BOSTON UNIVERSITY
GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

THE INTEGRATION AND SEGMENTATION OF VISUAL MOTION SIGNALS: EXPERIMENTS AND A COMPUTATIONAL MODEL OF CORTICAL MECHANISMS

by

LARS H. LIDEN

B.S., University of California, San Diego, 1994

Submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

1999
Approved by

First Reader

Ennio, Mingolla, Ph.D.
Associate Professor of Cognitive and Neural Systems
and of Psychology

Second Reader

Takeo Watanabe, Ph.D.
Associate Professor of Psychology

Third Reader

Stephen Grossberg, Ph.D.
Wang Professor of Cognitive and Neural Systems
Acknowledgments

I wish to thank my advisors, Professor Ennio Mingolla, Professor Takeo Watanabe, and Professor Stephen Grossberg, as well as the rest of the faculty of the Department of Cognitive and Neural systems, for their guidance during my course of study. For the past four years, my research was supported by the Krell Institute’s Computational Science Graduate Fellowship Program. I owe thanks for their generous support. I also wish to thank my parents, Sam and Giselsa, and my wife Sarah for their constant support.
THE INTEGRATION AND SEGMENTATION OF VISUAL MOTION SIGNALS:
EXPERIMENTS AND A COMPUTATIONAL MODEL OF CORTICAL
MECHANISMS

LARS LIDÉN
Boston University Graduate School of Arts and Sciences, 1999
Major Professor: Ennio Mingolla, Associate Professor of Cognitive and Neural Systems

ABSTRACT

The computational principles that govern the integration and segmentation of motion signals extracted from the visual array are examined using two methods. First, psychophysical studies of human subjects examine the variables that control how motion signals interact across visual space. These studies investigate both the role of similarity between motion signals in different areas of space and the role played by objects in the intervening space, such as occluding patches, in determining how signals from different locations are combined into a coherent motion percept. The present work shows that motion signals are preferentially combined between areas of space that are similar in depth, spatial frequency, color or contrast. It is also shown that several kinds of occlusion cues (depth, T-junctions, texture) are capable of influencing the degree to which motion signals near regions of occlusion contribute to the final motion percept. A neural network model is developed to account for the major published findings regarding motion integration and
segmentation, such as those concerning the effects of terminator salience and stimulus arrangement on the perceived direction of motion. The model provides mechanistic accounts, illustrated in computer simulations, of the chopsticks illusion, the occluded diamond illusion and the barber pole illusion. The model consists of integration and segmentation sub-systems, each composed of interacting cells that are tuned to the direction of motion. Integration is necessary to overcome noise and the inherent ambiguity in locally measured motion direction (the aperture problem). Competitive and cooperative interactions between integration cells serve to overcome noise in the input signal and propagate disambiguated motion signals across visual space. Segmentation is required to detect the presence of motion discontinuities and prevent spurious integration of motion signals between objects with different trajectories. The model’s segmentation cells form motion borders, which in turn serve as barriers for the propagation of motion signals by the integration cells. Possible links to the neurophysiology of the middle temporal visual area (MT, also known as V5), an area thought to be important for motion processing, are suggested.
## Contents

1. Introduction ................................................................. 1
   1.1 The Computation of Visual Motion ................................. 1
   1.2 Background ............................................................. 3
       1.2.1 Local Signals are Noisy ....................................... 3
       1.2.2 Motion Integration ............................................. 5
       1.2.3 Motion Segmentation ........................................... 7
       1.2.4 The Problem: When to Segment vs. Integrate? ................. 8

2. Factors Influencing Motion Integration Across Visual Space ...... 10
   2.1 Introduction ............................................................ 10
       2.1.1 Plaid Motion ..................................................... 11
       2.1.2 The Barber Pole Illusion ...................................... 12
   2.2 Experimental Series A ............................................... 14
       2.2.1 Series A: Methods .............................................. 15
       2.2.2 Experiment A1: Depth Manipulation ............................ 16
       2.2.3 Experiment A2: Contrast Manipulation ........................ 18
       2.2.4 Experiment A3: Phase Manipulation ........................... 19
       2.2.5 Experiment A4: Spatial Frequency Manipulation ............... 20
       2.2.6 Series A: Discussion ........................................... 22
   2.3 Experimental Series B ................................................. 24
       2.3.1 Series B: Methods ............................................... 25
2.3.2 Similarity in Opposition: Experiment B1a & B1b .......................... 28
2.3.3 Terminators and Similarity: Experiment B2a & B2b ....................... 33
2.3.4 Contrast Effects: Experiments B3a & B3b ................................. 39
2.3.5 Series B: Discussion ....................................................... 40
2.4 General Discussion ......................................................... 42

3. Monocular Occlusion Cues Alter the Influence of Terminator Motion in the Barber Pole Phenomenon .......................................................... 44
3.1 Introduction .................................................................. 44
3.1.1 Experiment 1: Textured Occluders With Depth ....................... 49
3.1.2 Experiment 2: Textured Occluders Without Depth .................. 54
3.1.3 Experiment 3A & 3B: Occluder Contrast ............................ 57
3.1.4 Experiment 4: Amodal Completion or Terminator Classification? ... 62
3.1.5 Experiment 5: Occlusion of One Edge ............................... 63
3.1.6 Experiment 6: Role of T-junctions .................................. 66
3.1.7 Experiment 7: Illusory Occlusion .................................. 71
3.2 General Discussion ....................................................... 73

4. A Model of Motion Segmentation and Integration .......................... 80
4.1 Modeling Introduction .................................................. 80
4.2 Review of other Models .............................................. 82
4.2.1 Segmentation or Integration ........................................ 82
4.2.2 Temporal Aspects of Motion Integration and Propagation ....... 86
4.3 Sketch of the Model ...................................................... 88
A.2 Predicting the Attenuation for One Side .......................... 146

Simulation Details 148

B.1 Input Nodes (V1) ......................................................... 148
B.2 Integration Nodes (MTi) ................................................. 150
B.3 Segmentation Nodes (MTs) .......................... 151
B.4 Summary of Variables .......................... 154
B.5 Summary of Parameters .......................... 156
List of Figures

Figure 1-1. Texture Accretion and Deletion .................................................................4

Figure 1-2. The Aperture Problem ..............................................................................5

Figure 2-1. The Flanking Barber Pole Paradigm ......................................................15

Figure 2-2. Experiment A1: Disparity Manipulation .................................................18

Figure 2-3. Experiment A2: Contrast Manipulation .................................................20

Figure 2-4. Experiment A3: Phase Manipulation ......................................................21

Figure 2-5. Experiment A4: Spatial Frequency Manipulation .................................22

Figure 2-6. The Effect of Occluders ..........................................................................26

Figure 2-7. Plus-shaped Stimulus ..............................................................................27

Figure 2-8. Cross-shaped Stimulus ...........................................................................28

Figure 2-9. Horizontal and Vertical Orientations ......................................................28

Figure 2-10. Color and Flank Depth in Opposition: Stimulus .................................31

Figure 2-11. Color and Flank Depth in Opposition: Results ....................................32

Figure 2-12. Occluder Configurations ......................................................................35

Figure 2-13. Sample Configurations for Free-fusing ..............................................36

Figure 2-14. Spatial Frequency and Terminators in Opposition: Stimulus ..........37

Figure 2-15. Spatial Frequency and Terminators in Opposition: Results ...............38

Figure 2-16. Central and Flank Contrast ..................................................................41
Figure 3-1. Three Stimulus Configurations Used for Experiment 1.............................47
Figure 3-2. Experiment 1: Textured Bars With Disparity Differences .....................53
Figure 3-3. Experiment 2: Textured Bars Without Disparity Differences ..................58
Figure 3-4. Experiment 3A & 3B: Grey Bars on a White Background.......................61
Figure 3-5. A Role of Amodal Completion? ..............................................................62
Figure 3-6. Experiment 4: Addition of Thin Bars .....................................................64
Figure 3-7. Experiment 5: Occlusion of One Side .....................................................67
Figure 3-8. T-Junctions are Indications of Occluding Surfaces ...............................68
Figure 3-9. Experiment 6: Elimination of T-Junctions .............................................70
Figure 3-10. Experiment 7: Addition of Occluding Bars .........................................74
Figure 3-11. All Conditions with Horizontal Occluders .........................................78
Figure 4-1. Integration Cell Architecture .................................................................93
Figure 4-2. Gating of Segmentation Cell Input .........................................................97
Figure 4-3. Real and Occluder Generated Terminators .........................................100
Figure 5-1. Translating Line: V1 Activation .............................................................109
Figure 5-2. Translating Line: MTi Activation .........................................................110
Figure 5-3. Key to Color Representation .................................................................111
Figure 5-4. Translating Line: Color Coded MTi Node Activity ...............................112
Figure 5-5. Line Capture .........................................................................................114
Figure 5-6. Line Capture Model Output .................................................................116
| Figure 5-7. | Barber Pole Illusion | 121 |
| Figure 5-8. | Diamond with Bar Occluders | 123 |
| Figure 5-9. | Diamond with Triagonal Occluders | 124 |
| Figure 5-10. | Ambiguous signals propagating towards one another | 126 |
| Figure 5-11. | Simulations of Divided Diamonds | 128 |
| Figure 5-12. | Crossing Bars | 130 |
| Figure 5-13. | Crossing Bars: No Segmentation Mechanisms | 131 |
| Figure 5-14. | Crossing Bars: Initial Motion Measurement | 132 |
| Figure 5-15. | Model output for Crossing Bars Stimulus | 133 |
| Figure 5-16. | Within-Object Occlusion | 134 |
| Figure 5-17. | Simulation: Translating Square | 135 |
| Figure 5-18. | Passing Diamonds Problem | 136 |
| Figure 5-19. | Simulation: Passing Diamonds | 137 |
| Figure 5-20. | Two Approaching Rectangles | 138 |
| Figure 5-21. | Incorrect Two Rectangle Output | 139 |
| Figure 5-22. | Three Objects | 140 |
| Figure 5-23. | Simulation: Two Objects | 141 |
| Figure 5-24. | Simulation: Three Objects | 142 |
### List Of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1</td>
<td>Primary Visual Area</td>
</tr>
<tr>
<td>V2</td>
<td>Secondary Visual Area</td>
</tr>
<tr>
<td>MST</td>
<td>Medial Superior Temporal Area</td>
</tr>
<tr>
<td>MT</td>
<td>Medial Temporal Area</td>
</tr>
<tr>
<td>MT_i</td>
<td>Integration Nodes</td>
</tr>
<tr>
<td>MT_s</td>
<td>Segmentation Nodes</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 The Computation of Visual Motion

The processing of motion is perhaps the most fundamental task of any visual system. The detection of motion is so important for survival that only humans and other evolved primates can respond to objects that don’t move (Kandel 1991). Frogs and other simple vertebrates cannot see objects which are stationary relative to their visual systems (Lettvin, Maturana, McCulloch & Pitts 1959). Even in humans, the ability to detect motion persists in areas of the peripheral visual field where static motion perception is no longer functional. Apart from its obvious applications, motion processing also plays a more general role, involved in tasks as diverse as time-to-collision, encoding of the third dimension, image segmentation and driving eye movements (Nakayama 1985). Although years of research have been performed, the computational principles behind aspects of motion processing are just beginning to be revealed.

Physiological and theoretical work suggests that the extraction of motion information takes place in at least two stages (Adelson & Movshon, 1982, Yuille & Grzywacz, 1988). The first stage extracts raw local motion measurements. The second stage combines local motion measurements to form a global coherent motion percept.

From a functional point of view it is advantageous to further sub-divide the second stage into two computational processes, an integration and segmentation process (Brad-
A motion processing system must both integrate motion signals, to overcome noise and the aperture problem, and segment motion signals, to detect the presence of motion discontinuities and prevent integration of motion signals between objects with different trajectories. Physiological (Tanaka, Hikosaka, Saito, Yukie, Fukada & Iwai, 1986; Allman & McGuinness, 1985; Born & Tootell, 1992 and Saito, 1993) and neuropsychological studies (Vaina & Grzywacz, 1992 and Vaina, Grzywacz & Kinkinis, 1994) support the idea that the second stage is sub-divided into these two computational processes, one serving to integrate motion signals across space, the other serving to detect motion discontinuities.

The first part of the dissertation consists of an examination of the motion integration and segmentation processes using visual psychophysics. The dissertation’s experiments are divided into two sets. The goal of the first set is to examine the role that similarity plays in motion segmentation and integration, specifically, how features such as contrast, spatial frequency, color and disparity influence the degree to which motion in one area of visual space affects motion perception in another (Chapter 2). The goal of the second set of experiments is to examine how objects in the intervening space influence the integration of motion signals. Specifically, the second set of experiments will examine the role of occlusion and cues, such as T-junctions, responsible for affecting motion integration (Chapter 3). The second part of the dissertation introduces a model for motion computations with interacting integration and segmentation sub-systems. Chapter 4 explains the computational principles and design of the model. Chapter 5 examines the performance of the model on a variety of motion stimuli.
1.2 Background

1.2.1 Local Signals are Noisy

There are many computational methods for extracting information about local motion direction from a moving image. (See Chapter 4 for detail.) However, there are several problems inherent to this process that invariably result in inaccurate and non-unique local motion measurements. In general, motion computation falls in the category of problems which are ill-posed (Poggio & Koch 1985).

The first major difficulty with acquiring motion measurements is that of surface specularity. As an object moves, the orientation of its surfaces changes with respect to the lighting source. Subsequently, the luminance of a particular point on the object is not consistent over time. As motion measurements depend on luminance consistency, such specularity introduces noise into any motion measurement.

A second important source of noise is that of texture accretion and deletion. As an object moves it covers background texture from view (deletion). At the same time previously occluded background texture is uncovered (accretion) (Kaplan, 1969). Such kinetic occlusion is a problem for local motion estimation procedures for several reasons: the intensity at the border is not conserved, the are multiple motions present at the edge and the flow field is discontinuous (Niyogi, 1995).

The third major difficulty with local motion estimation is the aperture problem (Marr & Ullman 1981). The aperture problem states that any localized motion sensor can only detect motion orthogonal to local contour. An example is depicted in Figure 1-2. Such
motion measurements are said to be *ambiguous* as 180° range of directions is compatible with the local motion measurement. A local motion detector, therefore, cannot generally compute the direction of motion of an object that is bigger than its own field of view. Since early detection of motion is carried out by neurons with relatively small receptive fields, the aperture problem is an immediate difficulty for biological visual systems.

In contrast to motion measurements along a contour, motion measurements at the termination of a contour are *unambiguous*, as only one direction of motion is compatible with the motion of a contour’s terminators. An effective motion processing system must use unambiguous motion signals in one region of the image to constrain the interpretations of ambiguous local motion directions in other regions. Examples of the aperture problem and corresponding sensor outputs are shown in Chapter 5.

Hildreth & Koch (1987) state “as a consequence of the aperture problem, the measurement of motion in the changing image requires two stages of analysis: The first stage
measures components of motion in the direction perpendicular to image features; the second combines these components of motion to compute the full 2-D pattern of movement in the images.”

1.2.2 Motion Integration

Although local information may be noisy or ambiguous, one can begin to resolve some of these deficits by combining (integrating) nearby motion signals. (See Chapter 4, Section 4.2 for a discussion of relevant models.) Spurious local motion signals can be eliminated by averaging motion signals over a small neighborhood. Ambiguous motion signals resulting from the aperture problem can be disambiguated by propagating unambiguous motion signals provided by image features such as terminators, corners and dots (Wallach 1935).

There are a large number of psychophysical phenomenon, such as motion capture, which support the ideas of motion integration and propagation (Ramachandran 1983). Motion capture occurs when motion in one area of the visual field results in another area
appearing to move in the same direction. For example, Ramachandran and Anstis (1983) described apparent motion displays with Necker cube-like bistability. The displays can be seen either as moving up/down, left/right, or rotating. When several of such stimuli are presented simultaneously, they are all seen to possess the same type of movement.

Another example occurs when two uncorrelated random dot displays are alternated. Usually no motion is seen, however when the same random dot displays are superimposed on a moving grating, the random dots are seen to move with the grating (Ramachandran & Inada 1985). The effect is so pronounced that random dot patterns that actually move in the opposite direction can still be captured by a moving grating.

*Random-dot cinematograms*, introduced by Anstis (1970), provide another example of how motions are integrated over spatial area. Random-dot cinematograms consist of random dot fields whose individual elements are displaced over time and whose trajectories are drawn from a uniform distribution. There is a tendency to see coherent flow in the mean direction of the dots even when a wide distribution of motions is employed (Williams and Sekuler 1984).

Chang and Julesz (1984) demonstrated that motion capture can occur in random-dot cinematograms using strips of alternating ambiguous and unambiguous horizontal random-dots. They found that ambiguous strips whose dots were statistically biased to move in the opposite direction of unambiguous strips can be perceived to move in the direction of the unambiguous strips. They argue for a cooperative process in which motion in unambiguous areas influences motion perception in ambiguous areas.
1.2.3 Motion Segmentation

Although the integration of local motion signals can help to overcome some of the inaccuracies in local motion measurements, integration produces problems of its own, particularly when there are multiple objects present in the visual array. The integration of motion signals arising from different objects leads to spurious results. A second process involving motion segmentation, for the detection of motion discontinuities, is equally important as motion integration.

Psychophysical experiments demonstrate the importance of segmentation. Nawrot and Sekuler (1990) examined what they termed heterokinesis and homokinesis. Heterokinesis is said to occur when a stimulus moving in one direction causes another stimulus to appear to move in the opposite. Homokinesis occurs when a stimulus moving in one direction causes another stimulus to appear to move in that same direction. They examined these two processes using cinematograms comprised of strips of moving dots which alternated between strips with random motion and strips whose motion was biased in a particular direction.

Nawrot and Sekuler (1990) found that the width of the strips was of critical importance in determining the perceived motion. When strips were narrow, motions were integrated across strips and assimilation occurred; the strips with random motion were seen as moving in the direction of the biased strips. When the strips were wide, contrast occurred; random strips were seen moving in the opposite direction of the biased strips. Spatial scale appeared to be a critical feature determining if an integration or segmentation of motion in strips occurred.
Marshak and Sekuler (1979) reported repulsion between moving visual targets. When two groups of dots move at different angles, subject overestimate the angle between the two sets of dots. The repulsion effect is greatest when the two dot sets move at the same speed. Marshak and Sekuler suggest that these direction shifts could be due to inhibition between elements sensitive to different directions of motion.

The localization of motion discontinuities also plays an important role in figure-ground segmentation. The sequential presentation of two random dot pattern which are identical apart from a central patch of dots that is shifted, creates a vivid impression of a central object with sharply defined borders (Anstis, 1970; Braddick, 1974). Objects which are otherwise invisible can be segmented when they move (Regan & Beverley 1984). Motion boundaries also often correspond to depth discontinuities in an image and can be used to refine estimations of surface discontinuities (Poggio, Gamble & Little 1988).

**1.2.4 The Problem: When to Segment vs. Integrate?**

Given that there are two competing computational requirements, integration of motion signals and segmentation at motion discontinuities, there must be some mechanism for resolving which process should dominate at each point in the visual array. Differences in nearby motion signals due to noise and the aperture problem should be smoothed. On the other hand, differences across motion discontinuities should not. How does the visual system know which process should dominate? How do the two processes interact? Whether unambiguous information is used for disambiguation depends on how the image is segmented. Motion signals which might otherwise influence motion percep-
tion in another area of space can be blocked when there are significant differences in the motion signals. What defines a significant difference? The following chapters will attempt to answer some of these questions.
Chapter 2

Factors Influencing Motion Integration Across Visual Space

2.1 Introduction

In order compute global motion, local motion processing units must interact to both integrate motions arising within objects and segment motion arising from different objects. To understand the computational principles behind these calculations, it is necessary to understand how motion processing units are connected and what stimulus features govern their interaction.

Research has examined how factors such as spatial frequency, disparity and contrast influence the integration and segmentation of motion signals at a single spatial location. However, less is known about how such factors influence motion processing of widely separated areas of visual space. The following chapter describes a series of experiments which examine the effects of these and other factors on the integration and segmentation of motions in separated spatial areas.

The nature of the interactions between motion processing units across space is unknown. Which attributes allow for the combination of motion signals and which create discontinuities? Work on transparent motion suggests there might be suppression between neurons tuned to opposite directions of motion when they have the same spatial frequency and disparity tuning (Qian, andersen & Adelson, 1994). Motion capture suggests that low spatial frequencies may have greater influence on higher frequency motion measurements.
than vice versa (Ramachandran & Inada, 1985).

The current study uses two sets of experiments to explicitly examine how factors such as spatial frequency, spatial phase, contrast, color, and disparity effect the combination of motion signals across disparate areas of space. It examines how these factors interact and how terminator classification plays a role. From such measurements one can make inferences about the nature of connections between motion processing units.

### 2.1.1 Plaid Motion

One method used to study the integration and segmentation of motion signals is that of plaid motion. Plaid stimuli consist of two overlapping sinusoidal or square-wave gratings viewed through an aperture. Under certain conditions, the two grating patterns are said to ‘cohere’, as subjects report seeing a single translating plaid pattern rather than two separate gratings translating in directions perpendicular to their orientation. Studies of plaid motion suggest that features such as spatial frequency, duration, contrast, eccentricity, disparity and relative component direction are critical for determining when such patterns cohere (Yo & Wilson, 1992; von Grunau, Dube & Kwas, 1993; Kim & Wilson 1993).

Unfortunately, when using plaid patterns, the motions that are integrated at times overlap in spatial position, as the gratings produce blob-like intersections. Experiment evidence strongly suggests that the perception of plaid direction is mediated in part by a blob tracking mechanism (Alais, Wenderoth & Burke 1994). Although some progress has been made through the use of multi-aperture plaid-like displays (Mingolla, Todd & Norman,
1992), plaid patterns are generally of limited utility for studying the way motion signals are propagated across space. Furthermore, motion integration takes time (Williams and Sekuler, 1984; Snowden and Braddick, 1989; Watanabe and Cole, 1995), and a full examination of the temporal aspects of motion propagation is also not possible when studying motions that overlap in spatial location as such a study requires a spatial separation.

2.1.2 The Barber Pole Illusion

Another classic phenomenon that may be better suited for studying motion integration across space is the barber pole illusion. Wallach (1935) illustrated that under normal conditions the perceived direction of motion of a diagonal grating pattern drifting behind a rectangular aperture depends upon the elongation of the aperture (Wuerger, Shapley & Rubin, 1996). The prevailing explanation for this phenomenon suggests that the perceived direction of motion results from the integration of motion signals from grating terminators at the edges of the aperture. Due to the aperture problem, the local velocity signals along the grating are inherently ambiguous. Terminators, however, produce unambiguous motion signals. As a vertically elongated aperture has larger number of terminators with vertical trajectories, vertical motion usually prevails.

Kirita (1988) studied the determinants of the barber pole illusion and found that the illusion decreased as the orientation of the aperture relative to the grating pattern increased, the speed of the gratings increased and the contrast of the gratings decreased. Kirita argued “from the standpoint of motion integration it can be assumed that the perceived....motion of the barber pole illusion is an outcome of [a] motion integration process
in the human visual system. The barberpole illusion may therefore reflect the way by which initial motion information is integrated” (p. 44). By carefully controlling parameters in the barber pole, one can determine how factors such as spatial frequency, phase, contrast, disparity, spatial separation and color effect the propagation of motion signals.

Adding a disambiguating signal to one area of a barber pole display can rapidly disambiguate the entire barber pole, giving it a unique direction of motion. For example, adding unambiguously translating dots to a barber pole display can override the influence of terminators, resulting in perceived barber pole motion in the direction of dot motion (Shiffrar, Li & Lorenceau 1995). Adding features such as indentations can abolish the barber pole illusion producing motion perpendicular to grating orientation (Kooi 1993).

However, features added to a barber pole display are not always integrated. Shiffrar et al. (1995) report that when subjects view a barber pole with dots, they are less likely to interpret barber pole gratings as moving in the same direction as the dots when the velocity and the depth of intrinsic grating terminators and the dots differs. The greater the difference, the less grating motion is seen in the direction of dot motion. Shiffrar et al. (1995) conclude that if locally unambiguous velocities differ, then those velocities are segmented and assigned to different objects. Local velocity information is only pooled between regions whose unambiguous motion is in sufficient “agreement”. The dimensions and relevant variables for determining the level of agreement, such as speed and direction, are not, however, fully understood. Furthermore, the amount of agreement does not necessarily coincide with the directional similarity of two regions.

There is also evidence that integration of motion signals is spatial frequency depen-
dent. Mulligan & Beutter (1994) showed subjects sinusoidal barber poles through elongated gaussian windows. Subjects were biased to see motion in the direction of elongation of the window, but the bias was shown to decreases smoothly with increasing spatial frequency of the barber pole stripes.

2.2 Experimental Series A

The goal of the first set of experiments was to determine the conditions under which unambiguously specified motion in one area of a motion display can constrain ambiguous motion signals in a spatially separated area. One method for such an experiment would be to create barber poles with unambiguous motion by adding dots or indentations to the display. These unambiguously moving regions could be placed next to ambiguously moving regions and their effect on the ambiguous regions observed. Unfortunately, when experimental parameters (spatial frequency, contrast, phase, etc.) are manipulated, the intervening border creates a discontinuity between the barber poles.

This problem can be remedied by presenting an occluding surface between the barber poles (See Figure 2-1). The occluders can also be given a different depth than the barber poles by presenting the display stereoscopically. Shimojo, Silverman & Nakayama (1989) used the fact that aperture aspect ratio generally determines the perceived direction of motion to examine the effect of surface depth on the barber pole illusion. When three barber pole stimuli are placed next to each other and divided by occluding bars as shown in Figure 2-1, the three surfaces can be viewed either as three separate horizontal entities, or as one large vertical entity occluded by bars.
Shimojo et al. (1989) manipulated the stereoscopic depth between the barber poles and the occluding bars. When bars were perceived as being in front of the barber poles, subjects reported more vertical motion. When bars were perceived at the same depth or lower than the barber poles, subjects reported more horizontal motion.

The flanking barber pole paradigm was employed in the first set of experiments as allows for an examination of the way in which motion signals are grouped (or integrated) across space while at the same time eliminating discontinuities at the edges of the barber poles.

2.2.1 Series A: Methods

Barber poles were presented through a mirror and prism haploscope with occluding bars as in Shimojo et al. (1989). All stimuli were displayed on a Silicon Graphics Reality Engine (model CMN-A011). The characteristics of the flanking barber poles were manipulated and the perceived direction of motion of the central barber pole measured. Subjects were informed to keep their eyes focused on the central barber pole and after motion direction had stabilized to report the dominant perceived direction of motion (horizontal or
vertical) for the central barber pole. A minimum viewing duration of 500 msec was required. The perception of vertical translation served as an indication that motion signals were being combined between the three barber poles (coherent motion). One the other hand, the perception of horizontal translation indicated that motion signals were not being combined between the barber poles (non-coherent motion).

Precautions were necessary to prevent difficulties with the motion aftereffect. After prolonged viewing of motion in one direction, one usually perceives motion in the opposite direction even for stationary stimuli. This phenomenon has been observed from at least the time of Aristotle (Mather 1980). For this reason for each stimulus configuration, four different stimuli were employed, two grating orientations (+45, -45 degrees) and two veridical motion directions, left and right. As the direction of motion changes frequently between trials, the motion aftereffect was no longer likely to play a significant factor. Subjects reported their perceived direction of motion by pressing one of four keys corresponding to up, down, left and right and the percentage of motion in each direction was recorded.

The first set of experiments were conducted on two subjects, the author and a naive graduate student. Before conducting experiments, subjects were given a depth test in order to ensure that they were able to achieve sufficient depth perception using the mirrored stereoscopic set up.

2.2.2 Experiment A1: Depth Manipulation

As previously mentioned, numerous studies have demonstrated that depth affects the
way in which motion signals within the same area of space are integrated. Additionally
physiological studies of the middle temporal visual area (thought to be responsible for
motion processing) have shown that of directional inhibition occurs mainly between
motion signals with similar disparities (Bradley, Qian & Andersen 1995). Experiment A1
was designed to determine whether these depth affects persist when motions are present in
different areas of space.

Shimojo, Silverman & Nakayama (1989) manipulated the depth of the non-moving
intervening bars in the flanking barber-pole paradigm and found that when the barber
poles appear in front of intervening bars less motion integration occurs. Although such
manipulations may say something about the influences of intrinsic (real object boundaries)
and extrinsic contours (those resulting from occlusion), they do not reveal how motion sig-
nals within different depths are combined.

Rather than manipulating the depth of the intervening bars, in Experiment A1, the
depth of the flanking barber poles was varied. In this way the experiment examined the
interaction of motion signals occurring at different disparities.

The results suggest that motion information is selectively integrated between regions
of the same depth even when motion signals arise from different areas of space (See Fig-
ure 2-2). As the depth difference between the central and flanking barber poles increased
the amount of coherent motion decreased. The amount of disparity is important, rather
than just the sign of disparity. The results are analogous to those found for plaids in which
coherent motion is more likely when component gratings share the same disparity (Von
Grunau, Dube & Kwas 1993).
2.2.3 Experiment A2: Contrast Manipulation

Previous experiments have examined the effect of changing the contrast of motion signals arising from nearby spatial locations. When gratings within a plaid are of different contrasts, the perceived direction is biased in the direction of the higher contrast grating (Stone, Watson & Mulligan 1990). Although greater differences in contrast result in decreasing coherency, it is only at the extreme low end of the contrast scale that these effects occur. The aim of Experiment A2 was to examine the effect of changing contrast in motion signals arising in different spatial locations. Experiment A2 manipulated the

Figure 2-2. Experiment A1: Disparity Manipulation. Percentage of coherent motion plotted against the disparity difference between the central and flanking barber poles. For clarity, the relative disparities are plotted schematically below as if one had been viewing the stimuli from the side rather from the top.

Disparity Manipulation

![Disparity Manipulation Graph]

---

**Figure 2-2.** Experiment A1: Disparity Manipulation. Percentage of coherent motion plotted against the disparity difference between the central and flanking barber poles. For clarity, the relative disparities are plotted schematically below as if one had been viewing the stimuli from the side rather from the top.
contrast of the flanking barber poles while the central barber pole was held at a fixed contrast.

As in experiment A1, the results are analogous to those found for the combination of motion signals within the same spatial location. As the contrast of the flanking barber poles is decreased, the central barber pole is seen to be moving with the flanks less frequently (See Figure 2-3). This indicates that as flanking barber poles contrast decreases the flanks have less influence on the central barber pole. Unlike the depth effects found in Experiment A1, the effect of contrast is only manifested at the very low end of the contrast scale. At more moderate contrast differences, the change has little influence on the perceived direction of motion.

2.2.4 Experiment A3: Phase Manipulation

The influence of depth and contrast on motion integration examined in the previous experiments has previously been examined, although only within a single spatial location, using plaids. Experiment A3, however, examined a factor, phase, which cannot be examined using the plaid paradigm. Specifically, the phase difference between the central and flanking barber poles was varied in Experiment A3.

The percentage of vertical motion was found to decrease as the phase difference between the central and flanking barber poles increased (See Figure 2-4). The results suggest that motion information is preferentially combined between areas whose spatial frequencies have similar phase. This result is somewhat surprising as one might not expect a set spatial frequency tuned filters to be sensitive to phase. One possible explanation for
the alignment of individual line of the barber pole across the occluders increases the likelihood of grouping the motion across the occluders.

2.2.5 Experiment A4: Spatial Frequency Manipulation

From the study of coherency in plaids, there is reason to believe that motion signals are preferentially combined between areas having similar spatial frequencies (at least within the same area of space). The purpose of Experiment A4 was to examine whether a similar rule applies to motion signals in different spatial locations.

As Experiment A3 demonstrated that phase can effect the integration of motion sig-
nals between the barber poles, the relative phase of the barber poles was randomized in Experiment A4. Otherwise changes in spatial frequency would be confounded with unavoidable changes in phase. Note that changes in spatial frequency are also unavoidably confounded with changes in temporal frequency, as the speed of the gratings was held constant.

The percentage of vertical motion was found to decrease as the frequency difference between the central and flanking barber poles increased, indicating that connections between motion processing units may be spatial frequency dependent across space.
2.2.6 Series A: Discussion

Experiment A1 demonstrated that in addition to playing a role in the integration of motion signals from the same spatial location, depth relationships are also important for the integration of motion signals arising from different areas visual space. Unlike the experimental work of Shimojo et. al (1989), the disparity between the moving areas was manipulated rather than the intervening bars. The results suggest that not only the sign of disparity is important for the way in which motion signals combine across space, but also that the magnitude of the disparity difference is critical. The more similar the depth, the
more motion signals are selectively combined between regions, independently of the depth of intervening bars.

Experiment A2 demonstrated that as flanking stimuli decrease in contrast, their influence on a central stimuli decreases. This may be related to the finding by Kirita (1988) that the strength of the barber illusion decreases with decreasing grating contrast. Flank influence drops off steeply at very low contrast. It is also worth noting that subjects reported a transparency phenomenon for low flank contrasts. Subjectively, at low contrasts, the peripheral stimuli appear to be covered by a transparent film.

There are two possible explanations for the weakening of flank influence as flank contrast decreases. The weakening could be the result of the lower contrast flanks having a smaller signal strength due to their lower contrast. Alternatively, since the lower contrast flanks are perceptually less similar to the central barber pole patch their influence may be weakened due to this perceptual difference. Section 2.3.4 examines these two possibilities.

Experiment A3 showed a relatively linear relationship between phase and the integration of motion signals. This result is somewhat surprising, for while it is known that motion selective neurons are selective to disparity and spatial frequency and that their response may be altered by the amount of contrast, it is not clear that motion selective neurons are selective to phase. Co-linearity of grating edges may be playing a role. As co-linearity is thought to be processed within the form system, this suggest form processes may be influencing processing within the motion system. Experiment A4 indicated that motion signals are preferentially combined between regions having similar spatial frequencies.
From a modeling perspective, these experiments provide information about how neurons tuned to features such as spatial frequency and disparity might interact and over what distances in frequency and disparity space these interactions take place. They suggest that motion processing units interacting across space excite each other selectively, showing more interaction when frequency and disparity are similar. Furthermore, the tuning is diffuse, showing something like a gaussian distribution of spatial frequency and disparity. Furthermore they suggest that the form system may have some influence on the motion processing system.

2.3 Experimental Series B

The experiments in Series B had several objectives. The first objective was to examine how the factors examined in Series A interact. Specifically, when factors are manipulated along more than one dimension, what are the effects on the perceived direction of motion? The second objective of Series B was to examine the role of extrinsic and intrinsic terminators in more detail. The results of Shimojo, Silverman & Nakayama (1989) suggest that the predominant factor influencing the perceived direction of motion for a barber pole pattern is the classification of the terminators. Given the importance of similarity on the perceived direction of motion demonstrated in Series A, it is important to know how similarity relates to terminator classification. Does the extrinsic/intrinsic classification of terminators override the effects of similarity, is similarity more important, or do the two factors interact?

The third and final objective of the experiments of Series B was to examine the role
that contrast plays in more detail. Experiment A2 demonstrated the importance of contrast for determining how motion signals are combined between two areas of space. However, it isn’t certain from Experiment A2 whether the effect of contrast on direction perception is due to a difference in contrast similarity between the flanks and the central barber pole, or simply a weakening of the motion signal originating from the flanks as a result of their lower contrast. The experiments in Series B attempt to resolve this questions with a new stimulus arrangement.

2.3.1 Series B: Methods

The experiments of Series A demonstrated that characteristics such as spacial frequency, contrast and depth have a profound effect on the way in which motion signals between disparate areas of space are integrated. However it should be noted that the stimulus arrangement pattern employed in experiment Series A suffers from one major confounding factor. As will be discussed in great detail in Chapter 3, the addition of occluding bars of texture is in itself enough to influence the perceived direction of motion. Given the arrangement shown in Figure 2-6 subjects’ predominant perceived direction of motion is along the vertical axis. Shimojo, Silverman & Nakayama (1989) demonstrated that occluder depth influences the perceived direction of motion in the flanking barber pole configuration (Figure 2-6, A). They conjectured that the three barber pole patterns are more likely to group when the occluders are seen in front of the barber pole. However, informal observations of the single barber pole configuration (Figure 2-6, B) suggest that grouping may not in fact be the important factor. The results of Shimojo, Silverman &
Nakayama (1989) persist when the flanking barber pole patterns are removed, suggesting it is the depth of the occluders themselves and not the presence of the barber pole flanks that influence the perception of motion.

This means that the configuration used in Series A created a vertical bias for all experimental conditions. The experiments in Series B overcome this obstacle using a stimulus configuration consisting of a plus pattern (See Figure 2-7). The new pattern contains occluders and flanks along both the horizontal and vertical axis. For this arrangement pattern, the central barber pole can be captured either by the horizontal or vertical flanks. The new arrangement pattern helps to reduce any effects of motion bias due the occluders themselves as occluders are present along both axis of motion. Another advantage of the new arrangement pattern is that it allows factors to be placed in opposition. For example, one can manipulate spatial frequency on the vertical axis and contrast on the horizontal axis.

There remains one discrepancy between the horizontal and vertical axis in the plus-
shaped stimulus. As the eyes are mainly sensitive to horizontal disparities, only the vertical occluders have a strong disparity signal along the edge of the occluding rock patterns. This aberration can be remedied by using a cross shaped stimulus as shown in Figure 2-8. However, pilot experiments demonstrated that this stimulus configuration introduces a more serious problem, namely, depth capture. Depth capture occurs when the perceived depth of a particular region appears to be the same as a surrounding region rather its true depth (Spillman & Redies, 1981; Ramachandran, 1986; Watanabe & Cavanagh, 1992). As the grating patterns in the cross-stimulus are horizontally oriented for half of the stimuli they contain no information about binocular disparity. This allows for depth capture by the surrounding textured square. All subjects reported strong depth capture in pilot experiments using the cross stimulus.

Another, although less serious, problem with the cross stimulus is that the barber pole phenomenon itself was found to be substantially weaker when the barber pole pattern is not oriented along one of the cardinal axis. Subjects were more likely to perceive motion perpendicular to the orientation of the gratings. They often reported all five barber pole
regions moving together; a situation which never occurred for the plus-shaped pattern. For these reasons the cross shaped stimulus was abandoned. A reported experiments employed the plus-shaped stimulus. To mitigate any discrepancy between the horizontal and vertical axis all manipulations were performed for both horizontal and vertical axis and the results averaged (See Figure 2-9).

![Cross-shaped Stimulus: Central area and four flanking areas are occluded by a square frame. Has the disadvantage that the depth of the central barber pole is captured by the depth of the occluder.](image)

Figure 2-8. Cross-shaped Stimulus: Central area and four flanking areas are occluded by a square frame. Has the disadvantage that the depth of the central barber pole is captured by the depth of the occluder.

![Vertical Flanks](image) ![Horizontal Flanks](image)

Figure 2-9. To mitigate differences in the strength of disparity cues along the horizontal and vertical axis, all manipulations were performed both in the vertical and horizontal orientation and averaged.

### 2.3.2 Similarity in Opposition: Experiment B1a & B1b

The first experiment of Series B examined what happens when two factors of similarity are placed in opposition to each other. Previous work has emphasized the importance
of retinal disparity for motion perception (Yo & Wilson, 1992; Von Grunau, 1993; Qian et al., 1994; Bradely et al., 1995). Although color has some influence on the motion system (Krauskopf & Farell, 1990; Culham & Cavanagh, 1994), its effects are thought to be weak relative to that of luminance (Ramachandran & Gregory, 1978; Cavanagh, Tyler & Favreau, 1984; Livingston & Hubel, 1988, Mullen & Boulton, 1992). Given these results one might expect color differences to be discounted in the face of contradictory depth information. Experiment B1a & B1b examined this hypothesis.

Methods

Four subjects participated in this and all following experiments; the author and three naive subjects. All stimuli were displayed on a Silicon Graphics Reality Engine (model CMN-A011). Observers viewed stimuli through a mirror and prism haploscope, 70cm from the monitor, in a darkened room. Each square barber pole patch subtended an area of 2 x 2 of arc. The speed of the grating was 2.5/sec. Occluders were seen in front of the grating pattern with a disparity of 17 min relative to the grating.

Experiment B1a

Experiment B1a served to test the effect of flank depth on the perceived direction of motion. Although Experiment A1 in the previous series of experiments already demonstrated an effect for depth, Experiment B1a examines whether the effect persists for the new stimulus configuration and serves as a baseline to measure the influence of color similarity for the subsequent experiment. As expected, the larger the deviation of the perceived depth of one set of flanks from that of the central barberpole the less likely subjects
are to see the motion of the central barber pole captured by those flanks (See Figure 2-11, B1a - bottom line). Note however that there appears to be a slight asymmetry between crossed and uncrossed disparities. Asymmetries between crossed and uncrossed disparities have been supported by other depth experiments (Takeichi, Watanabe, & Shimojo, 1992; Hakkinen & Nyman, 1996).

**Experiment B1b**

Experiment B1b examined what happens when color similarity cues are placed in opposition to depth similarity cues. All conditions for Experiment B1b were the same as Experiment B1a apart from colors which were added to all the barber pole patches. The colors were chosen such in a manner that similarity of color between one pair of flanks and central barber pole was in opposition to similarity in depth between the other pair of flanks and the central barber pole (See Figure 2-10).

**Results**

As previously mentioned, to mitigate any discrepancies due to differences in binocular signals along the horizontal and vertical axis, for half the trials, depth similarity was on the vertical axis while color similarity was on the horizontal axis. The remaining half of the trials reversed the similarity axis. Experimental results from the reversed orientation trials were flipped and combined with normally oriented trials.

Examining the results, the first thing to note is that color similarity does affect whether one area of motion will capture another. When no color is present and the flanks are all have the same disparity as the central barber pole, motion is ambiguous. Subjects
are at about the 50 percent level for reporting horizontal motion (See Figure 2-11, B1a - bottom, Equal Disparity). However, when color cues are added subjects report motion along the axis of similar color on an average of 80 percent of the time. So although motion sensitive cells in brain are generally not color sensitive (DeYoe & Van Essen, 1985; Livingstone & Hubel, 1988) color can nevertheless have an impact on the perceived direction of motion.

The second major feature to be noted in the results is that when color cues are placed in opposition to depth cues they partially negate the effect of the depth cues. This effect is somewhat surprising as cells in the brain area thought to handle motion integration (MT)
are sensitive to depth (Maunsell & Van Essen 1983b) but not to color (Zeki 1974b). Interpolating from the data one can see that there is a point at which opposing depth and color cues cancel each other out (approximately -7 degrees of arc in Figure 2-11). At this point, despite both differences in color and depth, the motion the central barber pole is ambiguous.

It is difficult to gauge the relative strength of these cues when it comes to their influence on motion perception as they are measured on different scales. Furthermore, no attempt was made to measure how different types color changes alter the strength of the
color influence. Only one color difference was tried and there was no principled reason for choosing the two colors. The important point is simply that color does have an effect and that its influence is on the same order of magnitude as disparity. Recent neurophysiological results support the importance of color. Ringach, Hawken & Shapley (1998) report color and luminance affect each other mutually and in a non-linear manner in V1 neurons, the primary source of input to MT.

2.3.3 Terminators and Similarity: Experiment B2a & B2b

Shimojo, Silverman & Nakayama (1989) established an important role for terminator classification in the perception of motion. They suggest it is more likely that motion is integrated between two regions when then terminators along the edges of the two regions are classified as extrinsic (resulting from occlusion). However, it is not clear how effects due to terminator classification within a region are modified by similarity between regions. Shimojo et al. (1989) suggest the propagating effect of extrinsic terminators is abolished. If this is true, are the effects of region similarity negated? If not, which has a greater effect on motion disambiguation, terminator classification or region similarity? Experiments B2a & B2b examine these questions.

One feature of the plus-shaped stimulus used in the experiments of Series B is that it allows for the manipulation of extrinsic and intrinsic terminators within the same arrangement. Three basic depth configurations were employed in this set of experiments. The square configuration employed a surrounding rock patch whose depth was similar along the vertical and horizontal axis (Figure 2-12, top). For this configuration when all barber
pole patches have identical characteristics, subject’s perception of motion direction is ambiguous. **Bar configuration I**, on the other hand, has different depths for the vertical and horizontal occluding patches; the vertical occluding patches are seen as nearer than the barber pole regions (Figure 2-12, middle). The horizontal occluding patches are seen at the same depth. This produces extrinsic terminators along the horizontal dimension and intrinsic terminators along the vertical dimension. In accord with the idea that extrinsic terminators are discounted, subjects report a predominance of horizontal motion when presented with this stimulus. **Bar configuration II** is the opposite of **bar configuration I**; the horizontal occluders are seen at a nearer depth than the barber poles while the vertical occluders are seen at the same depth (Figure 2-12, bottom). Presented with this stimulus, subjects report a predominance of vertical motion. Sample stimuli for free-fusers are shown in Figure 2-13.

**Experiment B2a**

Experiment B2a served as a baseline. The square configuration was employed so terminators along both the vertical and horizontal axis were extrinsic. A similarity cue, spatial frequency, was chosen which was known to have a powerful effect on the integration of motion between areas (See Experiment A4). As expected, the results indicate that the perceived motion of the central barber pole patch is preferentially captured by flanks of the most similar spatial frequency (Figure 2-15, lower line). When all flanks have the same spatial frequency, the average perceived motion is near the point of ambiguity (there is a slight bias for horizontal motion). As the difference in spatial frequency between the cen-
tral barber pole and one set of flanks increases, the likelihood of motion capture by the opposing flanks increases. Experiment B2a reconfirmed the results of experiment A4 for the new stimulus configuration, namely, that motion is preferentially integrated between areas of the same spatial frequency.

**Experiment B2b**
The second part of the experiment examined the role of extrinsic and intrinsic terminators with respect to similarity using *bar configurations I* and *II*. The axis of spacial frequency similarity was placed in opposition to extrinsic terminator classification. In other words, for each bar configuration, similarity in spatial frequency was arranged in such a way that it was likely to capture motion along one axis of the plus, whereas the arrange-
ment of terminators was likely to capture motion along the opposite axis (Figure 2-14). The spatial frequency of the flanks was then manipulated as in Experiment B2a. All other conditions were the same as Experiment B2a.

Results

The first thing to note is that when all barber pole spatial frequencies are the same, the predominant perceived direction of motion for the central barber pole is along the axis that passes under the extrinsic terminators. This is to be expected given prior results which demonstrate the extrinsic terminators are discounted. However, the addition of spatial frequency differences along that axis results in a change of perception. For large changes in spatial frequency (low and high) the predominant perceived direction of motion is
reversed (Figure 2-15). Subjects appear to discount terminator classification, perceiving motion along the axis of spatial frequency similarity. As in experiments B1a & B1b, one can interpolate the data, finding a point at which terminator classification and spatial frequency similarity are in balance. At this point the motion is ambiguous (near 16 and 60 cycles/degrees).

Shimojo et al. (1989) suggest that “3 segmented apertures.....can be integrated into an elongated larger aperture only if the regions in-between the apertures have crossed dispar-
ities relative to the aperture and the diagonal grating” p. 623. Although they also suggest that similarity may play a role, the results of experiment B1a & B1b, go even further, demonstrating that similarity between regions can in fact override the effects of disparity. The resulting perception is one which takes all factors into account.

### 2.3.4 Contrast Effects: Experiments B3a & B3b

Experiments B3a & B3b examine how contrast affects the integration of motion signals in different areas of space. In the previous series of experiments, Experiment A2 demonstrated that contrast does have an effect on the integration of spatially separated motions. Motion of a central region is less likely to be captured by lower contrast flanks. However, from the results of Experiment A2, one can not conclude that the effect is the result of a perceived difference in similarity between the regions. The resulting effect could also be due to the fact that lower contrast flanks produce a weaker motion signal and therefore lower contrast flanks are less likely to capture the central barber pole. Experiments B3a & B3b were designed to clarify this difference.

**Experiment B3a: Flank Contrast Only**

Experiment B3a served as a control and as a replication of Experiment A2 but using the new plus-stimulus. The contrast of one set of flanks was manipulated. When all flanks are at full contrast, motion was found to be ambiguous. Subjects were equally likely to report vertical and horizontal motion. As the contrast of the flanks along one axis of the plus decreased subjects were more likely to see motion along the opposite axis (Figure 2-
16, bottom line). These results are in accord with those found in Experiment A2.

**Experiment B3b: Central and Flank Contrast**

To determine whether the important factor for the effect of contrast on the likelihood of integration is similarity or signal strength, Experiment B3a was repeated with the addition of contrast changes to the central barber pole. The central barber pole patch was given the same contrast as the manipulated flanks. If the previously measured contrast effects were due to signal strength than one would expect the results of Experiment B3b to follow the same general pattern as B3a. The central barber pole patch would be captured by the high contrast flanks regardless of its own contrast as the high contrast flanks carry a large signal. On the other hand, if similarity is the important factor one would expect the opposite effect. The central barber pole patch would be captured by the flanks with the same contrast as the central patch regardless of signal strength.

**Results**

The central barber pole patch is preferentially captured by the flanks with the most similar contrast, not the larger signal strength (Figure 2-16, top line). This suggests that similarity, not signal strength is the predominant factor.

**2.3.5 Series B: Discussion**

Experiments B1a and B1b demonstrated that color similarity has a role in determining how motion in different regions of space is combined. Color similarity can override other cues such as disparity which are thought to play a major role in motion integration.
Experiments B2a & B2b demonstrated that extrinsic and intrinsic terminators, when present, are not the sole deciding factor for determining when grouping between regions should take place. Region similarity also plays an important role and can even override the perceptual organization provided by terminators. Terminator classification, rather than determining grouping, is one of many factors which contribute to the way in which motion signals in differing areas of visual space are combined. Experiments B3a & B3b demonstrated that perceptual similarity between regions rather than differences in signal strength...
determines whether two regions of motion will be integrated.

### 2.4 General Discussion

Recently another study has examined the way in which motion signals are integrated between separated spatial regions. Kim and Wilson (1997) examined the effect of adding a ring containing a moving grating around a small circle containing a grating moving in a different direction. They manipulated the contrast, speed and direction of the surrounding grating, measuring its effect on the perceived direction of motion in the central circle. Kim and Wilson found that the perceived direction of the central circle was shifted away from the direction of motion of the surrounding ring. Such *motion repulsion* has previously been demonstrated within the same spatial region using random dot patterns (Marshak & Sekuler, 1979). Kim and Wilson found that the degree of the motion shift depended on the contrast and spatial frequency of the ring. The effect is strongest when the surrounding ring has the same spatial frequency and contrast and the central circle.

The results of Kim and Wilson (1997) complement those of the current study. Their study examined antagonistic processes between motion in one region of space and motion in another when the two regions move in *different* directions. They demonstrated that the antagonistic effect was strongest when the two regions matched in similarity (of spatial frequency, contrast and speed). The current study, on the other hand, examined facilitory processing between two separated regions of motion in the *same* direction. Like the study of Kim and Wilson, the strongest effects were found when the regions of space were the most similar. Together these studies suggest there are both antagonistic and facilitory
interactions between widely separated motion processing units. The strength of these excitatory and facilitory effects depends on visual similarity between the regions in terms of features such as color, depth, and spatial frequency and whether the two regions move in similar or different directions.
Chapter 3

Monocular Occlusion Cues Alter the Influence of Terminator Motion in the Barber Pole Phenomenon

3.1 Introduction

Determining whether an image boundary is inherent to a surface itself or the product of occlusion by a second surface is critical to understanding the spatial layout of visual scenes (Rubin, 1915, 1958). When two areas, figure and ground, share a common boundary, the figure region is said to “own” the boundary. The figure’s appearance, i.e. its perceived shape, depth, or lightness, is in turn strongly affected by the boundary, while ground is only minimally affected. (Koffka, 1935; Coren, 1969; Nakayama & Shimojo, 1992). Furthermore, there is a tendency to perceive the ground as extending behind the figure. While border ownership is often examined in static displays, it can also play an important role in the perception of moving displays.

The barber pole illusion offers a compelling example of the importance of boundary classification. Wallach (1935) showed that under normal conditions the perceived direction of motion of a diagonal grating pattern drifting behind a rectangular aperture depends upon the direction of elongation of the aperture (Wuerger, Shapley & Rubin, 1996). The grating pattern is seen to move along the axis of the longer dimension of the rectangle. The prevailing explanation for this phenomenon suggests that the perceived direction of motion results from the integration of motion signals from the areas where the grating pat-
tern is terminated by the edges of the aperture. Due to the aperture problem, the local velocity signals along the grating are inherently ambiguous. Terminators, however, produce unambiguous motion signals. As a vertically elongated aperture has larger number of terminators with vertical trajectories, vertical motion prevails.

Not all line terminators signal ends of physical objects, however. Wallach (1935) noted that “free-line” endings, those which are not associated with any aperture edge, are treated differently than line endings at the edge of an aperture (Wuerger et al. 1996). More recently, Shimojo, Silverman & Nakayama (1989) demonstrated that terminator classification plays an important role in the aperture problem. Vertical motion no longer dominates when a vertically elongated barber pole is presented with uncrossed disparities (so the diagonal lines appear to be seen through an aperture). Evidently, intrinsic boundaries (real object boundaries) are treated differently than extrinsic boundaries (those resulting from occlusion). Terminators along extrinsic boundaries are excluded when processing the direction of motion for the barber pole region, as the motion of such terminators may be unrelated to the actual global motion of that region.

Shimojo et al. (1989) suggest that it is not disparity information that is the critical factor, but rather depth information, which can be carried by disparity or other factors including monocular depth cues. However, they argue that the effect of monocular depth cues is weaker than that of disparity cues. Specifically, they demonstrate that directional bias in the barber pole effect is reduced when the pattern is viewed monocularly. For the binocular condition, they suggest monocular occlusion cues, specifically local T-junctions at the end of individual lines of the grating (which signal occlusion), are overridden by zero-dis-
parity binocular cues (which fail to signal occlusion). Shimojo et al. conclude that although disparity is not critical for terminator classification, its effect is significantly stronger than that of monocular depth cues.

There is, however, evidence that monocular segmentation cues can have a pronounced effect on motion processing. Lorenceau and Shiffrar (1992) demonstrated that integration across space is more likely when the salience of terminators is decreased through jagged apertures, low contrast, viewing at large eccentricity, or when terminators are isoluminant with the background (Shiffrar & Lorenceau 1996). Additionally, binocular depth cues are not always exploited by the motion system. Watamaniuk & McKee (1995) examined subjects’ ability to track dot motion behind occluders. In a two-interval forced-choice procedure, subjects reported which of two stimuli contained a signal trajectory. Unlike monocular occlusion cues, such as the addition of opaque occluders or motion discontinuities, differences in depth were not sufficient to suppress motion noise and allow for signal detectability.

There are three primary purposes to our study. The first is to examine the importance of monocular occlusion cues using experiments within a single paradigm. We measure change in the average perceived motion direction resulting from the addition of occluding patches along either the vertical or horizontal edges of a barber pole pattern (See Figure 3-1). Experiment 1 measures the influence of occluders at a different disparity than the barber pole patch. Experiment 2 measures the influence of occluders with zero-disparity difference. Our results demonstrate that although the addition of a disparity difference results in a change in the direction of motion perception, the change is almost as strong
when only monocular cues indicate occlusion.

The second purpose of our study is to measure the degree to which motion signals from extrinsically classified terminators are discounted. Shimojo et al. (1989) suggest that intrinsic terminators produce strong propagation of velocity signals to the interior of lines, where they exert a biasing effect on local ambiguous signals. On the other hand, they suggest the propagating effect of extrinsic terminators is abolished. Rubin & Hochstein (1993), however, provide evidence that terminators do still retain some influence even when a barber pole is presented with uncrossed disparities. The perceived direction of lines viewed through an aperture is not perpendicular to the lines, as would be predicted if motion signals from extrinsic terminators were ignored, but deviates in the direction of terminator motion.

Figure 3-1. Three stimulus configurations used for Experiment 1. The isolated configuration, (a), employed isolated barber pole patches, the vertical configuration, (b), employed barber pole patches with vertical occluders on the left and right sides of the barber pole, and the horizontal configurations, (c), employed horizontal occluders on the top and bottom of a barber pole patch. The occluding patch for (b) and (c) was textured with a rock pattern.

The second purpose of our study is to measure the degree to which motion signals from extrinsically classified terminators are discounted. Shimojo et al. (1989) suggest that intrinsic terminators produce strong propagation of velocity signals to the interior of lines, where they exert a biasing effect on local ambiguous signals. On the other hand, they suggest the propagating effect of extrinsic terminators is abolished. Rubin & Hochstein (1993), however, provide evidence that terminators do still retain some influence even when a barber pole is presented with uncrossed disparities. The perceived direction of lines viewed through an aperture is not perpendicular to the lines, as would be predicted if motion signals from extrinsic terminators were ignored, but deviates in the direction of terminator motion.
We attempt to answer several questions. Is the extrinsic/intrinsic classification an all-or-none attribute? Can one terminator be classified as having stronger extrinsic characteristics than another, and if so how will this differential classification affect motion integration? In our paradigm only half the terminators (vertical or horizontal) are extrinsic, allowing for a comparison with the remaining intrinsic terminators and a measurement of extrinsic terminator influence relative to that of intrinsic terminators.

Our results demonstrate that even terminators with the strongest extrinsic characteristics still have an influence on motion processing. Secondly we demonstrate that terminators can be classified as more or less extrinsic, resulting in a continuum of effects on motion perception. When there is strong evidence for extrinsic terminators, there is a considerable change in the perception of motion direction. When extrinsic terminators are weakly supported, there is a correspondingly smaller change in the perceived direction of motion.

The third purpose of our study is to determine the relative importance of the different monocular occlusion cues since the monocular conditions that give rise to terminators being classified as extrinsic have not been adequately characterized to date. Although T-junctions have been suggested as the relevant monocular depth cue (Guzman, 1969, Shimojo et al., 1989; Rubin, 1997), we seek to explicitly determine which of the numerous monocular depth cues are important for extrinsic/intrinsic terminator classification and the degree to which each affects motion integration. We attempt to find the minimal stimulus necessary to produce a change in the average perceived direction of motion for a barber pole pattern. The role of contrast relationships between background and occluding pat-
terns, texture, amodal completion, T-junctions and illusory occlusion are all examined.

The principle result of the monocular experiments is that no single monocular occlusion cue is responsible for terminator classification or changing the perceived direction of motion. Rather a multitude of cues participate in the perception of occlusion and their differing strengths result in a continuum of effects on terminator classification and motion perception. Although T-junctions play a prominent role, they are not strictly necessary for occluders to have an influence on motion processing.

3.1.1 Experiment 1: Textured Occluders With Depth

The purpose of the first experiment was to measure how the influence of terminators on the perceived motion of ambiguous line interiors is attenuated when the terminators are extrinsic. To do so, we created barber pole stimuli in which only one axis of terminators (vertical or horizontal) was extrinsic (See Figure 3-1). By measuring subjects’ dominant perceived direction of motion, we were able to find the barber pole aspect ratio at which horizontal and vertical motion were in balance, allowing us to calculate the relative attenuation in the influence of the extrinsic terminators with respect to the intrinsic terminators.

Method

Subjects. Four subjects were employed for this experiment, one of the authors and three others who were naive to the purpose of the experiment. The same subjects participated in all subsequent experiments. All subjects had normal or corrected to normal vision and those other than the author received compensation of $8/hour for their participation in the experiment.
Simuli.

The stimuli consisted of translating square-wave gratings oriented at 45° viewed through rectangular apertures of different aspect ratios. The light and dark stripes of the grating were 1.6 and 81.2 cd/m² respectively. The experimental factors were stimulus configuration and aspect ratio of the barber pole. Three stimulus configurations were employed: the barber poles were presented either in isolation, sandwiched between two vertical textured bars, or sandwiched between two horizontal textured bars. For the sandwiched cases, vertical textured bars were placed on the left and right sides of the barber pole and horizontal textured bars placed on the top and bottom of the barber pole (See Figure 3-1).

The aspect ratio of the barber pole was varied. For the isolated case, thirteen levels were used, with the aspect ratio of the barber pole varying from 2.4:1 to 1:2.4. Only six levels were used for the occluded conditions, as an increase in the aspect ratio along the occluded dimension would be hidden by the occluder. For the vertical case, the aspect ratio varied from 2.4:1 to 1:1. For the horizontal case, the aspect ratio varied from 1:1 to 1:2.4. (See Legend, Figure 3-2).

The isolated configuration served as a baseline for the barber pole effect. The vertical and horizontal configurations served to measure the influence of occlusion on the perception of grating motion. As the occluding bars were presented with one orientation (vertical or horizontal), they created extrinsic terminators along one orientation of the barber pole, leaving terminators along the opposite orientation as intrinsic. We expected that the influence of the terminators along the occluded edge would be weakened relative to
non-occluded edge, resulting in an increase in motion perceived in the direction orthogonal to the occluders.

Two grating orientations (45° and -45°) and two directions of motion (left and right) were employed for each stimulus configuration, making four different stimuli per configuration. Stimuli with motion in the vertical direction (up and down) were not necessary, as each upward or downward motion looks identical to one of the horizontal motions. The two veridical directions of motion served to minimize any impact of the motion after effect.

All stimuli were displayed on a Silicon Graphics Reality Engine (model CMN-A011). Observers viewed stimuli through a mirror and prism haploscope, 70cm from the monitor, in a darkened room. The square barber pole (aspect ratio of 1:1) subtended an area of 2° x 2° of arc. The largest barber pole patch (aspect ratio of 2.4:1) subtended an area of 2° x 4.8°. The speed of the grating was 2.5°/sec. The occluders subtended a 2° x 5.2° patch. They were seen in front of the grating pattern with a disparity of 17 min relative to the grating.

**Procedure**

Each pattern was presented for 300 msec followed by a black screen containing a fixation point. The fixation point served as cue for subjects to respond. Given four keys (one for each of the cardinal directions), subjects were instructed to press the one that corresponded best to their perceived direction of motion for the grating pattern. The key press also initiated the next trial after a small delay. Each subject completed 6 blocks of 50 trials. All factors were intermixed and presented in random order.
Results and discussion

For the isolated configuration trials, Wallach’s barber pole phenomenon was confirmed. There were some individual differences between subjects, some showing an overall bias for vertical motion, others showing an overall bias for horizontal motion. In all cases, however, the greater the horizontal size of the grating patch, the more likely subjects were to report horizontal motion and the greater the vertical size, the more likely subjects were to report vertical motion. For an aspect ratio of 1:1, average subjects’ response was near the point of ambiguity (perceived horizontal motion 50% of the time) (See Figure 3-2).

A clear bias was found in subjects’ perception of the barber pole direction of motion when textured bars were added to the sides of the barber pole. When vertical textured bars were added to the left and right of the barber pole, subjects were more likely to see horizontal motion (ANOVA: Isolated vertical vs. vertical with occluders: $F(1,3) = 198.87, p < 0.0008$). A corresponding bias for vertical motion was found for the addition of horizontal textured bars (Isolated horizontal vs. horizontal with occluders: $F(1,3) = 70.21, p < 0.0036$) (See Figure 3-2).

Note that the addition of occluding bars did not abolish the influence of terminators along the occluded edge. Rather the influence of extrinsic terminators was merely attenuated with respect to the intrinsic terminators. Elongation of the barber pole aperture still influenced the perceived direction of motion. This enabled us to measure the attenuation of the extrinsic terminators with respect to the intrinsic terminators by comparing the aspect ratio at which a balance is achieved between horizontal on vertical motion.
Figure 3-2: EXPERIMENT 1: The addition of textured bars with disparity differences. The average percent horizontal motion reported by subjects is plotted against barber pole aspect ratio with solid lines for the isolated (♦), vertical (▲), and horizontal (■) configurations. The legend below contains icons depicting the corresponding aspect ratios employed in the experiment. Dotted lines are sigmoid estimations for the response curves. Note that only portions of the sigmoid are visible for the vertical and horizontal configurations. For the isolated configuration 13 different aspect ratios were employed ranging from 2.4:1 to 1:2.4. For the vertical and horizontal configurations, 6 different aspect ratios were employed. Note that the set of aspect ratios used for the vertical and horizontal configurations only overlap for the 1:1 case, as an increase in aspect ratio along the occluded sides is invisible. To reduce the number of experimental trials, only the isolated configuration employed the 1.2:1 and 1:1.2 aspect ratio.
As a heuristic for determining attenuation, subjects’ average percent horizontal motion was pooled and estimated with a sigmoid function using the simplex search method for function minimization (Nelder & Mead, 1965). The specific sigmoid function used is given by:

\[
\frac{100}{1 + e^{-\alpha x + \beta}}
\]  

(EQ 1)

Where \( \alpha \) controls the slope and \( \beta \) the position of the sigmoid. Note that only the tail ends of the sigmoid are visible for the occluded configurations (Figure 3-2, dotted lines). To find a rough estimate of the aspect ratios that result in a balance of horizontal and vertical motion, the points at which the function crossed 50% were calculated. Accordingly, an estimated balance for vertical and horizontal motion for a barber pole with vertical occluders occurred at an aspect ratio of 2.6:1, and a balance for a barber pole with horizontal occluders occurred at an aspect ratio of 1:2.5. Taking any individual subjects’ initial horizontal or vertical bias for the isolated barber pole into account, these aspect ratios correspond to a 61% and 60% reduction of the influence of terminators along the occluded sides respectively. (See Appendix A1 for details.)

3.1.2 Experiment 2: Textured Occluders Without Depth

The first experiment demonstrated that the addition of occluding surfaces, with crossed disparity along either the horizontal or vertical axis of a barber pole, biases the perception of barber pole motion in the direction orthogonal to occluder orientation. Pre-
sumably this effect results from the classification of the terminators along the occluded edge as extrinsic, thereby weakening their contribution to motion disambiguation.

Shimojo & Nakayama (1990) argue for a close interaction between depth and motion mechanisms, suggesting a role for disparity sensitive-cells in MT. They also note that monocular depth cues may still influence motion disambiguation, but argue that such cues may be significantly weaker. However, there is convincing evidence that surface decomposition occurs rapidly, strongly affecting the early stages of visual processing even for stimuli without binocular depth cues (Watanabe & Cavanagh, 1993b). Monocular occlusion cues can change the perception of apparent motion (Ramachandran & Anstis, 1986; Ramachandran, Inada & Kiama, 1986), and when monocular occluders flank an apparent motion display, they can alter the perception of the entire display (Watanabe & Cole, 1995). Experiments 2 was designed to examine the importance of disparity information by comparing Experiment 1 with a case in which the binocular depth cues suggest there is no disparity difference between the barber poles and the occluders.

Method

The second experiment was conducted under the same conditions as the first, although subjects did not view stimuli through a haploscope. Thus, the textured bars were seen at the same depth of the barber poles. In addition to the same four subjects reported in Experiment 1, two more subjects participated in the second experiment. As subjects in Experiment 2 were not using a haploscope, viewing distance was not strictly enforced, but the stimuli subtended roughly the same angles as in the previous experiment.
The sequence of experiments reported here reflects the order in which the authors felt made for the most clear presentation rather than the order in which subjects participated in the experiments. There was no chance of contamination of the zero-disparity conditions by the binocular condition, as all subjects participated in the zero-disparity conditions before the binocular condition. There remains some possibility that there may have been ordering effects among other experiments. The experiments presented in this paper were conducted over a period of five months. As the results of the initial experiments drove the design of later experiments, it was not possible to fully counter-balance the conditions across experiments.

Results and discussion

Despite the use of zero-disparity, the addition of occluders still produced significant biases for both the vertical (vertical isolated vs. vertical with occluders: F(1,5) = 71.26, p < 0.011, Figure 3-3b) and horizontal (horizontal isolated vs. horizontal with occluders: F(1,5) = 39.76, p < 0.0032, Figure 3-3c) occluders. The bias was reduced in comparison with those seen in Experiment 1, but the reduction was not significant (horizontal occluders, Experiment 1 vs. Experiment 2: F(1,5) = 4.37, p > 0.1278, vertical occluders, Experiment 1 vs. Experiment 2: F(1,5) = 7.94, p > 0.0668). Approximating subjects’ results with a sigmoid function and reading off the balance points as before, a balance in vertical and horizontal motion was found for an aspect ratio of 2.3:1 for vertical occluders and 1:2.2 for horizontal occluders, corresponding to an attenuation in terminator influence of 57% and 55% respectively.

The changes in the motion biases were relatively small from 61% and 60% for
occluders with a disparity difference to 57% and 55% without. The occluders maintain almost all of their ability to influence motion disambiguation despite disparity cues which indicate the barber poles and occluders are at the same depth. This suggests that although disparity is an important cue for determining the relevance of terminators in motion integration, it is not strictly necessary, nor is it even necessarily the strongest such cue in a given display. In this particular case, monocular occlusion cues (indicating occlusion), override disparity cues (indicating the barber pole is in the same depth plane). Compare this with the case of Shimojo et al. (1989) in which disparity cues dominate. When zero-disparity barber pole patterns (without occluders) are employed, the barber pole effect is weakened as binocular cues override monocular cues.

Shimojo et al. (1989) suggest two possible mechanisms by which depth information could influence motion perception. Either other cortical areas that process form cues can influence motion processing areas, or disparity-specific cells within the motion processing areas selectively inhibit end-stopped cells that code terminator motion. The results of Experiment 2 support the idea of influence from other cortical areas. The existence of selectively inhibiting, disparity-specific cells can not be solely responsible, as disparity cues in Experiment 2 do not support occlusion.

3.1.3 Experiment 3A & 3B: Occluder Contrast

Experiment 2 would also seem to suggest that form perception has an early and significant influence on motion disambiguation, as the bias occurs within the first 300 msec of stimulus presentation. Despite the lack of disparity cues, the addition of occluding sur-
Figure 3-3: EXPERIMENT 2: The addition of textured bars without disparity differences. The average percent horizontal motion reported by subjects is plotted against barber pole aspect ratio with solid lines for the isolated (●), vertical (▲), and horizontal (■) configurations. The legend below contains icons depicting the corresponding aspect ratios employed in the experiment. Dotted lines are sigmoid estimations for the response curves. Note that only portions of the sigmoid are visible for the vertical and horizontal configurations.
faces influences motion disambiguation. Presumably this occurs as the terminators along
the occluded edge appear to be occluded and their influence is therefore attenuated.

Alternatively, the addition of texture near the occluded terminators might somehow
have diminished the salience of the nearby terminators. Experiments 3A and 3B were
control conditions to ensure that the bias found in Experiment 2 was not due to an interac-
tion between the motion and texture patterns. In particular, we examined the effect for
non-textured occluding patches. Additionally, we changed the relative contrast between
the occluded and unoccluded sides of the barber poles to ensure that contrast polarity was
not a factor.

Method

Simuli. Stimuli sizes and configurations were the same as Experiment 2 and the same
set of subjects were employed. However, we used non-textured uniform bars. Experiment
3A, employed a white background and intermediate grey occluders. Experiment 3B,
employed an intermediate grey background and white occluders (See Legend, Figure 3-4).

Results and discussion

No difference was found between the biasing effect of white bars on a grey back-
ground and grey bars on a white background (white horizontal vs. grey horizontal: F(1,4)
= 0.13, p > 0.73; white vertical vs. grey vertical: F(1,4) = 2.22, p > 0.2). However, the
effect of the white and grey bars was somewhat weaker than that of the textured bars of the
previous experiment. (See Figure 3-4). This difference was significant for three of the
four cases (grey horizontal vs. textured horizontal: F(1,4) = 14.44, p < 0.0191; grey verti-
cal vs. textured vertical: F(1,4) = 5.28, p > 0.0832; white horizontal vs. textured horizon-
tal: $F(1,4) = 8.13, p < 0.0464$, white vertical vs. textured vertical: $F(1,4) = 13.71, p < 0.0208$).

As before, the attenuation of terminator influence was estimated by finding the balance point between horizontal and vertical motion. The resulting balance ratios were 2.0:1 and 1:1.7 for white occluders and 2.0:1 & 1:2.1 for grey occluders. The white occluders produced a 48% reduction in the influence of vertical terminators and a 41% reduction for horizontal terminators. Likewise, the grey occluders produced at 49% reduction in the influence of vertical terminators and a 52% reduction for horizontal terminators.

These results support the idea that the bias in the perceived direction of motion in Experiment 2 was not merely result of a change in salience of the terminators along the occluding edge resulting from contrast differences between the terminators and the occluding edge or interference from the texture pattern. As no disparity cues were available to signal occlusion, the results further support the idea that monocular depth cues can have an early and strong influence on motion disambiguation.

The results of Experiments 3a & 3b do, however, indicate that texture played a small role in the previous experiments, as the bias for untextured occluders was slightly less than for textured occluders. This suggests either that the addition of texture created a stronger perception of occlusion, or that the texture itself makes motion signals from terminators along the occluded border somewhat less efficacious.

### 3.1.4 Experiment 4: Amodal Completion or Terminator Classification?
Figure 3-4.EXPERIMENTS 3A & 3B: The addition of grey bars on a white background, and white bars on a grey background. The average percent horizontal motion reported by subjects for both background types is plotted against barber pole aspect ratio with solid lines for the isolated (●), vertical (▲, ■), and horizontal (◆, □) configurations. The legend below contains icons depicting the corresponding aspect ratios employed in the experiment. Dotted lines are sigmoid estimations for the response curves.
One of the consequences of border assignment is that there is a tendency to see the ground (the object that does not ‘own’ the border) as extending beneath the figure (the object that owns the border) (Rubin, 1915, 1958). The previous experiments indicate that an occluding form can have a major effect on motion disambiguation. However, it remains to be seen whether the border between the occluding form and the barber pole is the only controlling factor. Experiment 4 examines whether the rest of the occluding form plays a role.

The figure-ground relationships suggest that the occluded barber poles are being amodally completed behind the occluders. An occluded barber pole may be seen as a larger unoccluded barber pole that continues behind the occluders (See Figure 3-5). Experiment 4 was designed to determine if border classification is the only controlling factor. In other words, do the occluding forms need to be wide enough to accommodate an amodal completion of the barber poles? Experiment 4 examined this question by decreasing occluder thickness.

![Figure 3-5](attachment:image.png)

Figure 3-5. A role of amodal completion? The figure on the left could be perceived as a horizontal barber pole extending beneath two vertical occluding patches (depicted in the center). The grey grating symbolizes the invisible, yet amodally competed percept. In some sense, this may be equivalent to the non-occluded barber pole depicted to the right.

**Method**

The same experimental conditions were used as before, but the occluding bars were
significantly thinner, subtending approximately a 0.4° x 4.8° patch. (See Legend, Figure 3-6.)

**Results and discussion**

The thin occluders produced a large bias (thin vertical vs. isolated vertical: $F(1,5) = 42.19$, $p < 0.0013$; thin horizontal vs. isolated horizontal: $F(1,5) = 39.72$, $p < 0.0015$). The measured effect was slightly less than that produced by the thick occluders. However, the difference between thick (Experiment 4) and thin (Experiment 2) bars failed to reach significance (thin vertical vs. thick vertical: $F(1,5) = 3.75$, $p > 0.1482$; thin horizontal vs. thick horizontal: $F(1,5) = 1.79$, $p > 0.2729$). Balance ratios were 2.1:1 and 1:2.2 for vertical and horizontal occluders respectively. Thin vertical occluders produced at 52% reduction in vertical terminators. Thin horizontal occluders produced a 55% reduction in horizontal terminators.

The results of Experiment 4 indicate that classification of the terminators at their border is the major factor influencing motion disambiguation. If indeed amodal completion does occur, there is little need for a wide zone of amodal completion behind the occluders.

**3.1.5 Experiment 5: Occlusion of One Edge**

The previous experiments demonstrated that the addition of occluding edges can bias the perceived direction of a barber pole pattern. The results are consistent with the idea that the influence of terminator motion is attenuated (but not abolished) when the terminators appear to have been created extrinsically. A sufficient number of extrinsic terminators can override intrinsic terminators and dominate motion disambiguation.
However, it remains to be seen whether a weakening of terminators is sufficient to explain the change in perception. To explore whether the bias is solely governed by the number of terminators and their classification, Experiment 5 examined barber pole patterns with only one occluded edge. We predicted that occluding only one edge would still
attenuate the influence of terminators along that edge, but the extent of the bias was unknown.

If an attenuation of terminator influence is sufficient to account for the bias, the results from Experiment 2 can be used to make a prediction for the aspect ratios required to balance horizontal and vertical motion when an occluder is added to only one side of a barber pole. Experiment 2 found a 57% reduction in the influence of vertical terminators, and a 55% reduction for horizontal terminators for the appropriately added occluders. If only terminator classification is responsible for the bias, we would then expect the same attenuation of terminator influence to occur, but for only half of the terminators. Taking the measured attenuation from Experiment 2 and applying it to half the terminators, we can predict a balance between horizontal and vertical motions at aspect ratios of approximately 1.4:1 and 1:1.38 respectively for barber poles with one side occluded. (See Appendix: Part II for details).

Method

Experimental conditions and stimuli were the same as Experiment 2, but only one side of the barber pole pattern was occluded (See Legend, Figure 3-7). Results for a horizontal occluder above the barber pole and a horizontal occluder below the barber pole were pooled as were occluders to the left and right of the barber pole for the vertical conditions.

Results and discussion

As expected the occlusion of only one side of the barber pole still results in a bias in
the perceived direction of motion (one-side vertical vs. isolated vertical: $F(1,4)=18.41, p<0.0127$; one-side horizontal vs. isolated horizontal: $F(1,4) = 18.52, p < 0.026$). However, the bias is significantly weaker than when two sides (Experiment 2) are occluded (two-sides vertical vs. one-side vertical: $F(1,4) = 19.19, p < 0.0220$; two-sides horizontal vs. one-side horizontal: $F(1,4) = 23.02, p < 0.0172$) (See Figure 3-7). A balance in horizontal and vertical motion was found for an aspect ratio of 1.7:1 and 1:1.8 for vertical and horizontal occluders. These values are larger than the predicted aspect ratios of 1.40:1 (vertical) and 1:1.38 (horizontal). Additionally, the corresponding reduction in terminator influence along the horizontal and vertical occluded sides was 84% and 91% respectively, considerably larger than the values of 49% and 52% calculated in Experiment 2. These values are much higher than would be expected given our previous results. The results suggest that a linear summation of terminator influence based on their classification is not sufficient to account for the perceived biases. Rather, the configuration of the stimulus as a whole influences the perceived direction of motion.

3.1.6 Experiment 6: Role of T-junctions

From the previous experiments, it is clear that the presence of monocular depth cues that support occlusion has a major influence on motion disambiguation. However, whether a single monocular cue is responsible, or the effect can be generated by an array of monocular cues, has yet to be established. Previous research has demonstrated that T-junctions are important cues for occlusion (Cavanagh, 1987; Anderson & Julesz, 1995). Other research has shown that the perception of occlusion depends on the presence of T-
Figure 3-7. EXPERIMENT 5: The occlusion of one side. The average percent horizontal motion reported by subjects is plotted against barber pole aspect ratio with solid lines for the isolated (●), vertical (▲), and horizontal (■) configurations. The legend below contains icons depicting the corresponding aspect ratios employed in the experiment. Although the icons only depict vertical occluders to the left of the barber pole, trials were run with both left and right vertical occluders and the results folded together. The same is true for the top and bottom horizontal occluders. Dotted lines are sigmoid estimations for the response curves.
or L-junctions even when other occlusion cues are present (Rubin, 1997). T-junctions have frequently been used in computational vision (Guzman, 1969; Grossberg, 1997). The top of the “T” is created by the occluding surface, the stem of the “T” is the occluded contour. The occluding surfaces in all of the previous experiments have created T-junctions between the barber pole pattern and the occluding surface (Figure 3-8). Experiment 6 was designed to determine if these T-junctions are critical for the introducing the bias.

Methods

New stimuli were created for Experiment 6 which differed from Experiment 1 in that the occluding bars were given the same length of the barber pole patch. This eliminated T-junctions whose tops are aligned with the occluding patches (See Legend, Figure 3-9).

Results and discussion

Although T-junctions whose tops are aligned with the occluding surface were eliminated, a bias was still observed (vertical occluders vs. vertical isolated: $F(1,5) = 6.68$, $p < 0.0472$; horizontal occluders vs. horizontal isolated: $F(1,5) = 16.58$, $p < 0.0096$). The bias was significantly weaker than found in Experiment 2, where T-junctions were present,
(no-T vs. T vertical: $F(1,5) = 29.47, p < 0.0123$; no-T vs. T horizontal: $F(1,5) = 57.85, p < 0.0047$). A balance in horizontal and vertical motion was found at an aspect ratio of 1.5:1 and 1:1.6 for vertical and horizontal occluders respectively. This corresponds to the vertical and horizontal terminators being attenuated by 32% and 38% respectively (See Figure 3-9).

McDermott, Weiss, & Adelson (1997) recently examined the effect of terminator classification on a different set of motion stimuli than our own. When overlapping vertical and horizontal bars oscillate sinusoidally, 90° out of phase, they can be seen moving coherently in a circular path, or separately as two bars sliding over each other. They demonstrated that adding a frame around the edges (create extrinsic terminators at edges of the two bars) improved coherence, in agreement with our results. They propose that a simple contour-based heuristic is used to segment motion signals. Our results, however, suggest that although T-junctions are a very important cue, they are not critical for the influence of terminators to be attenuated. A significant motion bias is still observed despite the elimination of T-junctions whose tops are aligned with the occluding surface.

Although local T-junctions between the individual grating bars and the occluding patch are still present in these stimuli, experimental observations suggest their influence is minimal. Shimojo et al. (1989) found that such local T-junctions are overridden by binocular information that suggests the barber pole patch is in the same plane as its surround. It should also be noted that Experiment 3B contained local T-junctions along the non-occluded side of the barber pole (between the grey background and the barber pole). The perception of motion in these stimuli was still one of motion under the occluding patches.
This suggests that local T-junctions are not an important factor. Pilot experiments also indicate that the bias in Experiment 6 persists when an intermediate grey-background is used. Such a background creates local T-junctions on both the occluded and non-occluded
sides, yet the bias is still for motion to be seen under the occluded side. Together, these observations suggest that it is unlikely that local T-junctions account for the remaining bias found in Experiment 6.

3.1.7 Experiment 7: Illusory Occlusion

The previous experiments demonstrated that the addition of occluding patches to a barber pole pattern alters the perception of motion. Disparity information and T-junctions were not necessary for the occluding patterns to influence motion disambiguation, nor was a specific contrast polarity between the occluders, background and barber poles. This suggests that boundary classification occurs prior to the resolution of motion ambiguity.

As we have previously suggested, these effects might be due to feedback from a higher level or parallel processes which interprets surface segmentation. If this is the case, the presence of a physical luminance change at the edge of a barber pole is not necessary to alter the perceived direction of motion, only the perception of an occluding edge is required. This idea is supported by Wallach (although no objective measurements were made). He noted that one of the perceptions of a moving grating with free-line ends visible is of an illusory surface appearing like a bright stripe under which the grating moves (Wuerger, Shapley & Rubin, 1996). Vallortigara & Bressan (1991) also report that terminators internal to a barber pole or plaid pattern are discounted when they aligned to form subjective occluding bars. Similar results are reported by Gurnsey & von Grünau (1997).

To further examine the role of illusory contours on motion disambiguation and to
quantify the phenomenon, we generated occluding surfaces using Kanizsa-like figures. Both the occluded and non-occluded edges of the barber pole had the same luminance profile. However, the occluded edge abutted an illusory contour created by a Kanizsa-like figure (Figure 3-10). As the previous experiments indicate that it is the perception of an occluding surface rather than a disparity or luminance change that results in a bias of motion disambiguation, we predicted that the Kaniza configuration would also result in a bias in the perception of barber pole patterns.

Methods

Six aspect ratios were employed for Experiment 7: 1:.40, 1:1.32, 1:1.16, 1:1.08, 1:1, 1.08:1, 1.16:1, 1.24:1, 1.32:1, 1.40:1. This smaller range was employed for two reasons. First, we expected the effect of illusory occluders to be significantly weaker than real for occluders. In informal observations, the authors found an illusory edge to be less salient than a real edge for the experimental configurations. Secondly, the aspect ratio was limited to prevent the illusory figure inducers from getting too far apart. As inducers of a fixed size are moved further apart, the perception of an illusory contour weakens (Watanabe & Oyama, 1988; Shipley & Kellman, 1992).

Results and discussion

A significant bias was found for the illusory occluders configuration (illusory occluders vs. no occluders, vertical: F(1,6) = 7.75, p < 0.0318; illusory occluders vs. no occluders, horizontal: F(1,6) = 0.0138, p < 0.0138). A balance in horizontal and vertical motion was found for ratios of 1.3:1 and 1.3:1 were found for the horizontal and vertical occluders.
respectively. The attenuation of vertical and horizontal occluders were both 23%. These results are strongly in support of input from other cortical areas which influences the motion processing areas (See Figure 3-10).

Figure 3-10.EXPERIMENT 7: The addition of occluding bars. The average percent horizontal motion reported by subjects is plotted against barber pole aspect ratio with solid lines for the isolated (●), vertical (▲), and horizontal (■) configurations. Dotted lines are sigmoid estimations for the response curves. A smaller range of aspect ratios was employed than the previous experiments. Four stimuli typical of those used in Experiment 7 are shown to the right of the plot. The occluders have the same luminance as the background.
3.2 General Discussion

There were three primary purposes to our study. The first was to compare the relative importance of disparity and monocular occlusion cues for motion perception. Are monocular occlusion cues significantly weaker than disparity cues? Secondly, we wished to determine the degree to which the influence of extrinsic terminators is discounted. Is it completely abolished? Is the extrinsic/intrinsic classification all-or-none? Finally, we wanted to determine the relative importance of different monocular depth cues. Do occluder and background contrast make a difference? Are T-junctions necessary?

The first purpose of our experiments, the comparison between disparity and monocular occlusion cues, provides insight into possible neural mechanisms involved the barber pole phenomenon. Recently, a physiological study demonstrated that stereo-based terminator classification affects the responses of cells in area MT (Stoner, Duncan & Albright, 1997). When the depth relationships of a barber pole configuration are such that they bias the perceived direction of motion in a cell’s preferred direction, the cell elicits its greatest response. There are two possible interpretations for this result. In one interpretation, cells coding terminator motion are suppressed when their terminators have crossed disparities and are seen as in-front. Such an interpretation does not require an analysis of occlusion, but merely the presence of the appropriate disparity cues. For the other interpretation, some manner of form processing is required. It is generally believed that there are separate visual pathways for form and motion (Maunsell & Newsome, 1987). The second interpretation suggests a second system, one that is involved in the analysis of occlusion, feeds into the motion processing areas and influences motion processing. These two pos-
sibilities were originally suggested by Shimojo et al. (1989) based on their psychophysical work.

Previous research has demonstrated cases in which disparity information overrides monocular depth cues (Shimojo et al. 1989). Such results are compatible with the first interpretation. Our results, on the other hand, support the latter interpretation, that there is input from a second system, which processes form. Our first major finding is that there is little difference in the influence of occluding patches with and without disparity differences (Experiments 1 and 2). In Experiment 2, monocular occlusion cues (indicating occlusion) override binocular depth cues (indicating no disparity difference). This suggests that the suppression of extrinsic terminators by disparity tuned motion sensitive cells is not a sufficient explanation of how terminators are suppressed.

Although Stoner, Duncan & Albright (1997) found that depth relationships in a barber pole configuration influence the response of MT cells, their result does not necessarily imply an input from the form to the motion pathway, as MT cells are tuned for disparity (Maunsell & Van Essen, 1983b). However, one could test for the influence of form, if the same experiments were performed using the range of stimuli which we have shown to influence motion disambiguation psychophysically.

Our second major finding is that the influence of terminators is not abolished but merely attenuated when terminators are classified as extrinsic. Even the strongest extrinsically classified terminators still retain some of their capacity to bias motion perception. This is particularly noticeable for the larger aspect ratios (See Figure 3-2). Additionally, the classification appears not to be a strict dichotomy but rather a smooth continuum in
which the influence of terminators varies with the strength of occlusion cues (See Figure 3-11 & Table 1). Terminators with monocular cues that strongly support extrinsic classification exert less of an influence on motion perception than those terminators with weaker support. The perceptual implication of this classification on a given trial is probabilistic. The perceived direction of motion for subjects (with a rare exception) is along a cardinal direction. However, the likelihood of a particular cardinal direction on a given trial depends on the degree to which various terminators can be thought of as extrinsic.

Our third major finding is that no single monocular cue is responsible for determining the degree to which terminators influence motion disambiguation. In addition to disparity, a multitude of cues exist (texture, T-junctions, etc.) which support the perception of occlusion and attenuate the influence of terminators to varying degrees (See Figure 3-11 & Table 1).

1. Note that only the illusory occlusion condition deviates from this smooth continuum. We suggest a possible reason for this is that as the aspect ratio of the barber pole changes, the distance between the barber pole and the illusory contour inducers also changes. In turn the strength of the illusory contour is altered for the different aspect ratios, giving different strengths to the perceived illusory surface.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Description of Stimuli</th>
<th>Horizontal Aspect Ratio</th>
<th>Horizontal Reduction</th>
<th>Vertical Aspect Ratio</th>
<th>Vertical Reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Textured Occluders with Depth</td>
<td>1:2.5</td>
<td>60%</td>
<td>2.6:1</td>
<td>61%</td>
</tr>
<tr>
<td>2</td>
<td>Textured Occluders without Depth</td>
<td>1:2.2</td>
<td>55%</td>
<td>2.3:1</td>
<td>57%</td>
</tr>
<tr>
<td>4</td>
<td>Thin Textured Occluders</td>
<td>1:2.2</td>
<td>55%</td>
<td>2.1:1</td>
<td>52%</td>
</tr>
<tr>
<td>3a</td>
<td>Grey Occluders / White Background</td>
<td>1:2.1</td>
<td>52%</td>
<td>2.0:1</td>
<td>49%</td>
</tr>
<tr>
<td>3b</td>
<td>White Occluders / Grey Background</td>
<td>1:1.7</td>
<td>41%</td>
<td>2.0:1</td>
<td>48%</td>
</tr>
<tr>
<td>5</td>
<td>Textured Occluders on One Side</td>
<td>1:1.8</td>
<td>91%</td>
<td>1.7:1</td>
<td>84%</td>
</tr>
<tr>
<td>6</td>
<td>Textured Occluders, no T-Junctions</td>
<td>1:1.6</td>
<td>38%</td>
<td>1.5:1</td>
<td>32%</td>
</tr>
<tr>
<td>7</td>
<td>Illusory Occluding Surface</td>
<td>1:1.3</td>
<td>23%</td>
<td>1.3:1</td>
<td>23%</td>
</tr>
</tbody>
</table>
Table 1). Although past research has emphasized the role of T-junctions, we found that although they are an important cue, they are not strictly necessary.

Many researchers have suggested the existence of separate visual pathways for form and motion (Maunsell & Newsome, 1987), for ‘spatial’ and ‘object’ vision (Mishkin, Ungerleider, & Macko, 1988) or the existence of separate color-opponent and broad-band channels (Schiller & Logothetis, 1990). Livingston & Hubel (1988) argued that motion and shape processing is segregated from the earliest levels in the brain to the highest levels that have been studied. However, the existence of an interaction between form and motion is supported by physiology showing the anatomical pathways in the primate visual system contain cross-connections between the streams (DeYoe & Van Essen, 1988).

Our psychophysical results support such an interaction between streams, demonstrating that form perception has an early and significant influence on the motion pathway, specifically in the disambiguation of motion information. Evidence suggests that cells in area V2 possess the ability to processes emergent visual forms such as the type of illusory contours in a Kanizsa display (von der Heydt, Peterhans, & Baumgartner, 1984). That Kaniza-like illusory contours can influence motion disambiguation (as shown in Experiment 7), suggests a role for the V2 -> MT connection. Recently, computational models have begun to incorporate such mechanisms (Grossberg, 1991; Francis & Grossberg, 1996).

Other recent psychophysical work also supports a strong interaction between the form and motion systems. In a related study we demonstrated that similarity (or grouping) based on form cues between spatially disparate areas mediates the disambiguation of
motion signals in a barber pole pattern (Lidén, Mingolla & Watanabe, 1997). In particular, contrast similarity rather than contrast magnitude determines the degree to which motion in one spatial area affects another. Croner & Albright (1997) found that form cues such as hue and texture can aid in the discrimination of motion direction, suggesting a strong influence of form cues on motion processing. Watanabe (in press) has demonstrated a reciprocal interaction between form and motion processing, with surface decom-

Figure 3-11. All conditions with horizontal occluders. The best fitting sigmoid estimation of the average responses of all subjects for each of the horizontal conditions is plotted. Note that for illustration purposes, the sigmoid plots extend from an aspect ration of 1.5:1 to 2.5:1, a bit beyond the actual measured values. The illusory occlusion case was measured using aspect ratios from 1.4:1 to 1:1.4 and all others were measured from an aspect ratio of 1:1 to and 2.4:1.
position affecting motion decomposition and vice versa.

Research also supports the idea that occlusion-related processing has an early effect on motion disambiguation. In the realm of apparent motion, Shimojo & Nakayama (1990) found that the addition of an intervening occluding surface in bi-stable apparent motion displays creates a bias for motion in the direction under the occluding surface. He & Nakayama (1994) found that an occluded ‘L’ shape has a bias to move between tokens of the same orientation only when the ‘L’ is seen in front of the occluder. The bias is reduced when the ‘L’ is seen behind. Presumably the ‘L’ is amodally completed behind the occluded patch. Together this work suggests a significant and essential role for information flow between the form and motion streams.

In summary, although previous research has suggested monocular occlusion cues are of less importance than binocular disparity for motion disambiguation in the barber pole phenomenon, we found monocular occlusion cues can profoundly influence motion perception even when zero-disparity is employed. Secondly, our results support the idea that extrinsic/intrinsic classification is not an all-or-none process; rather, a continuum exists in which one terminator can be classified as having stronger extrinsic properties than another. Thirdly, we demonstrated that a variety of occlusion cues, both monocular and binocular, influence the disambiguation of motion signals. No single cue is responsible. Finally our results support the role of input from the form system to MT for mediating the extrinsic/intrinsic classification rather than a disparity based suppression of cells in MT.
Chapter 4

A Model of Motion Segmentation and Integration

4.1 Modeling Introduction

The remaining chapters present a computational model which addresses both the integration and segmentation of motion signals, as well as some of the temporal properties of motion processing. The premise for the modeling work is that any system which distinguishes the motion of objects from that of whole fields must employ two separate but interacting computational processes, one for motion integration and one for motion segmentation. As explained in Chapter 1, the motion system, must both integrate and segment motion signals, to overcome noise and the aperture problem and to detect the presence of objects with different trajectories, respectively.

The computational requirements for these two processes are quite different. The integration process must combine local motion signals, computing some sort of average over space. This process tends to uniformize motion signals, eliminating local differences and smoothing over discontinuities. The segmentation process, on the other hand, must enhance local differences and identify motion discontinuities. It is not immediately apparent how to reconcile the conflicting requirements of these two processes, let alone explain how the two processes interact to create a unified interpretation of visual motion.

The model presented in this chapter demonstrates not only how these two processes can co-occur but also how they can complement one another through mutual interactions.
The model consists of two sets of nodes. These nodes are theorized to correspond biologically to cells or groups of cells. One set acts to integrate local motion signals, eliminating local differences and propagating motion signals across space. The second set of nodes localizes motion discontinuities and creates motion borders across which the propagation of motion signals is not allowed. Together the two set of cells create a unified interpretation of motion.

The previous chapters demonstrated that terminator classification and T-junctions are important for motion perception. The model discussed in the remaining chapters illustrates how terminator classification and T-junctions play an important role in the computation of motion processing.

Although many motion processing models have been developed, how to achieve both integration and segmentation within the motion domain has yet to be completely resolved. As a result, each model overlooks some important data. Section 4.2 describes how various models fail to achieve both an adequate integration and segmentation of motion signals. Many models also ignore the temporal aspects of motion integration, such as the improvement of performance over time in motion detection tasks and the apparent propagation of motion signals across space.

As an immense amount of research has been performed in the area of motion perception, the scope of the model is limited to the study of the factors that govern integration and segmentation of motion signals. The model is not concerned with how the initial motion signals are extracted. Accordingly, a simplified “front-end” for motion detection is employed in simulations. For convenience, this front-end will be referred to as the ‘V1’
component of the model. However, as this is a simplified version of what must actually be
taking place in biology, no specification is made as to the actual source of these signals.
They are assumed to originate from motion sensitive cells in either the primary or second-
ary visual areas. Also, although form cues play a role in determining the contribution of
individual motion signals and how they are allowed to propagate across space, the model
only addresses the effects of some hypothesized form inputs, not how the form cues are
generated. The model is also restricted to motion processing within the two-dimensional
domain.

4.2 Review of other Models

4.2.1 Segmentation or Integration

Models that perform motion integration often fail to achieve adequate image segmen-
tation. One method for overcoming noise and the aperture problem is to integrate motion
signals over a spatial neighborhood of a fixed distance (Hildreth and Koch, 1987; Bulthoff,
Little & Poggio, 1989a & 1989b). When motions from different objects fall within the
same spatial neighborhood, however, these models integrate motions which should not be
associated. Another method is that of regularization (Horn & Schunck, 1981; Yuille &
methods restrict the number of solutions to motion disambiguation by choosing a cost
function to be minimized, such as a motion discontinuity penalty. The cost function is cho-
sen by an analogy to a conservative physical system. Unfortunately, such methods produce
incorrect results for multi-object images or images with simple occlusions. Regularization methods fail to localize discontinuities as they smooth over them, only allowing the presence of one motion direction at discontinuities (Poggio, Torre & Koch. 1985).

Some modified regularization models come closer to addressing both segmentation and integration. Hutchinson, Kock, Luo and Mead (1988) presented a hardware implementation of regularization method which specifically allows for the smoothness assumption to be ‘turned-off’ between adjacent pixels when there is evidence for the presence of a motion discontinuity. This is accomplished by adding terms to the energy function which incorporate information about luminance edges from zero-crossings and motion discontinuities by the presence of large motion gradients. Harris, Kock, Staats & Lou (1990) followed up with an analogue version of the model with similar properties.

However, even if a modified regularization process can be implemented that does account for discontinuities, such a process may have little to do with human motion perception. Braddick (1993) states “to argue that vision solves a minimization problem does not tell us how this process could be implemented in the brain.”

There are a number of biologically inspired models of motion perception that attempt to determine how the brain might process motion. Nakayama & Loomis (1974) propose the idea of a convexity function which compares motion direction over its center and surround in different directions. Other models, such as Sachtlter & Zaidi (1995) use similar center-surround receptive fields to detect motion discontinuities. Although useful tools, without the addition of other mechanisms, these center-surround models are limited in scope and fail to deal with issues that are addressed by smoothing algorithms (such a noise
in the optic array) and difficulties introduced in motion measurements by the aperture problem.

Another way to view the problem of resolving motion ambiguities and the detection of motion boundaries involves the propagation of constraints on motion signals. Marr (1982) notes that although the aperture problem results in ambiguous motion signals, it also provides a range of motions which are inconsistent with the local motion. By combining these so-called “forbidden” directions of motion with forbidden directions in neighboring locations, one can find a smaller subset of motions which are compatible with the two areas. If combination with a neighboring region results in all motions being forbidden, it is likely that the initial measurements of the two regions belong to different objects. Although such a method might be effective for performing motion segmentation in a limited domain, like center-surround models, it fails to address motion integration.

One model which achieves a crude segmentation is that of Nowlan & Sejnowski (1994). Their filter selection model segments an image into multiple regions of support for each motion in the image. However, it does so by ignoring bad estimates of motion. It assumes that a small number of motion processes are present and ignores velocity measurement that don’t sufficiently agree with them. Thus, the support regions for each object are a small fraction of the local-velocity measurements. The final output of the model has no spatial resolution, as it consists of one set of units which represents the entire visual field. Some additional process would be required for true object segmentation. The model also does not deal with the temporal aspects of motion perception.

Chey, Grossberg & Mingolla (1997, In Press) present a model that extracts initial
motion measurements, accounts for much of the psychophysical data on speed tuning, and data on the coherence and incoherence of plaid patterns. However, the current version of the model does not adequately localize motion discontinuities for all stimuli. For example, for overlapping regions with different trajectories of motion, such as components of an incoherent plaid, the model produces a continuum of motion directions at their intersection rather than a sharp motion discontinuity. Phenomenologically, the motion of two crossing objects (such as bars) produces a clear segmentation into two motion directions with no such continuum at their intersection. Furthermore, psychophysical experiments using random dot stimuli show that, despite the summation of motion signals over large spatial and temporal ranges, the perception of motion edges is nonetheless sharp (Regan & Beverley 1984).

Some of the difficulties in the Chey et al. model may be rectified through the addition of form-to-motion interactions which act to segment the image (Francis & Grossberg, 1996; Baloch & Grossberg 1997) or through the use of 3D motion mechanisms. These segmentation processes operate within a separate form system rather than within the motion system itself. Form-to-motion interactions are important. However, only segmentation and integration processes that operate within the motion system are considered for the model examined in this chapter. There is ample evidence that a segmentation processes only within the motion system, serving to identify motion boundaries and block the integration of motion signals. Many cells within the motion processing areas of the brain are selective for motion discontinuities, showing no response to other types of discontinuities (Marcar and Cowey, 1992). Such evidence is discussed in detail in Section 4.3.3.
4.2.2 Temporal Aspects of Motion Integration and Propagation

A second aspect of most motion models is that they fail to deal with the temporal aspects of motion integration and propagation (but see Chey et al. 1997). Psychophysical evidence suggests that the integration of motion signals takes time (hundreds of milliseconds) and can have effects across the entire visual field, and also that such propagation can occur around 30 deg/sec. (Williams & Sekuler, 1984; Ramachandran & Anstis, 1986b; Watanabe & Cole 1994). Williams and Sekuler (1984) examine the perception of coherent global motion in dynamic random-dot cinematograms. They conclude that individual motion signals are independently detected and then pooled over both space and time to generate the final coherent motion percept. Similar findings are reported by Van Doorn & Koenderink (1984).

Snowden and Braddick (1989) use the term *temporal recruitment* to describe the improvement in direction discrimination performance in random-dot cinematograms with increasing number of frames. They conclude that the parameters of temporal recruitment are not easily explained by any combination of limits on integration in time or space. Instead they suggest “a cooperative process of inhibitory and excitatory interconnections between motion detectors...within a local area.”

Watanabe and Cole (1995) examine the propagation of local motion signals in apparent motion displays. The edges of evenly spaced dot fields are always seen to oscillate, while the interiors can be seen to move rightward or oscillate depending on presentation duration. Their results indicate that signals for oscillatory motion gradually propagate to central units with ambiguous motion signals. With short durations the unambiguous infor-
mation from the edge of the display does not have time to propagate to the ambiguous interior. Watanabe and Cole conclude that the “result of the experiment of this study suggest that the information exchange between distant neurons within MT may be accomplished by a gradual propagation by means of iterative local interactions between neurons rather than direction connected, long-range interactions.” The dissertation work takes a similar point of view, that the phenomenological experience of motion propagation is accomplished by a literal propagation of motion signals. Information exchange between distant neurons is not solely the product of large spatial kernels.

Another phenomenon which indicates the importance of the temporal aspects of motion integration is the ability of humans to detect a single dot moving on a trajectory embedded in noise. For short durations, such a dot is imperceptible; its motion integrated with surrounding noise (Watamaniuk, McKee & Grzywacz 1995). However, detection improves with increasing stimulus duration. Watamaniuk et al. suggest a network of interconnected motion-energy units could be performing a directionally selective temporal and spatial recruitment via facilitory signals in the direction of motion.

The importance of temporal integration is lost on the regularization models discussed in the previous section. Such methods fail to explain how humans can track the trajectory of a single dot among a field of moving random dots (Watamaniuk, McKee & Grzywacz 1995). Regularization results in the obliteration of a single signal dot in noise.

Of the models discussed thus far, only the model of Chey et al. (1997) includes the potential to model some important aspects of the temporal integration of motion. The model incorporates neurons which act as leaky integrators, summing their inputs over
time, as well as lateral recurrent feedback. Together these mechanisms allow motion propagation to occur, notably from feature tracking locations that compute unambiguous motion directions to locations where only ambiguous computations occur due to the aperture problem.

### 4.3 Sketch of the Model

The present model consists of a neural network with two primary systems each composed of interconnected cells. The first system comprises a set of integration cells which act to integrate local motion signals. These cells possess a surround which is tuned in the same direction as the center. The cells respond best when motion in both their center and surround are in the preferred direction of the cell. This type of cell performs a smoothing or integration process which is used to overcome noise and the aperture problem by propagating motion signals across space.

The second system comprises a set of segmentation cells which act to segment motion signals. These cells possess a surround which facilitates cell activity when stimuli in the surround are moving opposite to the cell’s directional preference. The surround also inhibits cell activity when stimuli in the surround move in the cell’s preferred direction. This type of cell is tuned to signal the presence of motion discontinuities and will be used to constrain the interactions of the integration cells.

Recurrence is an important part of this model. Interactions between integration cells serve to enhance the activity of disambiguated motion signals and to propagate unambiguous motion signals across space. As the size of the kernels are relatively small with respect
to the size of the image, and as only relatively nearby nodes in the model are connected, in
order for a disambiguating signal to ‘travel’ to ambiguous areas and for distant objects to
have an effect upon one another, recurrent iterative actions are required. The use of small
kernels helps to ensure that nearby objects can be segmented.

Additionally, the interaction between segmentation and integration cells is necessarily
recurrent. The spread of signal from integration cells depends on the activity of segmenta-
tion cells as they constrain the propagation of motion signals between integration cells.
However, the state of segmentation cells depends critically on input from integration cells.
Neither of these processes could be achieved with a feed-forward network.

4.3.1 Simplified V1 Representation

A number of motion models have been proposed for the initial extraction of motion
signals from the retinal image. These models can be classified into two basic types,
energy models (Fennema & Thompson, 1979; Marr & Ullman, 1981; Adelson & Bergen,
1985; Grossberg & Rudd, 1989, 1992) and correlational models (Reichardt, 1961; Van
Santen & Sperling, 1984, 1985). Such models do not process the ambiguities resulting
from the initial extraction. However, each may be suitable as a front-end for models
which do compute how signals can be integrated and segmented during subsequent pro-
cessing stages. As currently implemented, the input to the model need only specify an
analog value corresponding to the evidence for motion in each direction.

As the specifics of the initial motion extraction mechanism are of little importance to
the current model, a simple correlation scheme was chosen for it computational simplicity.
Two successive image frames are used to compute correlation. For each position in space, a small window of pixels is chosen from the first frame. The grey-level pixel values for the window in the first frame are compared to the grey-level values for shifted windows in the second frame. The correlation between the two grey-levels is used as a measurement of the motion in the direction of the shift. For each position eight shift directions were employed. For each direction, shifts of one and two pixels were measured and the resulting correlation values summed. The use of two shift sizes was employed to capture a larger range of speeds. Details of these calculations can be found in Appendix B2.

4.3.2 Model of Integration

*Background*

Psychophysical experiments have revealed some of the mechanisms that may be involved in the integration of local motion signals. The study of random dot cinemograms suggests that there is a pooling of responses between cells that have similar directional tuning (Williams & Sekuler 1984). Similarly, Kim and Wilson (1993) found that coherency of plaid patterns depends on relative component motion direction and suggest a facilitation between units tuned to similar directions.

There also appears to be an antagonistic mechanism in place. Marshak and Sekuler (1979) report an exaggeration of the perceived angular difference between random dots patterns that move in different directions. Snowden (1989) examined the effect of adding vertical pattern motion on the perception of horizontally moving patterns. He found strong suppression from the vertical direction suggesting that output from local detectors
is passed into a system of interactions among all local signals, similar directions being facilitory and other directions being inhibitory.

Wilson & Cowan (1973) proposed a mathematical cooperative/competitive neural network model with inhibitory and excitatory processes. Williams, Phillips & Sekuler (1986) expanded the model, adding two primary computational principles, a non-linear excitatory interaction between nodes of similar directional tuning and non-linear inhibition among nodes with different directional tuning. These two computational principles have been the basis for many biologically motivated models of global motion processing (Nawrot & Sekuler, 1990; Grossberg & Mingolla, 1993; Nowlan & Sejnowski, 1994; Grunewald, 1996; Chey et al., 1997).

There are, however, many different ways in which such computations have been implemented. Models vary in the range of directions for which facilitory and inhibitory mechanisms are active and the spatial extent of such interactions. Some models, such as Williams et al. (1986) make the computations recurrently and contain separate channels for inhibitory and excitatory activity. Other models, such as that of Chey et al. (1997) perform cooperative and competitive steps in consecutive layers. The integration cells in the model examined in the dissertation are recurrent; all computations are made within the same layer. It uses one directionally dependent excitation process and two directionally dependent inhibition processes, one short-range and one long-range.

Integration Cells

The dynamics of the integration cells depends on three computational principles (See Figure 4-1):
[1] A directionally dependent inhibition exists between cells with different direction tuning at a given spatial location weighted by the directional difference. In this way, cells in ambiguous locations (areas containing multiple active directions) are preferentially suppressed with respect to cells in unambiguous locations (areas with one or few active directions). It has been suggested that such preferential weighting of unambiguous motion signals is used by the visual system to overcome the aperture problem (Shiffrar & Pavel, 1991; Lorenceau & Shiffrar, 1992). Chey et al. (1997) have demonstrated how such preferential weighting of unambiguous motions is one of the effects of inhibition across direction within spatial location.

[2] A directionally dependent excitation across space, weighted by the magnitude of directional difference, operates between nearby cells. In the model this excitation signal serves to propagate disambiguated motion signals across space. Such propagation can explain some of the psychophysical phenomenon discussed in Chapter 2 such as motion capture and random-dot cinematograms.

[3] A directionally dependent long range inhibition, weighted by the magnitude of directional difference, operates between cells of different directional tuning. This inhibition indirectly implements the constraint propagation proposed by Marr (1982), inhibiting spatially distant nodes whose signals are inconsistent with local nodes. When combination with neighboring regions result in all motions being suppressed, it is likely that the node is positioned at a motion discontinuity, and its response is suppressed. In this way the model can still achieve a rudimentary segmentation even when only the integration cells are included (Lidén 1997).
The importance of long range inhibition has been demonstrated for the early stages of motion processing in both the retina (Barlow & Levick, 1965) and the primary visual cortex (Sillito, 1975). Long range inhibition has been employed by many other motion models. The model of Qian, Andersen & Adelson (1994b) employs long-range directional inhibition between opposite directions of motion. Grunewald (1996) and Chey et al. (1997) also employ a version of directionally-dependent long-range inhibition. Details of computations for the integration cells can be found in Appendix B3.

Figure 4-1. Integration Cell Architecture: Each position in space is indicated by a vertical column. Note there the three processes: [1] inhibition within the same spatial location across directions, [2] excitation between nearby nodes with similar directional tuning and [3] far reaching inhibition to nodes with different directional tuning. Note that all connection are weighted by the directional difference.
**4.3.3 Model of Segmentation**

*Background*

An array of evidence supports the existence of a specialized system for processing motion boundaries. Lamme, Dijk, & Spekreijse (1993) recorded visually evoked potentials (VEP) in man and monkey for a random dot segmentation task and found response components specific to motion contours. Not only were the VEP components correlated to subject performance, but they were not present for non-motion contours or non-contour motion stimuli.

Sachtler and Zaidi (1994) examined motion detection in random-dot patterns with square wave, sine wave and cusp velocity profiles. Subjects were found to have greater sensitivity to square wave velocity profiles and showed no reduction in sensitivity to cusp stimuli. Sachtler and Zaidi conclude that sensitivity is largely determined by the presence of motion boundaries and that the steepness of the velocity gradient determines sensitivity.

Lesion studies indicate the middle temporal area has a role in the detection of kinetic boundaries. Marcar and Cowey (1992) examined monkeys trained to perform shape discrimination tasks. After lesions to MT, the monkeys were found to have unimpaired performance on shape discrimination tasks using luminance boundaries, but severe impairment with shape discrimination tasks using kinetic boundaries. They claim their experiments “support the view that area MT plays an indispensable role in the integration of kinetic boundaries into the form channel.”

Single-unit recording experiments also support the existence of cells which respond
to motion boundaries. Frost & Nakayama (1983) recorded from cells in the deeper layers of the pigeon optic tectum. They found cells which respond best when textured backgrounds move in the opposite direction of a moving test spot. The cells were not selective for the direction of the spot as long as the surround moved in the opposite direction of the test spot. Frost & Nakayama suggest that such cells could play a critical role in the delineation of object boundaries, signalling the movement of an object relative to its surround. Tanaka, Hikosaka, Saito, Yukie, Fukada & Iwai (1986) and Allman, Miezin & McGuinness (1985) report finding a similar cell in area MT in macaque monkey. The cells have no conventional directional preference for the center and surround, but rather respond when the center and surround are moving in opposite direction.

**Segmentation Cells**

The second system in the model is a segmentation system, consisting of cells that are center-surround and respond greatest to motion discontinuities. Such center-surround cells have been employed in models such as Nakayama & Loomis (1974) and Sachtler & Zaidi (1995). However, unlike these models, which generate center-surround receptive fields through a feed-forward weighting of inputs, the center-surround cells in the dissertation’s model arise from recurrent interaction within the network, as in Chey et al.’s (1997) model.

For each position in visual space there is a set of segmentation cells which are tuned for the full set of possible directions of motion (in this case 8 directions are used).

There are three sources of excitatory input for the segmentation cells:

1. **Center Surround Excitation:** Segmentation cells receive center-surround input
from the integration cells. They are excited by integration cells of the same preferred direction in their receptive field center and by integration cells that possess the opposite preferred direction in their receptive field surround. In this way, segmentation cells are excited by direction discontinuities in the pattern of integration cell activity.

[2] *Excitation from nearby segmentation cells:* Segmentation cells also receive a gating signal from other nearby excitation cells of the same directional tuning. A gating signal is one which gates input to the cell. When the gate is closed, other inputs to the cell are blocked. This gating allows for the development of motion borders signaling the presence of motion discontinuities. In this way segmentation cells reinforce each other’s activity and grow motion borders across the image when there is support for a motion discontinuity.

[3] “V1”: Segmentation cells also receive a non-direction specific gating signal from local motion cells. The gate is open whenever bottom up motion signals from V1 are present.

A key computational mechanism for the segmentation cells is that their input is gated. A segmentation cell can not be activated by center-surround input unless it also receives a gating signal from either V1 or other segmentation cells (See Figure 4-2). The propagation of activity in the integration cell system allows for sub-threshold integration cell activity *even when there is no underlying V1 activation.* The gating mechanism ensures that motion borders only develop in the presence of V1 activation or when a such a border is supported by a border in another location where V1 is active. If this gating were not present erroneous motion borders could develop based on segmentation cell
activity alone. Example simulations in the next chapter will help to clarify the importance and behavior of the gating signals.

There are two sources of inhibition for segmentation cells, an inherent decay and inhibition which occurs when integration cells in the segmentation cell’s surround that possess the same preferred direction as the segmentation cell’s center are active. This has the effect of suppressing the cell’s activity when presented with a large field of uniform motion. Note that unlike the excitatory input, neither of the inhibitory inputs are gated. Details of these calculations can be found in Appendix B4.

4.3.4 Segmentation & Integration System Interaction

One way in which the segmentation and integration cells interact has already been discussed. Namely, the activity of the segmentation cells is determined by integration cell activity through input channels with center-surround receptive fields. However, a key mechanism is still not in place. One of the primary differences between the current model and existing models is that the segmentation cells regulate the action of the integration
cells. The segmentation cells excite nearby segmentation cells, forming motion borders. These segmentation cells serve as barriers for the spread of motion signals by the integration cells in qualitatively similar way that cells in the Boundary Contour System (BCS) block the spread of form signals in the Feature Contour System (FCS) in the model of Grossberg & Todorović (1988).

Although such interactions are not strictly necessary for all motion stimuli, they become critical for any model when multiple moving objects are present in the visual array. This is particularly true when such objects overlap, as integration should only occur between related objects. The role of segmentation cells is to block the spread of activity of integration cells. By doing so the segmentation cells prevent the integration of motion signals between unrelated objects.

Segmentation cells prevent the spread of integration by suppressing integration cells that share nearby spatial locations. Although there is a set of segmentation cells representing the full array of preferred directions at each spatial position, the suppression is directionally non-specific, as any discontinuity (regardless of its direction) is relevant.

One phenomenon that supports the idea that the propagation of motion signals are blocked by a segmentation systems is that in some special cases the motion system can ‘break-down’. When segmentation fails, propagating motion information may ‘leak’ from objects which should contain them. Watanabe & Cavanagh (1991) reported an analogue to neon color spreading in which a static or moving texture cross was inserted at the intersection of an Ehrenstein figure. When presented in the periphery, they found that the texture and its motion appears to spread outside the cross. Such failures of the motion system
illustrate the importance of segmentation cues which serve to moderate the interaction of motion signals.

4.3.5 The Role of Form

Psychophysical data suggest that surface perception profoundly alters the integration of motion signals (Stoner & Albright, 1993). Local motion information near line edges usually provides unambiguous motion information that can be used to disambiguate other motion signals. However, as discussed in previous chapters, terminator information can be discounted based on surface segmentation (Shimojo, Silverman & Nakayama, 1989). Ramachandran and Anstis (1986a) demonstrated that occlusion can alter the perception of apparent motion displays consisting of sparse random dot patterns with an empty (dot-free) square patch. Whether subjects perceive the square as an occluding surface, which moves back and forth hiding dots beneath it, or as a patch of dots which jump back and forth, alternately filling the empty spaces, depends on form cues. Faubert & Grunau (1995) and Tse, Cavanagh & Nakayama (In Press) have demonstrated how the shape of apparent motion cues can radically change the perception of motion. Trueswell & Hayhoe (1993) have shown that surface segmentation cues are used to determine whether or not motion information arising from various contours are pooled together.

A number of models have started to examine how form cues can be integrated into the motion system (Poggio, Gamble & Little, 1988; Weiss & Adelson, 1994; Francis & Grossberg, 1996; Baloch & Grossberg, 1997). Although clearly, surface segmentation plays an important role in some of the examples mentioned above, one of the purposes of the cur-
rent study was to examine how much can be accomplished within the motion system itself to disambiguate motion signals with the least amount of external input.

At a minimum, it is critical that any motion processing system suppress motion signals from spatial locations where occlusion is present, as occlusion produces spurious motion signals. When any part of an object passes beneath another, all motion signals orthogonal to the orientation of the occluding contour are lost. Only motion signals parallel to the edge are preserved (Figure 4-3). Furthermore, such spurious motion signals are unambiguous as the aperture problem only applies when a single edge is present. If such signals were allowed to survive, they might lead to an incorrect disambiguation of object motion.

As a final element of the model, inputs from an unspecified form system suppress input to integration nodes near locations of occlusion. The mechanisms for the generation of such form signals have been addressed in other modeling work (Grossberg & Mingolla 1985b; Grossberg 1994, Grossberg 1997) and are beyond the scope of this model. T-junctions are employed as indications of occlusion. Rather than simulate the form processing involved in identifying T-junctions, the localization and identification of T-junctions is
performed manually. A mask is composed representing the location and position of T-junctions and then used to suppress input to I cells in the presence of T-junctions

**4.4 Neurophysiological Interpretation**

Although neurophysiological knowledge of cortical motion processing is still incomplete, one can begin to speculate about how such a model might be related to the way the brain computes motion signals. The brain area known as the middle temporal area (MT) is the most likely physiological correlate for the model.

The initial extraction of motion in primates signals is thought to take place in the primary visual cortex (V1). Layer 4b of the primary visual cortex contains a particularly large number of direction-selective neurons (Dow, 1974; Blasdel & Fitzpatrick 1984) and gives rise to the greatest number of projections to MT (Maunsell & Van Essen 1983c). MT also receives input from thick stripes in the secondary visual cortex (V2) where a large proportion of neurons are direction-selective (DeYoe & Van Essen, 1985; Shipp & Zeki 1985).

Rodman & Albright (1989) suggest that the middle temporal cortical area (MT) is the first cortical area in which the individual components of a complex stimulus are combined to form a representation of the global motion. At the same time, Qian & Anderson (1994) argue that MT is the first cortical region in which the suppression of local motion signals is employed to achieve a reduction in the noise of a motion stimulus. Area MT would therefore seem to be the first likely candidate for a neural area capable of simultaneously performing integration and segmentation of motion signals.
MT is comprised of motion selective cells (Zeki, 1974a&b) with a full topographical representation of the visual field (Gattass & Gross, 1981). Interactions in the model, such as inhibition between directions, are supported by psychophysical experiments (Marshak & Sekuler, 1979; Snowden, 1989; Grunewald & Lankheet, 1996) and are important for generating MT response properties (Snowden, Treue, Erickson & Andersen, 1991; Bradley, Qian et al., 1995). Anatomical studies of MT also report numerous fibers, oriented laterally with respect to cortical surface, that could support the interactions used in the model (Van Essen, Maunsell, Bixby 1981). Evidence also suggests that the receptive field structure of segmentation cells seen in MT is unlikely to arise from the convergence of afferent connection since neurons in the input layer of MT are the least likely to have surrounds in their receptive fields and then only possess the smallest and weakest surrounds. Rather it is likely to arise through inhibitory synapses within MT, as in the proposed model (Raiguel, Van Hulle, Xiao, Marcar & Orban 1995).

One possible interpretation is that the proposed segmentation and integration node types of the model are analogous with two cell types found in MT (Allman, Miezin & McGuinness, 1985; Tanaka, Hikosaka, Saito, Yukie, Fukada & Iwai, 1986; Born & Tootell, 1992; Saito, 1993). Tanaka et al. (1986) reported finding two classes of directionally selective cells in MT. The classes differ in their surround inhibition ($SI$). Many cells, designated $SI^+$, were found to respond to a bar moving in a specific direction and were suppressed by a dot pattern moving in the same direction in their surround. The response was sometimes facilitated by movement in the opposite direction. They suggest that such cells could be detecting differences between movements of an object and its background. The
other class of cells, designated SI-, have no surround suppression. Although different nomenclatures are employed, similar results were reported by Allman et al. (1985) and Saito (1993). Born and Tootell (1992) labeled MT with 2-deoxyglucose in the presence of large-field random-dot patterns and found two neuron types, interband (contrast) and band (summation), that were segregated in a columnar fashion.

It is also known that opponent cells in MT are capable of signalling a kinetic boundary, as would be important for the segmentation nodes in the model (Marcar, Xiao, Raiguel, Maes & Orban, 1995). However, these MT cells are unable to accurately code the orientation of a kinetic boundary or the position of the boundary within their receptive fields. Marcar et al (1995) ask how such cells could encode sharp motion boundaries? The model, however, demonstrates that such coding of orientation and position within the receptive field of segmentation cells is unimportant. The suppression of integration nodes by the segmentation nodes in the model is directionally non-specific and spatially diffuse. Furthermore, it is supposed that the activation of the segmentation cells in the model is not consciously perceived. The perception of boundaries results from the activity of integration nodes, which are consciously perceived, rather than the diffuse activity of segmentation cells.

4.4.1 Relationship to MST

One of the major output areas for MT is another brain area known as MST (Maunsell & Van Essen, 1983c). MST appears to be further subdivided into at least two areas (Komatsu & Wurtz, 1988). MSTv(l) cells have small receptive fields and are presumably
used for smooth pursuit (Tanaka, Sugita, Moriya & Saito, 1993). MSTd cells have large, complex receptive fields and are presumably used for the analysis of self-motion (Wurtz, Yamasaki, Duffy & Roy, 1990; Duffy & Wurtz, 1991a&b) and complex object motion (Graziano, Andersen & Snowden, 1994; Geesaman & Anderson, 1996; Ferrera & Lisberger, 1997).

Recently, Born, Zhang & Berezovskii (1997) presented preliminary evidence that projections are made preferentially between MT band cells and MST large-field cells and between MT interband cells and MST small-field cells. These results are in accord with the model and suggest that MT segmentation cells provide MSTv with motion discontinuity information for smooth pursuit, while MT integration cells provide MSTd with disambiguated motion signals for the processing of optic-flow and object motion.

There are several reasons to believe that MSTd cells process both optic flow and object motion. Graziano, Andersen & Snowden (1994) suggest MSTd neurons encode patterns of motion regardless of whether these motions are generated by moving objects or by motion induced by the observer. MSTd neurons are position invariant, meaning they cannot give precise information about the retinal location of the focus of expansion. This means that individual neurons cannot accurately encode the direction of heading. Positional invariance suggests other possible roles for MSTd than optic flow as there is no computational advantage in positional invariance for calculating optical flow.

Graziano et al. suggest that there were probably some artifacts in the earlier studies by Saito & Tanaka, that claimed only very large field motion elicits cell response in MSTd, as these studies employed anesthetized animals. When anesthesia is employed, larger input is
needed to get cells to respond. In Graziano at al. (1994) cells were found to respond to much smaller stimuli than previously reported. A full response was found for stimuli of a size of 10° of visual angle. They argue that a patch of 10° does not resemble the flow field produced by an observer. They suggest instead that perhaps the function of MSTd is to analyze the motion of individual objects in the environment.

Geesaman & Andersen (1996) looked to see if MSTd is involved in the generic function of complex motion pattern processing rather than just the analysis of self motion. They found neurons respond to a diverse set of stimuli (dots, solid and outlined squares, flicker, non-fourier motion, etc.). MSTd was found to be form/cue invariant and to respond to stimuli as small at 5 degrees. They claim that it is an appropriate candidate for analysis of object motion as well as motion induced by observer translation.

Geesaman et al. (1996) suggest that object and self-motion information coexists in MSTd and that parsing of the motion signals is delegated to other cortical areas. Psychophysical experiments show that in certain conditions the motion of objects over optical flow patterns influences the perceived location of the focus of expansion (Royden & Hildreth, 1994 & 1996, Warren & Saunders 1995). Physiologically, Ferrera & Lisberger (1997) report MSTd cells which respond to any motion inside of the receptive field and do not distinguish between target and background motion. The targets employed in these experiments were only 0.9 degrees in size.

The lateral suprasylvian visual area (LS) of the cat is the analogous area to MT/MST. The response properties of LS led researchers to suggest that it plays a major role in heading perception. However, on closer examination, with more complicated stimuli combin-
ing both global motion and object motion, it was revealed that LS does not respond solely to global motion but can respond specifically to particular objects embedded in an optic flow display (Sherk, Mulligan & Kim 1997). Research with these more complicated stimuli have yet to be performed in MSTd in monkey.

Even if MSTd were only processing pure optic flow, the addition of an earlier process handling motion disambiguation would be beneficial. Imagine a texture composed of randomly placed diagonally oriented lines translating to the right. Due to the aperture problem motion along the edge of the lines is ambiguous. Apart from the terminators (which comprise a small portion of the motion signals) the majority of energy in the motion signal is diagonal. Thus the input to MSTd would be composed of mainly diagonal motion. On the other had, if the motion signals were disambiguated before reaching MSTd (presumably in MT), the input to MSTd would be pure rightward motion - the true direction of optic flow.

Another possible interpretation of the present model is that its two computational processes take place within only the band or interband column. In neurophysiology, the two cell types are not all-or-none, but rather a continuum of surrounds exists, from very weak to very strong (Born & Tootell, 1992). When examined closely, few cells show a total lack of inhibition in their surrounds (although the surround of some may be quite weak). And even within an interband (opponent) column, there are many cells which fit the classification of summation cells (Raiguel, Van Hulle, Xiao, Marcar & Orban 1995). Further physiological research will need to be conducted to determine which interpretation of the model is the most accurate.
4.5 Summary

The model presented in this chapter consists of two computational processes implemented by two separate, but interacting sets of nodes. It explains how two seemingly contradictory processes, integration and segmentation, can interact to create a unified interpretation of visual motion. Integration nodes act to smooth over noise in the input image, overcome the aperture problem and propagate motion signals across visual space. The dynamics of the integration nodes depend on three computational processes, a directionally dependent inhibition within spatial location, a directionally dependent excitation across space and a directionally dependent long range inhibition.

Segmentation nodes act to detect motion discontinuities through the use of motion center surround receptive fields. These cells are only activated if, in addition to their center surround input, they also receive input from local motion detectors or lateral input from other segmentation cells. This gating process acts to grow motion borders along discontinuity edges. The segmentation cells in turn inhibit the integration cells preventing the propagation of motion signals across motion discontinuities. Chapter 5 demonstrates how the model behaves on a number of classic motion stimuli.
Chapter 5

Simulations

5.1 Translating Line

As previously discussed, due to the aperture problem, the direction of motion of a translating line is ambiguous along the length of the line. Measurements about the true direction of motion can only be obtained at the line’s terminators. Figure 5-1 depicts the output of V1 cells for a horizontal line translating at 45 degrees relative to its orientation using the ‘star’ representation. Motion of the line is ambiguous along the edge of the line as cells with upward, up-to-the-left and up-to-the-right directional preferences are equally active. Along the terminators, however, unambiguous information about the true direction of motion is available in some locations, as only cells coding motion up-to-the-right are active.

One proposal for a method to recover the true direction of line motion is for unambiguous motions signals generated by the terminators propagate along the contour of the line where they act to disambiguate motion (Wallach, 1976; Hildreth, 1983; Nakayama & Silverman, 1988a,b; Lorenceau, Shiffrar, Wells, & Castet 1993). One model that has successfully used this method is that of Chey, Grossberg & Mingolla (1997). The method is also a feature which results from the dynamics in the present model.

An examination of the evolving activity of the integration cells (MT₁) demonstrates how the unambiguous motion information available from V1 nodes centered over the
line’s terminators is able to disambiguate the motion of the entire line. Initially, the activity of MT_i nodes mirrors that of the V1 nodes (Figure 5-2, A). However, as time evolves the activity of nodes at the terminators is enhanced with respect to nodes along the rest of the line (Figure 5-2, B). This occurs due to the directional competition of MT_i nodes within spatial location. The long-range inhibition and short-range excitation interactions of the MT_i nodes serve to gradually propagate the unambiguous activation along the length of the entire line. Eventually the activity of all nodes along the line are disambiguated and the entire line is coded as moving up and to the right (Figure 5-2, C).

Figures 5-1 and 5-2 use a ‘star’ representation to depict node activity. Unfortunately
this representation is inadequate for larger, more complicated stimuli as it takes too much room to display. In addition since it only allows one to look at small portions of an image at a time, visualization of the evolution of the network’s activity is difficult. A color coding scheme will be used as an additional means for visualizing network activity. Rather
than a star, this method uses a color to represent each location in space. Hue represents direction; saturation is used to represent magnitude. Figure 5-3 is a key for all subsequent diagrams in which color is employed.

Figure 5-3. Key to Color Representation: Hue represents directions. Saturation represents magnitude.

The evolution of MT$_i$ activity for translating lines with different terminator orientations are depicted in Figure 5-4 using the color scheme. The color at each point in space is determined by calculating a vector average based on the activity and directional selectivity of all nodes located at that point in space. Note how unambiguous terminator information at the edges fills into the center. The local orientation of the terminators makes no difference.

The results of the simulation are in agreement with experiments that examine the per-
ception of translating line segments (Lorenceau, Shiffrar, Wells & Castet 1993). For short durations (100 to 400 msec) observers presented with a translating line perceive motion in a direction normal to the line orientation rather than in the veridical direction of motion (as signaled by the terminators). Only at longer durations is the line seen to move in the direction of its terminators. For low contrast, long duration presentations, observers reported perceiving motion normal to line orientation at the beginning of motion which would smoothly turn into motion in the direction of the terminators.

Although segmentation nodes (MT_s) were present in the discussed simulations, their activity levels are not shown as they are negligible for this particular stimulus. The motion of a translating line does not contain motion discontinuities and thus does not evoke significant activity from the MT_s nodes. Subsequent simulations will demonstrate situations in which the MT_s cells are active and play an important role in constraining the evolving
activity of the MT; nodes

5.2 Line Capture

As demonstrated in earlier chapters, the addition of an occluding surface to an image can have a major effect on the perceived motion within that image. For a translating line, if the terminators of the line are obscured, the true motion of the line is undefined. However, if one adds unambiguous motion to a different area of the image, the perceived motion of the ambiguous line can be captured by that unambiguous motion. This can occur even when the unambiguously moving area is not spatially connected to the ambiguous moving area.

Two examples of line capture are shown in Figure 5-5. The motion of the line on the left, “A”, is ambiguous as the white diamonds obscure motion of the terminators. However, when the same line is paired with a rightward moving line, “B”, the perceived direction of the ambiguous line is also rightward. On the other hand, if the ambiguously moving line is paired with an upwardly moving line, “C”, the ambiguous line appears to move upward. The only difference between the stimuli depicted in “B” and “C” is the motion of the terminators of the paired line; line edge motion is the same in both. Despite the small area of change, the perceived direction of motion of both lines as a whole is enormously different.

Model Account of Line Capture

Simulation of motion capture depends on two additional mechanisms that were
present but played a negligible part in the single line simulation. First, line capture depends on a suppression of motion information from occluded terminators. Secondly, line capture depends on sub-threshold activation and propagation of MTi node activity.

As seen in Chapter 3, one of the primary cues for occlusion is the presence of T-junctions produced by the intersection of two objects. Although clearly other cues, such as depth discontinuities and changes in texture, are involved, T-junctions are quite reliable as a first pass. Consequently, for the model, T-junctions serve to identify points of occlusion and in turn determine locations in which input from V1 to MTi is suppressed.

The detection of T-junctions is beyond the scope of the current study as it lies the realm of form processing. However, once identified and localized, T-junctions play an important part in the model. Rather than simulate the form processing involved in T-junctions, the localization and identification of T-junctions was performed manually. A mask was composed representing the location and position of T-junctions and then applied to V1 activity, thereby suppressing V1 input to MTi in the presence of T-junctions. One could easily imagine a process by which T-junctions are identified and localized (Guzman

Figure 5-5. Line Capture. (A) An ambiguous line. (B) Horizontal line capture. (C) Vertical line capture.

A more complete model would incorporate such T-junction processing.

The second mechanism that plays an important role in the line-capture example is the propagation of MT_i activity across regions in which V1 input is absent. MT_i node activity in regions lacking direct V1 input will be referred to as sub-threshold as the allowed maximal activation level of these nodes is less than when V1 input is present. This sub-threshold activity allows disambiguating motion signals from one area of space to influence motion in another even when the motion signals are spatially disconnected. An examination of the evolving activity of the model will help to clarify how sub-threshold activation operates.

Figure 5-6 shows the evolving activity of MT_i nodes for the horizontal (left two columns) and vertical (right two-columns) line capture examples. Both sub-threshold and threshold activity is displayed. Initially, as seen in snapshot ‘A’, there is little difference in MT_i activity for horizontal and vertical line capture. The only distinction occurs for the terminators of the non-occluded line, one of which shows rightward motion (red) and the other upward motion (lime-green). As in the single line example, motion along center of the lines is ambiguous (orange). MT_i output for the two occluded lines is initially identical. Notice also that there is no motion information at the tips of the occluded lines, as V1 input near the occluded terminators is suppressed.

As network activity evolves, snapshots ‘B’ & ‘C, disambiguated terminator motion in the non-occluded bar gradually disambiguates the non-occluded bar. Sub-threshold motion signals from both bars also propagates beyond the bars. The occluded bar remains
Figure 5-6. Line Capture Model Output: Eight snapshots of the evolving sub-threshold and threshold MT$_i$ node activity for horizontal (left two columns) and vertical (right two columns) line capture.
unchanged as no disambiguation signal is present. However, by snapshot ‘D’, disambiguating sub-threshold activation from non-occluded bar begins to reach the occluded bar. This signal begins disambiguating the occluded bar. By the final snapshot, ‘H’, both bars have been entirely disambiguated. The occluded bars move in the same direction as their non-occluded neighbors.

5.3 The Barber Pole Illusion

The previous examples have demonstrated how unambiguous motion information from a line’s terminators can disambiguate motion along the rest of the line and how lines in different areas of space can influence each other’s motion. What happens when multiple disambiguating signals are present that are not in agreement?

The barber pole offers an example of a situation in which a competition between terminators exists. Terminators along the horizontal axis of a barber pole signal a different direction of motion than terminators along the vertical axis. As discussed in Chapter 3, the dominant perceived direction of motion is generally along the longer axis of the barber pole. Presumably this occurs as more terminators are present along the longer axis.

The model demonstrates that this is a viable explanation. Figure 5-7: ‘A’, displays MT_i cell activity for a barber pole with a longer horizontal than vertical axis. Terminators along the vertical edges signal unambiguous upward motion (lime-green). Terminators along the horizontal edges signal unambiguous rightward motion (red). Motion in the interior is ambiguous and orthogonal to the orientation of the lines (orange). As the network’s activity evolves, both the upward and rightward disambiguating signals are propa-
gated inwards. However, as there are more rightward moving terminators, the rightward signal is stronger and rightward motion comes to dominate the entire barber pole display (left column, bottom).

The inclusion of additional moving stimuli, such as dots, can add another source of disambiguating signal to a barber pole display. As previously discussed, such signals may result in the barber pole being perceived to move in the direction of dot motion (Shiffrar, Li & Lorenceau 1995). Figure 5-7: ‘B’, demonstrates how the addition of two upward moving dots can change the behavior of the model. As in the previous example there are still more rightward moving terminators than upward moving terminators. However, the addition of two upward moving dots creates a second source of disambiguated upward signal. This upward signal comes to dominate the entire barber pole.

Shiffrar et al. (1995) note that such capture is less likely to occur when the dots have a significantly different speed than the barber pole. As only a single range of speeds is used in the model, such capture also occurs if dot speed is outside this speed range. It is well known the cells in MT are tuned to speed (Maunsell & Van Essen, 1983a; Felleman & Kaas, 1984). An expanded version of the model with speed tuned MT cells could account for this data by having connections which were weighted by similarity in speed tuning. Cells coding different speed would have significantly less influence on each other.

Shiffrar et al. (1995) also report that dot influence decreases as depth differences between the dots and the barber pole increase. Binocular disparity is not within in the scope of the model as currently implemented. However, it is also known that MT cells are tuned to disparity (Maunsell & Van Essen, 1983b). Were the model expanded to include
depth tuning of MT₁ nodes and interactions weighted by depth plane, it could also account for these results.

Chapter 3 demonstrated how the addition of occluders to a barber pole pattern can change its perceived direction of motion. It was postulated that this change was a result of a suppression of motion information from the occluded terminators. Figure 5-7: ‘C’ demonstrates how such suppression can allow occluders to change the perceived direction of motion of a barber pole. A mask was composed based on the T-junctions along the occluded edges of the barber pole. As unambiguous terminator information is no longer available for the horizontal edges of the barber pole, upward motion of the vertical terminators comes to dominate.

One limitation of the current model is that it cannot account the fluctuation of an individual’s perception of motion in the barber pole illusion. As demonstrated in Chapter 3, although motion along one axis tends to dominate, motion along the other is occasionally seen. The extent of the directional bias depends on the aspect ratio of the barber pole as well as the strength of occluders when present. A likely explanation for the flipping of directions in the barber pole illusion is motion adaptation. After prolonged viewing in one direction, the motion system adapts to that direction, and the perception of other motion directions becomes more likely. A natural extension of the model, which could account for the fluctuations in motion perception, would be the addition of an adaptation mechanism. One such device for adaptation is transmitter gating (Grossberg 1980). The addition of transmitter gating to the model would allow the model to flip back and forth between the two directions as the cells coding motion in the directions alternately adapt.
A second important addition to the model would be the use of variable strength motion suppression in the presence of occluders. Chapter 3 demonstrated that the amount of suppression depends on the effectiveness of occlusion cues. Currently the model uses an all-or-none mask to implement such suppression. A more complete model would include form input that suppresses V1 motion input to varying degrees depending on the strength of occlusion cues.
Figure 5-7. The Barber Pole Illusion: (A) Horizontal Barber Pole (B) Barber Pole with disambiguating dots (C) Barber pole with occluders
5.4 Divided Diamonds

Experimental evidence suggests that the existence of reliable terminator information may in itself act to block further integration of motion signals. Lorenceau and Shiffrar (1992) examined subjects’ ability to judge the direction of rotation of a diamond shaped figure viewed through four (invisible) apertures. Subjects’ poor performance reflected their inability to integrate motion signals from lines within the four apertures.

However, when noise was added to ambiguate terminator information, subjects were able to perform the direction discrimination task. Different types of ambiguity were introduced. The apertures were made visible producing T-junctions indicative of extrinsic terminators, the terminators were viewed at large eccentricities, or were given reduced contrast or noisy lengths. All conditions produced ambiguous terminator information and resulted in significantly improved subject performance.

Shiffrar and Lorenceau (1996) examined the effect of isoluminance, stimulus duration and line width using the same paradigm. Similarly, they found that decreasing terminator information increased integration across contours. Performance on the direction discrimination task improved when isoluminance conditions were approached, line widths were thinned or presentation time decreased. These results are compatible with Lorenceau, Shiffrar, Wells & Castet (1993) who demonstrated that direction discrimination for moving lines deteriorates at low contrast and short duration.

Lorenceau and Shiffrar (1992) suggest that the motion system is biased towards local interpretations rather than more global interpretations. Unambiguous motion of terminators may inhibit the integration of motion signals across contours. However, the integra-
tion of motion across contours is facilitated when terminators are ambiguous.

These effects were examined using two sets of stimuli. The first set of stimuli were generated by placing opaque bars over a rightward translating diamond (See Figure 5-8, 1). When the opaque bars have the same contrast as the background, rendering them invisible, (2a), the individual parts of the diamond appear to break into four separately invisible, (2a), the individual parts of the diamond appear to break into four separately
moving pieces (2b). The left two pieces appear to move down and up, approaching each other. The right two pieces appear to move up and down, away from each other. However, when the same stimulus is shown with visible opaque bars, (3a), the perception is of a rigid structure moving rightward (3b).

The second set of stimuli were similar, however, rather than placing opaque bars over the ends of the triangle, four triangular patterns were used to cover the diamond’s corners (See Figure 5-9, 1). As before, when the opaque occluders are visible, (2a), coherent

1.

2a. 2b.

3a. 3b.

Stimulus Motion Perceived Motion

Figure 5-9. Diamond With Triangular Occluders: Stimuli are generated by translating a diamond pattern behind opaque triangles. Invisible portions of the diamond are shown in light grey (1). When occluders are invisible, (2a), the visible portions of the diamond are seen to move incoherently (2b). When visible occluders (3a) are employed, the visible portions are seen to move together in the true direction of motion.
rightward motion is seen, (2b). When they are invisible, (3a), the parts of the diamond are seen as separately moving pieces (3b). However, unlike the first set of stimuli, when the pieces are seen to separate, they appear to move diagonally rather than up and down. This suggests that the important factor for determining the perceived direction of motion is the orientation of the occluding patches which in turn determines the motion signal present at the line terminators.

First, let us examine the instances in which the occluders are invisible (Figure 5-8, 2a and Figure 5-9, 2a). There are two sources of disambiguating signal for these stimuli. Each piece of the diamond consists of a short line segment. Like the line segments previously discussed, motion along the edges of the lines is ambiguous. Each terminator, however, provides an unambiguous motion signal in a direction orthogonal to its (invisible) occluding edge. These signals correspond to the local signals, those within each diamond piece, referred to by Lorenceau and Shiffrar (1992).

The second source of disambiguating signals arises from a combination of signals across space. As discussed in the line-capture example, motion signals in the model are propagated across space. Eventually signals propagating from each of the four line segments reach one another. When they do, particular directions of motion are incompatible (see Figure 5-10). The competitive dynamics of the model ensure that these incompatible motion directions are suppressed while compatible motion directions are enhanced. The surviving motion serves as a second source of disambiguating signal. In terms of Lorenceau and Shiffrar (1992) these signals correspond to the direction of global motion.

Which of these two (incompatible) signals is used to disambiguate the line? The
model’s evolving activity supports Lorenceau and Shiffrar’s suggestion that unambiguous motion of terminators inhibits the integration of motion signals between spatially separated contours. Due to the aperture problem the model’s output along the edge of each line is initially in a direction orthogonal to the orientation of the line (Figure 5-11, A&C, left-most images). At the terminators, however, the output follows the direction of motion for the terminators. Like the translating line example, activity of MT$_i$ nodes along the edge of the line is suppressed, since motion along the edge is ambiguous and multiple MT$_i$ nodes are active. The activity of MT$_i$ nodes at the terminators, however, is unambiguous and thus enhanced. The model’s propagation dynamics lead to a gradual filling-in and eventually the output for each line is in the direction of its terminators (see right-most
images). As each individual line’s terminators possess different motion signals, the lines end up with different motion directions.

The results are quite different when the occluders are visible (Figure 5-8, 1a and Figure 5-9, 1a). As previously discussed, the assumption is made that V1 input to the MT cells is suppressed when occlusion information (T-junctions) are present. Therefore unambiguous motion information at the terminators is no longer present and filling-in to the ambiguous centers no longer occurs (See Figure 5-11, B & D). The second disambiguating signal which arises from the combination of motion signals across space now plays an important role. In the given examples, only rightward motion is compatible with the motion signals propagating from all four lines. This signal serves to disambiguate all four line segments. The model’s behavior is in agreement with Lorenceau and Shiffrar’s suggestion that the integration of motion across contours is facilitated when terminators are ambiguous.
A) Diamond with invisible bar occluders

B) Diamond with visible bar occluders

C) Diamond with invisible triangle occluders

D) Diamond with visible triangle occluders

Figure 5-11. Simulations of divided diamonds. Model output is shown for 6 different time snapshots. When occluders are invisible (1 & 3) the model’s output shows the diamond parts moving incoherently. When occluders are visible (2 & 4) the model signals the true direction of motion. See text for details.
5.5 Crossing Lines

The motion simulations presented thus far have involved only single objects or cases in which there were no discontinuities in the motion signals arising within objects. Although segmentation cells (MTs) were present in the model, they played an insignificant role as their activation was minimal. In this section a simulation is presented in which the activity of the segmentation cells is critical for producing correct output.

One of the most fundamental and difficult problems in motion processing exists when multiple conflicting motion signals are present in an image. A simple example of this difficult problem is what is known as ‘the cross problem’ (Weiss & Adelson, 1994). The cross-problem consists of two orthogonal lines translating past one another (Figure 5-12, 1). The importance of the segmentation system can be understood by looking at unambiguous signals within the cross display. Unlike previous examples, there are two conflicting sources of unambiguous motion information in the display. As in the previous examples, motion is ambiguous along the edges of the lines, whereas motion at the line terminators is unambiguous. Additionally, a second source of unambiguous motion signals is present at the intersection of the two lines. These two sources of unambiguous signals are incompatible. Motion at the terminators signals the true direction of line motion while motion at the intersection signals upward motion. How does the motion processing system know which source should be employed to determine the direction of motion?

A variation on the cross-problem occurs when occluding patches are added to the top and bottom of the moving lines (Figure 5-12, 2a). When occluders are present, the vertical motion of the lines is no longer perceived. Rather, the two lines appear to move
upward as a single rigid object (Figure 5-12, 2b).

Observations of motion perception suggest that when terminator information is readily available (Figure 5-12, 1) the disambiguating signal from the intersection is ignored. On the other hand when terminator motion is unreliable (Figure 5-12, 2), motion signal from the intersection is used to determine the perceived direction of motion.

How does the model behave with only integration cells? When terminator information is available, there are two sources of unambiguous motion information, at the terminators and at intersection of the lines. Both disambiguating signals attempt to fill-in the ambiguous motion areas along the edge of each line. As the two sources of disambiguating signal are not in agreement, the end result is a line whose tips move in one direction and whose center moves in another (See Figure 5-13).

Figure 5-12. Crossing Bars: Stimuli consist of two lines translating in opposite directions When occluders are absent, (1a), the visible portions of the diamond are seen to move in their veridical direction (1b). When occluders are employed, (2a), the lines are seen to move coherently upward as a single rigid object.
The difficulty with such a stimulus is that unlike those that have been previously discussed there are two separately moving objects in the image. An integration system can adequately process stimuli with single stimuli. However, as just demonstrated, a segmentation system is required when multiple, differently moving objects are present in order to prevent integration of motion signals between these separate objects.

The segmentation system functions by looking for motion discontinuities in the image. Such discontinuities are indicative of points of intersection between separately moving objects. Segmentation cells with center-surround receptive fields are activated by the presence of discontinuities. Their activity in-turn suppresses the activity of the integration cells at nearby locations, thus preventing the integration of motion signals in the presence of motion discontinuities.

For the cross stimulus, there are strong motion discontinuities at the intersection of the two lines (See Figure 5-14). $MT_s$ cells are activated by this discontinuity and subsequently suppresses $MT_i$ cell activity near the intersection. This suppression prevents unambiguous upward motion signals from the intersection from propagating outward. $MT_s$ cells taking input from propagating sub-threshold $MT_i$ cell are also activated via the
gating signal from active MT_s cells. This allows a motion border to grow outward separating the different motion areas and preventing the integration of sub-threshold MT_i cells. Consequently, unambiguous motion signals propagating inward from the line terminators come to dominate. The evolving activity of the segmentation and integration cells for the unoccluded cross stimulus is depicted in the left two columns of Figure 5-15. Segmentation cell activity is depicted in the first column. The darker the red the greater segmentation cell activity. Integration cell activity is displayed in the second column. Note that integration cell activity is suppressed at locations where segmentation cells are active. The model’s final output depicts the two lines moving in separate directions.

When occluders are present (Figure 5-15, right two columns), unambiguous motion signals from the terminators are unavailable. Now upward signals come to dominate. As intersection is disambiguated, the motion discontinuity disappears and consequently MT_s cell activity is in turn reduced. This release from suppression allows MT_i cells at the intersection to be active. The final model output is one of a rigid cross object moving upwards.
Figure 5-15. Model output for the non-occluded (right) and occluded (left) cross stimulus. Both segmentation (MT_s) and integration (MT_i) node activity is shown. See text for details.
5.6 Discontinuities Within An Object

The previous example demonstrated how motion discontinuity localization prevents spurious integration between different objects. In some cases, however, discontinuities are present within a single rigid object. For example, a translating diamond or rectangle contains motion discontinuities near each corner (Figure 5-16). The object, however, when properly disambiguated should consist of one, not four separate pieces.

Figure 5-16. Within-Object Discontinuities: Many objects, such as a translating rectangle, contain discontinuities. In the pictured example, motion discontinuities exist at each of the four corners of the rectangle (discontinuities are illustrated only at two of the corners).

Figure 5-17 demonstrates how the model handles such cases. Initially motion along all four sides of the square is ambiguous (Figure 5-17, A). Discontinuities at the corners activate the segmentation cells (Figure 5-17, B). Segmentation borders begin to grow inward as sub-threshold activation is propagated inward. Unambiguous motion signals from the lower right hand corner begin to disambiguate that corner (Figure 5-17, C). The
lower right hand corner disambiguates first as the motion of the rectangle is up and to the right. Gradually all corners begin to disambiguate (Figure 5-17, D). The disambiguation causes segmentation cell activity in the corners to collapse (Figure 5-17, E-G). The sides are disambiguated and the motion border continues to collapse. Finally, all motion borders have collapsed and the entire square is disambiguated (Figure 5-17, J).

<table>
<thead>
<tr>
<th></th>
<th>MT_s</th>
<th>MT_i Sub-Threshold</th>
<th>MT_i Threshold</th>
<th>MT_s</th>
<th>MT_i Sub-Threshold</th>
<th>MT_i Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.</td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
</tr>
<tr>
<td>B.</td>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
<td><img src="image9.png" alt="Image" /></td>
<td><img src="image10.png" alt="Image" /></td>
<td><img src="image11.png" alt="Image" /></td>
<td><img src="image12.png" alt="Image" /></td>
</tr>
<tr>
<td>C.</td>
<td><img src="image13.png" alt="Image" /></td>
<td><img src="image14.png" alt="Image" /></td>
<td><img src="image15.png" alt="Image" /></td>
<td><img src="image16.png" alt="Image" /></td>
<td><img src="image17.png" alt="Image" /></td>
<td><img src="image18.png" alt="Image" /></td>
</tr>
<tr>
<td>D.</td>
<td><img src="image19.png" alt="Image" /></td>
<td><img src="image20.png" alt="Image" /></td>
<td><img src="image21.png" alt="Image" /></td>
<td><img src="image22.png" alt="Image" /></td>
<td><img src="image23.png" alt="Image" /></td>
<td><img src="image24.png" alt="Image" /></td>
</tr>
<tr>
<td>E.</td>
<td><img src="image25.png" alt="Image" /></td>
<td><img src="image26.png" alt="Image" /></td>
<td><img src="image27.png" alt="Image" /></td>
<td><img src="image28.png" alt="Image" /></td>
<td><img src="image29.png" alt="Image" /></td>
<td><img src="image30.png" alt="Image" /></td>
</tr>
</tbody>
</table>

Figure 5-17. Translating Square: Evolving segmentation and integration cell activity (both threshold and sub-threshold) are shown for a rectangle translating up and to the right.
5.7 The Diamond Problem: What to integrate?

One challenge that has been posed for motion processing models is the correct disambiguation of two passing diamonds without knowledge of the underlying diamond structure. Although the motion of a diamond’s edges is ambiguous due to the aperture problem, the true direction of motion can be determined by integrating motion signals between edges. When only one diamond is present the method succeeds. However, when more than one diamond is present, the question arises as to which motion signals should be integrated.

The problem is depicted in Figure 5-18. When an upward and downward moving diamond pass each other there are various ways the motion signals can be combined. If the
left and right motion pairs are combined, a correct disambiguation results. On the other hand combining the top and bottom motion pairs results in spurious motion directions.

The model demonstrates that the passing diamond problem is remedied when terminator motion and motion discontinuities are taken into account. Both computational processes, segmentation and integration, as well as their interactions are important for this example. Figure 5-19 depicts how the model resolves the passing diamonds stimulus. As

---

**Figure 5-19. Two Passing Diamonds: Evolving network activity for the segmentation and integration cells is shown for two passing diamonds. See text for details.**
illustrated in the previous square example, initially segmentation cells are activated by motion discontinuities in the corners of each object (Figure 5-19, A). The discontinuity between the diamonds also begins to activate other segmentation cells between the two diamonds (B-C). As in the single rectangle example, the corners are disambiguated. The disambiguation results in a collapse of segmentation cell activity as motion discontinuities in the corners fade (D). On the other hand, segmentation activity along the border between the two diamonds continues to grow as no disambiguating signal is present (E). Finally both diamonds are disambiguated. The remaining active segmentation cells divide the space into two separate regions, one for each diamond (G).

5.8 Multiple Object Displays

This section demonstrates how the model can cope with multiple overlapping objects. The first example consists of two approaching rectangles (Figure 5-20, left). The initial motion measurement (V1) for each line-edge pair is distinct and ambiguous. Furthermore, no motion signals exist in the interior of the two objects as their surfaces are uniform (Figure 5-20, right).

Figure 5-20. Two approaching rectangles. Input is shown on the left, V1 output on the right.
Correctly disambiguating the motion of these two objects is critically dependent on two mechanisms. First, spurious motion signals from V1 near the intersection of the two objects must be suppressed. This was accomplished as before by constructing a mask at the location of T-junctions in the image. Secondly, the segmentation system must suppress MT₁ activity in the presence of motion discontinuities. Figure 5-21 shows the model output when either the mask or the segmentation system were turned off.

Figure 5-23 shows how the full model is able to segment the motion of the two rectangles. Initially (snapshots A & B) segmentation cells are activated by discontinuities at the corners of the rectangles and at the intersections between the rectangles. As sub-threshold motion signals are propagated across the image, motion borders grow inward to the center of and between the two rectangles (snapshots C-E). When the corners of the rectangles are disambiguated (snapshot F), the motion discontinuity erodes, and segmentation cells at the corners are deactivated. Motion borders within the rectangles are no longer supported and begin to degrade (snapshots G-I). By the final snapshot, both rectangles have been completely disambiguated. Note that the remaining segmentation border divides the space
into two regions containing the two different object motions (J).

The second example consists of three objects moving in different directions (See Figure 5-22). Figure 5-24 depicts how the model disambiguates and separates the motion of the three objects. By the final snapshot, model output shows all objects moving in their correct direction. Segmentation cell activity divides the space into three regions corresponding to the three objects.

Figure 5-22. Three Objects. Input stimulus.
Figure 5-23. Ten snapshots of MT_s and MT_i cell output for two moving rectangles. See text for details.
Figure 5-24. Ten snapshots of $MT_s$ and $MT_i$ cell output for three moving objects. See text for details.
5.9 Comparison to Chey, Grossberg & Mingolla (1997, 1998)

As previously mentioned, the model of Chey, Grossberg & Mingolla (1997, 1998) is perhaps the most developed of all the biologically motivated models of motion perception. The model not only extracts initial motion measurements, accounts for much of the psychophysical data on speed tuning and data on the coherence and incoherence of plaid patterns, but also does so in a biologically plausible fashion. The motion of multiple objects, however, is beyond the scope of the current version of the model. Although the Chey et. al model examines coherent and incoherent plaids, its output for incoherent plaids consists of a continuum of motion directions at the intersection of the plaid components. This is because the model does not contain any method for segmentation.

Recent work suggests that the Chey et al. model may be extended through the addition of form-to-motion interactions which act to segment an image (Francis & Grossberg, 1996; Baloch & Grossberg 1997). With form-to-motion interactions the Chey et al. model may be extended to account for some of the multiple-object simulations simulated by the model discussed in this chapter. It should be noted, however, that these two models take fundamentally different approaches to the generation of image segmentations. The Chey et al. extension would presumably use form/edge information to break an image into its constituent parts. The segmentation information generated in the form system would then be fed to the motion system where it would acts to influence how motion information in combined.

The model discussed in this chapter, on the other hand, does not generate segmentations within the form system. Apart from a general suppression of motion signals in the
presence of T-junctions, the motion model knows nothing about edges, their orientation, contrast, or any other aspect of form. Rather, all processing occurs on the raw motion signals. Segmentations are generated by the model by looking for motion discontinuities within the raw motion signals.

In biology it is likely that both form-to-motion segmentation and within-motion segmentation processes play a role in motion perception. It is a useful exercise, however, to see how much can be accomplished by one system alone. The model presented herein attempted to see how much the motion system can accomplish in isolation. The simulations presented in Chapter 5 clearly demonstrate that the motion system is capable of processing a wide range of motion stimuli with the minimal of form-to-motion interactions. The addition of form-to-motion actions such as proposed by Baloch and Grossberg (1997) and Francis and Grossberg (1996), when combined with the 3-D figure-ground separation mechanisms described in Grossberg (1994) and McLoughlin and Grossberg (1997) would help to further improve the model’s output. In particular, although segmentation is achieved in the crossing lines example (section 5.5), form-to-motion input is probably necessary to account for the perception of motion within the intersection.

5.10 Summary

This chapter demonstrated how a motion processing model with two processing sub-systems, one for integration and one for segmentation, can interpret a wide range of motion stimuli. The model produces output compatible to the perceived motion of human subjects for a variety of motion stimuli such as the barber pole illusion, motion capture
and crossing lines. Although the segmentation system does not play an important role for stimuli involving single objects, both the segmentation and integration systems are required for stimuli involving multiple objects with different trajectories. The integration system serves to overcome the aperture problem and noise in the image. The segmentation system prevents the integration of motion signals between objects. Despite seemingly contradictory roles, the model demonstrates how they complement each other to produce a unified interpretation of motion stimuli.
Appendix A

Terminator Attenuation

A.1 Calculating the Attenuation of Terminators

When the perception of horizontal and vertical motion is equally likely, the influence of the horizontal and vertical terminators is presumably equal. If $x$ is the length of the occluded side, $y$ is the length of the non-occluded side and $a$ is the attenuation of terminator influence due to the presence of the occluders, then terminator influence is in balanced when:

$$a2x = 2y$$  \hspace{1cm} (EQ 2)

The length of the non-occluded side, $y$, was held constant in all experiments. The value of $x$ can be calculated by finding the 50% level on the sigmoid which best fits the subject data. The attenuation of the occluded sides is then found by solving for $a$:

$$a = \frac{y}{x}$$  \hspace{1cm} (EQ 3)

The percent reduction of terminator influence is then given by $(1-a)$.

A.2 Predicting the Attenuation for One Side

When one side of the barber pole is occluded, only one set of terminators along the $x$-
dimension is attenuated, so a balance between horizontal and vertical motion presumably occurs when:

$$ax + x = 2y$$  \hspace{1cm} (EQ 4)

The attenuation of the occluded terminators is then found by solving for $a$:

$$a = \frac{2y}{x} - 1$$  \hspace{1cm} (EQ 5)

Using the attenuation level calculated for the occlusion of both sides and the fact that length of the unoccluded side, $y$, is held constant in all experiments, one can also make a prediction for the length of the occluded side, $x$, that will result in a balance in the perception of vertical and horizontal motion:

$$x = \frac{2y}{a + 1}$$  \hspace{1cm} (EQ 6)
Appendix B

Simulation Details

The following appendix describes the equations used to implement the model discussed in Chapter 4 and Chapter 5. The first three sections describe the calculations used for the input nodes (V1), the integration nodes, (MT_i) and the segmentation nodes (MT_s) respectively. For convenience, a summary of all variables and their associated meanings is provided in Appendix B.4. The model was instantiated in the C language and all simulations were run on a Silicon Graphics workstation running the IRIX operating system. Equations were numerically integrated using floating point values.

In the following set of equations, position is represented by superscripts whereas directional preference is represented by subscripts.

B.1 Input Nodes (V1)

The activity of input nodes is calculated using a weighted correlational scheme. For each spatial location there are eight V1 cells, one for each direction of motion. For each position in space, the activity of V1 cells is calculated by comparing a weighted window of grey-level pixel values around that position to eight different shifts of the same window at later points in time. The activity of a given V1 cell, \( v \) with position subscripts, \( x, y \), and directional preference superscript, \( n \), is given by:
\[ y^n_{x,y} = \left(1 - \sum_{(i,j) \in K} k_{i,j} I(t)(x+i, y+j) - I(t+1)(x+i + \gamma^n_i, y+j + \gamma^n_j)\right) + \left(1 - \sum_{(i,j) \in K} k_{i,j} I(t)(x+i, y+j) - I(t+2)(x+i + \gamma^n_i, y+j + \gamma^n_j)\right) \] (EQ 7)

\[ I(t)_{x,y} \] corresponds to the grey-level pixel intensity in the input image at position \( x, y \) and time \( t \). \( K \) is the spatial extent of the correlation kernel. \( \gamma^n_i, \gamma^n_j \) are the x and y positional shifts for each point within the kernel for direction \( n \). \( k_{i,j} \) is the roughly gaussian weight matrix corresponding to the correlation kernel. The the weights were normalized so:

\[ \sum_{(i,j) \in K} k_{i,j} = 1 \] (EQ 8)

When perfect correlation occurs for a shift in a specific direction, the grey-scale values of one window matches the grey-scale values of the window shifted in that direction at a later time step. The summation value is therefore zero. On the other hand when perfect anti-correlation occurs, the weighted sum is one, since the weight matrix is normalized. By subtracting the summation from a value of one, the result is assured to be one when perfect correlation occurs and zero when perfect anti-correlation occurs.

Two points in time are used to capture the fact that the multiple speeds may be present in the input stimulus. A more complete model would use multiple sets of V1 cells each tuned to different speeds rather than the simple summation employed here.
B.2 Integration Nodes (MT_i)

The dynamics of the integration system are implemented by a single equation which incorporates all three computational principles: inhibition within spatial location, $\Gamma$, short range excitation, $\Lambda$ and long range inhibition, $\Omega$, as well as inhibition from the segmentation nodes, $\chi$ (See Figure 4-1). The activity of an integration node, with position subscripts, $x$, $y$ and direction selectivity superscript $n$ is given by:

$$\frac{d}{dt}m^n_{xy} = \left(1 - m^n_{xy}\right)\left(Bv^n_{xy} + E\Lambda^n_{xy}\right) - \left(m^n_{xy}\right)\left(Dm + C\Gamma^n_{xy} + A\Omega^n_{xy} + F\chi_{xy}\right)$$  \hspace{1cm} (EQ 9)

The equation consists of excitatory and inhibitory terms gated by the node’s own activity. This keeps the activation of each integration node in the range $[0,1]$. Input from V1, is denoted $\nu$ and weighted by parameter $B$. Each node also has a decay rate $D_m$. The three computational processes, $\Gamma$, $\Lambda$ and $\Omega$ are weighted by parameters $E$, $C$ and $A$ respectively and are given by:

1. **Inhibition within spatial location across direction**, $\Gamma$, with an inter-directional connectivity matrix $\alpha$ where $k$ and $n$ are direction indices.

$$\Gamma^n_{xy} = \sum_{k=1}^{8} m^k_{xy}\alpha^{kn}$$  \hspace{1cm} (EQ 10)

2. **Short range excitation**, $\Lambda$, with an inter-directional connectivity matrix $\mu_1$ and an inter-spatial connectivity matrix $\mu_2$ with direction indices $k$ and $n$, and positional indices $x, y, i$ and $j$. 
(3) Long range inhibition, \( \Omega \), With an inter-directional connectivity matrix \( \phi_1 \) and an inter-spatial connectivity matrix \( \phi_2 \).

\[
\Omega_{xy}^n = \sum_{(i, j) \in \Omega} \left( \sum_{k=1}^{8} m_{ij}^{k} \phi_1^{kn} \right) \phi_2^{ijxy}
\]  

(EQ 12)

All inter-directional connectivity matrices, \( \alpha, \mu_1 \) and \( \phi_1 \) are roughly Gaussian in shape. See Appendix B.5 for details.

Inhibition by the segmentation nodes, \( \chi \) is weighted by parameter \( F \) and is simply the sum of all segmentation node activity at that position given by:

\[
\chi = \sum_{n=1}^{8} s_{xy}^n,
\]

(EQ 13)

where \( s \) is a segmentation nodes at position \( x, y \) with directional tuning \( n \).

B.3 Segmentation Nodes (MTs)

For each position in space there is a set of eight segmentation nodes, \( s \) with position \( i, j \) and directional selectivity \( n \). \( \Psi \), the lateral excitation from other segmentation nodes, is defined as:
\[ \Psi_{xy}^n = \sum_{(i,j) \in \psi} s_{ij}^n, \]  
(EQ 14)

where \( \psi \) is the spatial extent of lateral excitation.

\( \chi \), the input from V1, is defined as:

\[ \chi_{xy} = \sum_{n=1}^{8} V_{xy}^n, \]  
(EQ 15)

where \( V_{xy}^n \) is the V1 input at spatial position \( x, y \) for direction \( n \). The summation results in directionally non-specific input.

Finally, the surround inhibition, \( \mathcal{S} \), is defined as:

\[ \mathcal{S}_{xy}^n = \sum_{(i,j) \in \epsilon} m_{ij}^n, \]  
(EQ 16)

where \( \epsilon \) is the spatial extent of the surround inhibition. In this way a segmentation cell is inhibited when integration cells in the surround with the same directional tuning are active. In other words, the segmentation cell is suppressed in the presence of whole field movements.

The activity of a segmentation node \( s \) with directional preference \( n \) at position \( x, y \) is given by:

\[ \frac{d}{dt} s_{xy}^n = \left( 1 - s_{xy}^n \right) \left( G\Theta_{xy}^n \right) \left( s_{xy}^n \right) \left( D_s + H\mathcal{S}_{xy}^n \right), \]  
(EQ 17)

where \( D_s \) is the decay rate. \( \Theta \) is the gated excitation of the segmentation nodes weighted by parameter \( G \) and \( \mathcal{S} \) is the inhibition of the segmentation nodes weighted by
parameter $H$.

Segmentation cell's input is gated (See Figure 4-2). In other words, its input is only effective if the cell also receives lateral excitation from either segmentation nodes or input from V1. The gated input, $\Theta$, is given by:

$$\Theta_{xy}^n = \begin{cases} 
\Upsilon_{xy}^n & \text{if } \Psi_{xy}^n > 0 \text{ or } \chi > 0 \\
0 & \text{otherwise}
\end{cases}$$

(EQ 18)

where $\Upsilon$ is the total excitatory input, $\Psi$ is the lateral excitation and $\chi$ is the input from V1. The total excitatory activity, $\Upsilon$, is given by:

$$\Upsilon_{xy}^n = \sum_{(i, j) \in \theta_c} \left[ m_{ij}^n - J \sum_{d \neq n} m_{ij}^d \right]^+ + \sum_{(i, j) \in \theta_s} m_{ij}^n$$

(EQ 19)

The first term summates over the kernel center, $\theta_c$. Integration cells, $m$, within the center that possess the same directional preferences, $n$, are excitatory. Those possessing all other directional preferences, $d \neq n$, are inhibitory and weighted by a parameter, $J$. $[ ]^+$ refers to a half-wave rectification that ensures that the sum is always positive. In this way the segmentation cells receive more excitement from disambiguated areas of visual space (i.e. locations in which fewer directions are active at the integration cell level). The second term summates over the receptive field surround, $\theta_s$. In the surround, integration cells, $m$, of the opposite directional preference, $\neg n$, are excitatory.

The conditional statement, equation 18, ensures that the excitatory input, $\Upsilon$, is only effective when either lateral excitation, $\Psi$ or input from V1, $\chi$ is also present.
B.4 Summary of Variables

Network Components

\( m \) An integration node (MT\(_i\))

\( s \) A segmentation node (MT\(_s\))

\( v \) An input node (V1)

\( I \) A pixel intensity (input image)

Indices

\( n, d \) - Directional preference

\( t \) - Time

\( x, i \) - Horizontal spatial position

\( y, j \) - Vertical spatial position

Equation Variables

\( \Lambda \) Short range excitation for integration cells

\( \Gamma \) Inhibition within location for integration cells

\( \Omega \) Long range inhibition for integration cells

\( \Theta \) Gated center-surround excitation for segmentation cells

\( \Upsilon \) Center-surround excitation for segmentation cells

\( \Upsilon \) Center-surround inhibition for segmentation cells

\( \Psi \) Lateral segmentation cell input for segmentation cells

Spatial Dimensions
\(\gamma_i^n\) Correlation shift in \(x\) component for direction preference \(n\) for input cells

\(\gamma_j^n\) Correlation shift in \(y\) component for direction preference \(n\) for input cells

\(K\) Spatial extent of correlation kernel of input cells

\(k\) Weight of a position within the correlation kernel of input cells

\(\alpha\) Connectivity matrix for inhibition within location \((\Gamma)\) of integration cells

\(\mu_1\) Inter-directional connect matrix for short range excitation \((\Lambda)\) of integration cells

\(\mu_2\) Inter-spatial connectivity matrix for short range excitation \((\Lambda)\) of integration cells

\(\phi_1\) Inter-directional connect matrix for long range inhibition \((\Omega)\) of integration cells

\(\phi_2\) Inter-spatial connectivity matrix for long range inhibition \((\Omega)\) of integration cells

\(\lambda\) Spatial extent of short range excitation \((\Lambda)\) for integration cells

\(\omega\) Spatial extent of long range inhibition \((\Omega)\) for integration cells

\(\psi\) Spatial extent of lateral excitation for segmentation cells

\(\theta_c\) Spatial extent of center for center-surround input for segmentation cells

\(\theta_s\) Spatial extent of surround for center-surround input for segmentation cells

\(\epsilon\) Spatial extent of surround inhibition for segmentation cells

Adjustable Parameters

A Weight parameter for long range inhibition \((\Omega)\) of integration cells

B Weight parameter for V1 input \((\nu)\) of integration cells

C Weight parameter for inhibition within spatial location \((\Gamma)\) of integration cells

E Weight parameter for short range excitation \((\Lambda)\) of integration cells

F Weight parameter for segmentation cell inhibition \((\chi)\) of integration cells
Decay rate parameter for integration cells

Decay rate parameter for segmentation cells

Weight parameter for center surround excitation (Θ) of segmentation cells

Weight parameter for center surround inhibition (Ξ) of segmentation cells

Weight parameter of center surround center inhibition of segmentation cells

B.5 Summary of Parameters

\[ A = 1.0, \ B = 0.8, \ C = 6, \ D_m = 0.5, \ D_s = 0.6, \ E = 0.7, \ F = 0.5, \ G = 1.7, \]

\[ H = 1.2, \ J = 0.4, \alpha, \phi_1 = \begin{cases} 
0.9^{n,n-3}, 0.5^{n,n-2}, 0.5^{n,n-1}, 0.0^{n,n} \\
0.5^{n,n+1}, 0.5^{n,n+2}, 0.9^{n,n+3}, 1.0^{n,n+4}
\end{cases} \]

As the kernel, \( \lambda \) consisted of only the most adjacent surrounding units, the inter-spatial connectivity matrix, \( \mu_2 \) was simply 1.0 for all values. Were larger spatial scales employed the connectivity matrix would need a gaussian fall off as distance increases. The kernel, \( \omega \), consisted of surrounding units of a distance of three units or less. \( \phi_2 \), the inter-spatial connectivity matrix was 0.5 for the most adjacent spatial position, 1.0 for an offset of one, and 0.5 for an offset of two spatial positions.
References


Vita

Lars Lidén received a Bachelor of Science in Cognitive Science from the University of California, San Diego in 1994, graduating with department honors with highest distinction. In 1994 he was awarded the Computational Science Graduate Fellowship from the Department of Energy and for the past four years he has been pursuing graduate study at Boston University in the Department of Cognitive and Neural Systems supported by this fellowship.