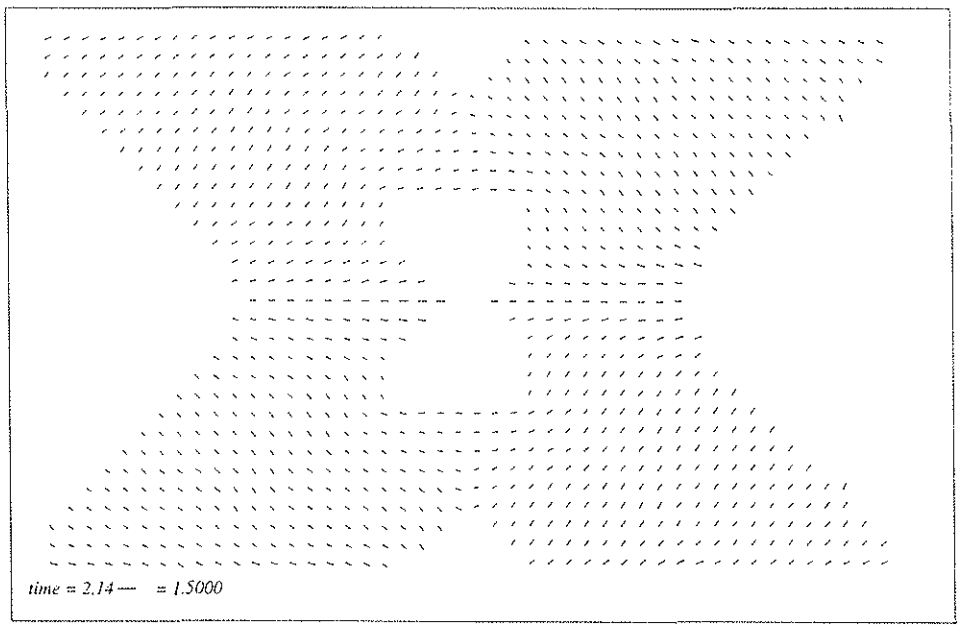


Figure 14. Long-range filter responses combining ON and OFF channels during simulated motion of a Type 1 symmetric plaid. Components are aligned at 45 degrees from vertical. Filter responses have been magnified to show ambiguous signals clearly. The large feature signals do not therefore appear at their true magnitude. Such feature signals are present at the four corners where the components intersect.



(a)



(b)

Figure 15. (a) Long-range filter outputs interpreted as motion vectors during the coherent plaid motion with components oriented at 45 degrees from vertical. Motion signals are almost entirely uniform across the pattern, indicating rigid motion. (b) Long-range filter outputs interpreted as motion vectors for a plaid pattern during incoherent motion. The two components move independently. The feature points still move horizontally, a common perceptual phenomenon when viewing incoherently moving plaid patterns.

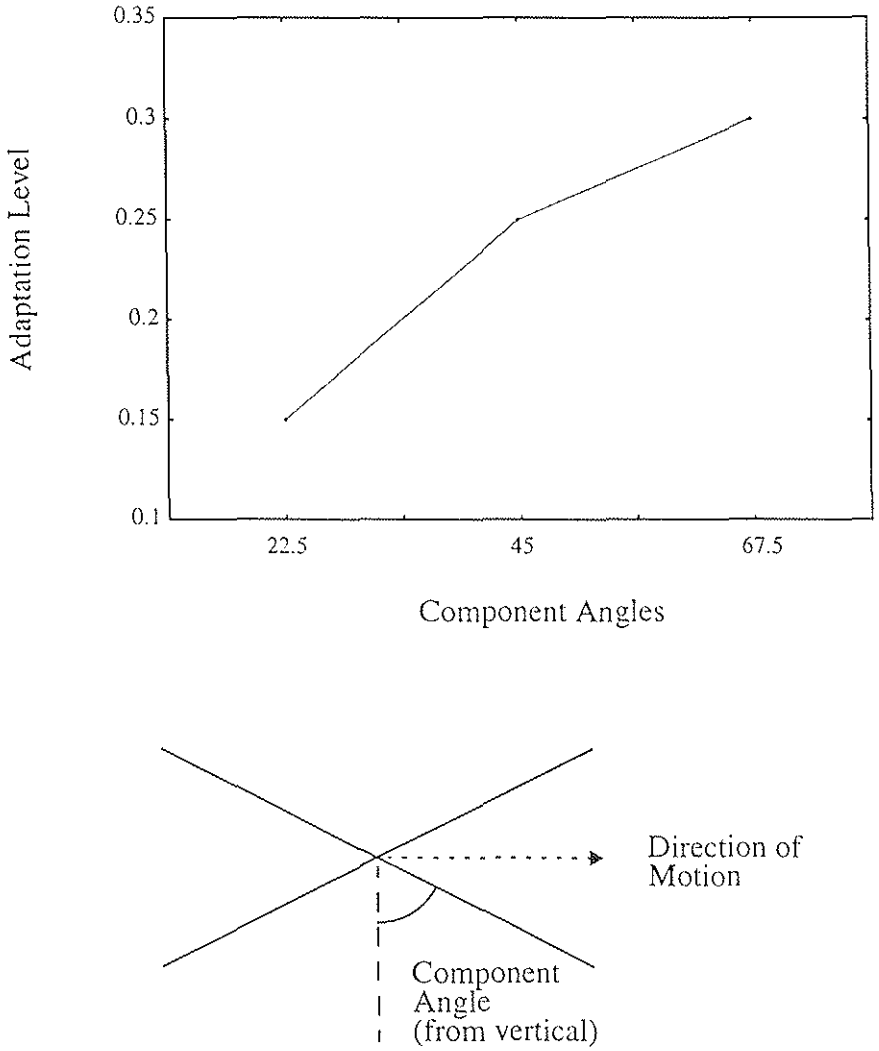


Figure 16. Adaptation levels at which simulated Type 1 symmetric plaids are perceived to move incoherently for a range of component angles from horizontal for a horizontally moving plaid. Each adaptation level is the ratio of the adapted to non-adapted long-range filter outputs. Only the rightward selective filters are adapted. When the component angles are near horizontal, greater adaptation is required for incoherent motion to be produced. This result is consistent with the smaller probability of seeing coherent motion with these components.

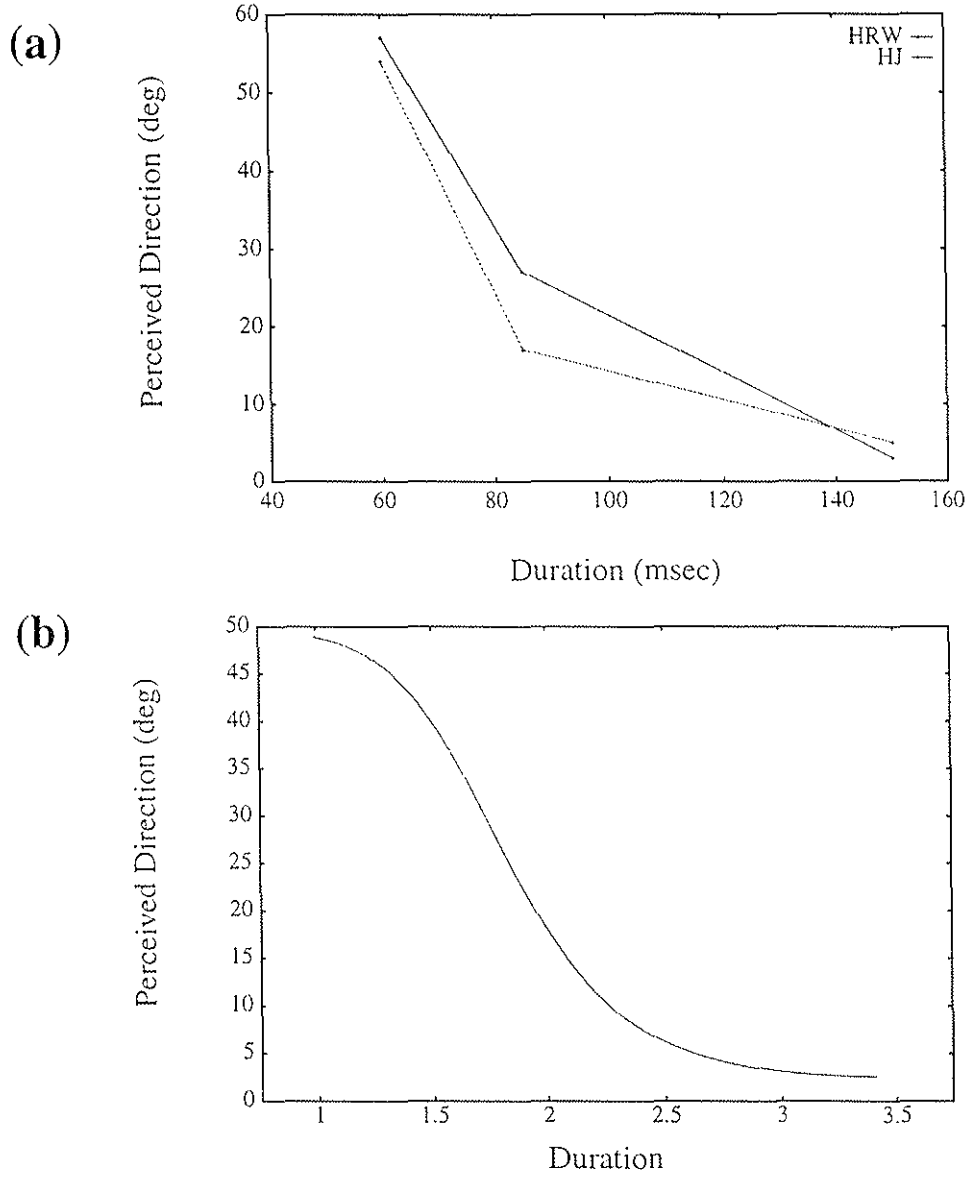


Figure 17. Perceived direction of Type 2 patterns is duration-dependent. At small exposure durations, patterns are perceived to move in the vector sum direction. For longer durations, the IOC direction dominates. Data in (a) are replotted from Yo and Wilson (1992), obtained using components whose normals are at 48.2 and 70.5 degrees from the direction of motion (vector sum 55.5). Simulated results are shown in (b), obtained using components oriented at 45 and 67.5 degrees (vector sum direction 56.25 degrees). The simulations use slightly different orientations to suit the Cartesian grid underlying the simulated input. In both cases component speeds are adjusted so that the IOC prescribed direction of motion of the plaid is at 0 degrees.

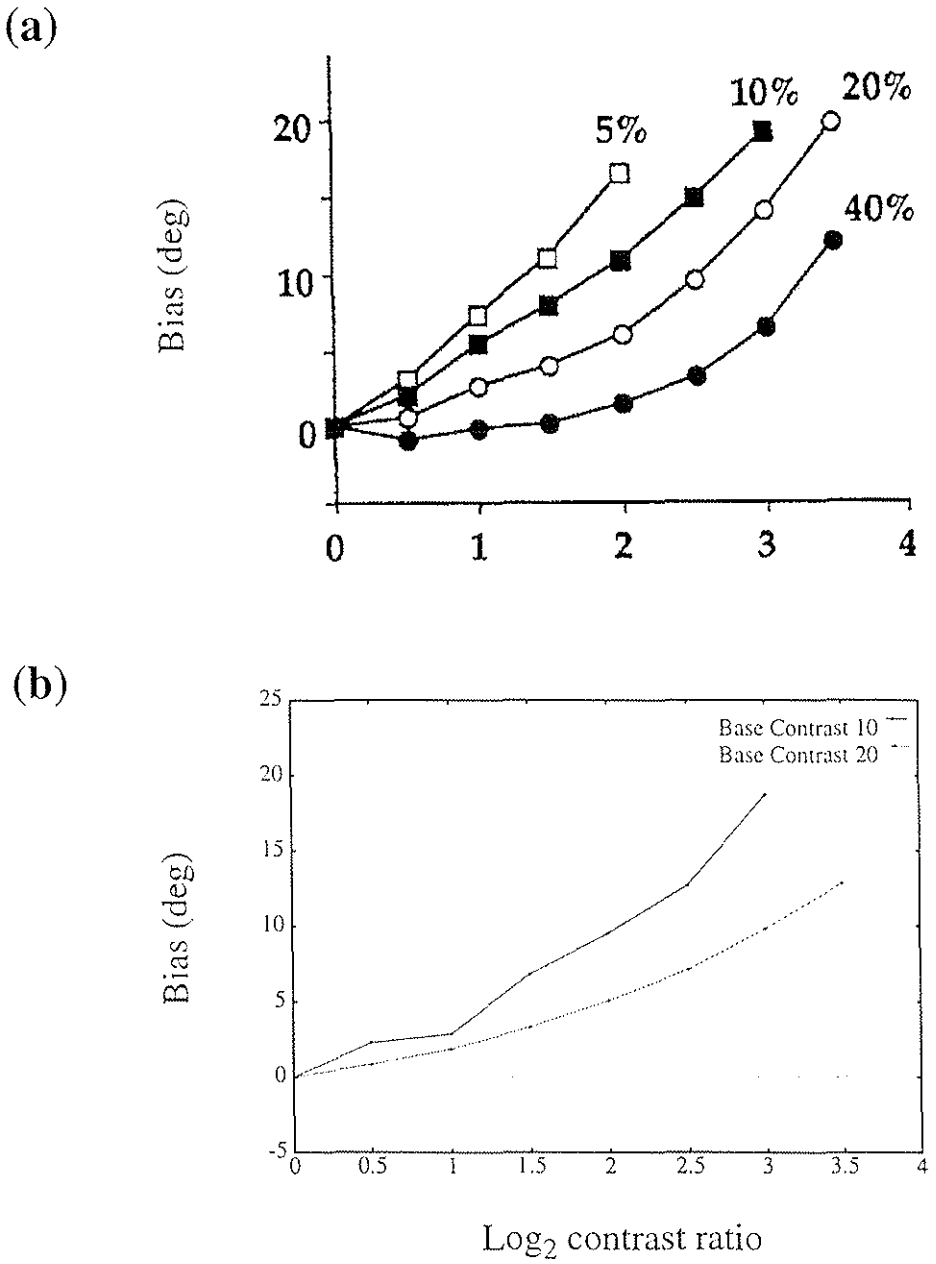


Figure 18. Perceived direction of Type 1 symmetric plaid patterns is dependent on component contrasts. Data from Stone *et al.* (1990) show bias in perceived direction as a function of the contrast ratio between the two components for four different base contrasts (a). Simulation results show the same qualitative biases for two different base contrasts (b). In both cases, components were oriented at 67.5 degrees from vertical.

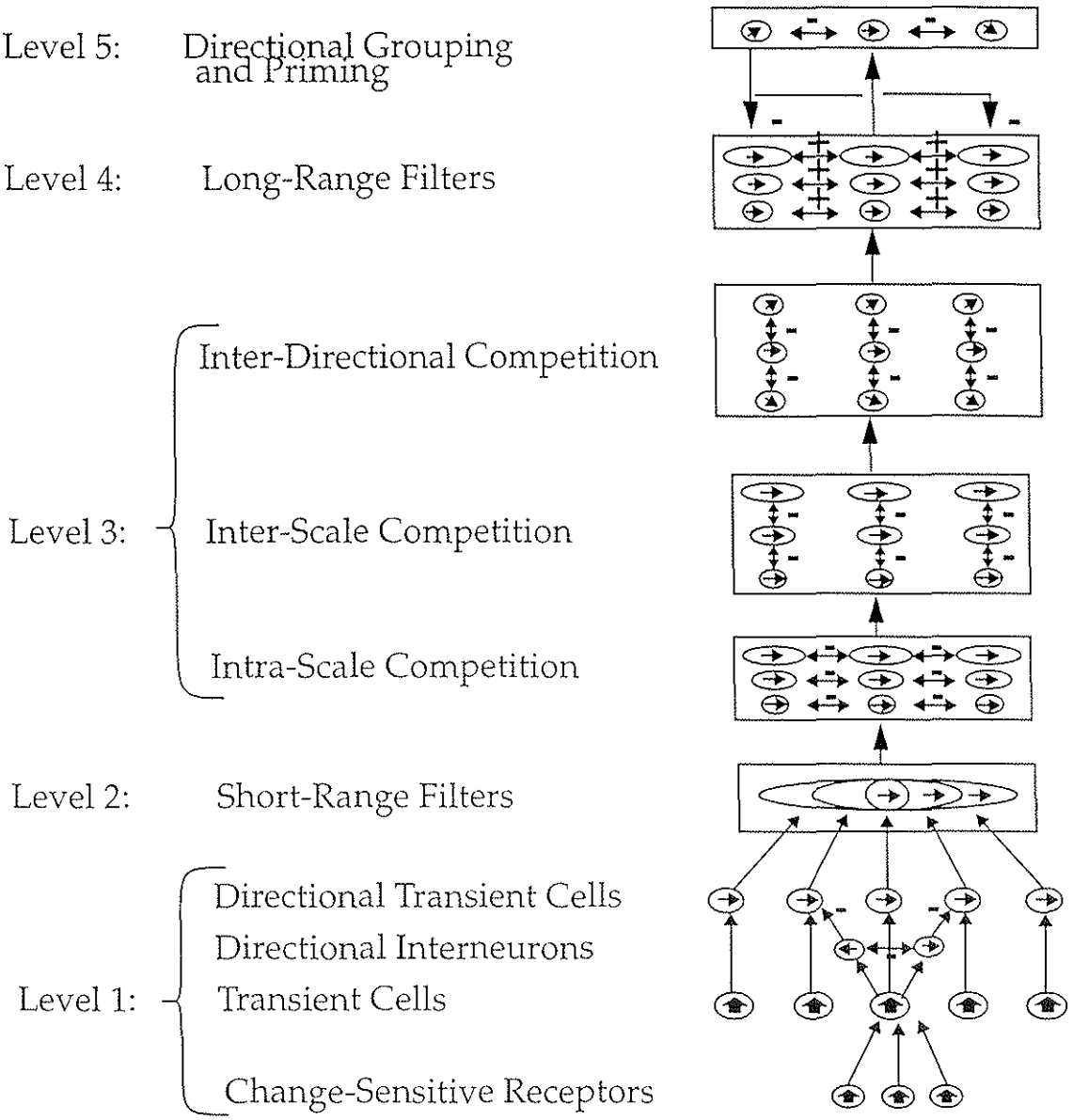


Figure 19. Model mechanisms. Level 1 consists of change-sensitive units that are transiently activated for fixed time intervals by a moving stimulus. Transient cells sum and time-average activities from fixed, non-overlapping sets of the change-sensitive units. Directional interneurons veto responses of directional transient cells to establish early directional selectivity. Multiple Level 2 short-range filters at each spatial position draw input from a set of directional transient cells, the size of which is determined by the spatial scale of the filter. The thresholded outputs of the short-range filters forms input to Level 3 intra-scale competition across space. Inter-scale competition then takes place between all scales at each spatial position. Competition between directions within each scale normalizes activity across directions at each location. Level 4 long-range filtering tracks the output of the inter-directional competition. Long-range filter outputs are selected through feedback from Level 5 long-range directional grouping cells.